Make, Buy, or Share: The Role of Interfirm Trade in Fracking Wastewater Reuse

Matthew O'Keefe*

December 30, 2024

Abstract

Wastewater reuse in the shale gas industry reduces firms' private costs and mitigates many of the local environmental harms associated with fracking. Most reuse occurs within the firm boundary, but rival operators often exchange (or "share") wastewater prior to reuse. I study the effect of wastewater sharing on reuse rates, transportation efficiency, and other outcomes. I find substantial private benefits from wastewater sharing, and modest external benefits. However, these benefits are diminished by the presence of large transaction costs. I explore the sources of these transaction costs and consider potential policy interventions to improve sharing markets and incentivize reuse.

Keywords: oil and gas extraction; produced water; firm boundaries; transaction costs; transferrable utility matching; empirical matching models

*Vanderbilt University. Email: matthew.okeefe@vanderbilt.edu. I am grateful for the generous support and guidance of Rob Porter, Mar Reguant, and Vivek Bhattacharya. I also thank Michael Powell, Gaston Illanes, Bill Rogerson, Tom Hubbard, Igal Hendel, Meghan Busse, Ryan Kellogg, Alex MacKay, Ameet Morjaria, Gaston Lopez, Francisco Pareschi, and numerous conference and seminar participants for helpful insights. Several individuals at the Pennsylvania Department of Environmental Protection and Susquehanna River Basin Commission provided invaluable background on wastewater management. An earlier version of this paper circulated under the title "Firm Boundaries and External Costs in Shale Gas Production" and appeared in the author's PhD Thesis at Northwestern University.

1 Introduction

Oil and gas wells produce large volumes of toxic wastewater. Prior to the shale boom in the United States, most oil and gas producers relied on a final disposal technology for wastewater known as injection well disposal. Due to regulatory pressure and technological developments, producers have increasingly turned towards alternative strategies for wastewater management. One such strategy is *reuse*, in which wastewater from one well is repurposed to complete new wells. In Pennsylvania, the second largest gas producing state, nearly 90% of wastewater is reused. In other oil and gas-producing regions, however, reuse remains relatively uncommon. Across all US shale basins, only 10-15% of wastewater is reused.[1]

Given the toxicity of wastewater and the enormous volumes of it that are produced by even a single well, positive externalities from reuse are potentially large. An understanding of the institutions that facilitate high rates of reuse in Pennsylvania is therefore useful when considering how to incentivize greater reuse in other regions. This paper contributes to this understanding by examining the role of trade in wastewater, known as *wastewater sharing*, in mediating reuse. Onshore oil and gas producers vary widely in size and sophistication (Small et al., 2014; Boomhower, 2019). In this context, wastewater sharing can significantly reduce the total costs of reuse, resulting in more extensive reuse and improved efficiency. A key question for policy is the extent to which reuse in Pennsylvania is attributable to sharing versus other factors, such as exceptionally high injection disposal costs.

Using administrative disposal records, I first show that approximately 10% of wastewater leaving unconventional gas wells in Pennsylvania is transferred to facilities associated with rival firms. This indicates that while sharing is not a rare occurrence, most reuse occurs within the firm boundary. At the same time, I find evidence that sharing is subject to substantial transaction costs. The extent to which observed sharing rates reflect modest gains from trade versus high transaction costs is therefore uncertain, making it difficult to assess the value of sharing. Moreover, the existence of transaction costs potentially has important

[1] Calculation based on figures reported in Groundwater Protection Council (2019).

implications for the magnitude of local environmental harms related to wastewater. Even if aggregate reuse rates are high, transaction costs could result in excessively long wastewater shipments and extended storage durations prior to reuse.

To better assess the value of wastewater sharing in Pennsylvania, I develop an empirical model of wastewater management in which the benefits of trade are balanced against Coasean transaction costs (Coase, 1937) and the costs of final disposal. To capture these tradeoffs, I adopt a transferrable utility (TU) matching framework (Choo and Siow, 2006; Galichon and Salanie, 2022). Wastewater production from a well declines over time. In the model, wastewater from many older wells is matched to a small number of ongoing completions that consume large volumes of water. Importantly, the matching framework allows me to rationalize the allocation of resources both within and across firms, as well as the "make-vs-buy" decisions that firms make when choosing to participate in the sharing market.

To estimate the model, I exploit the close connection between TU matching models and gravity models in the international trade literature. The model estimates imply that the elasticity of reuse with respect to distance is close to -1, similar to many prior estimates of the elasticity of trade in international as well as domestic settings (Head and Meyer, 2014; Atalay et al., 2019). At the same time, I find that transaction costs are economically large—equivalent in magnitude to the cost of shipping a truckload of wastewater halfway across the state, several times the mean shipment distance. Variation in the estimates suggests that transaction costs might arise from a variety of mechanisms, such as liability aversion, a desire to protect trade secrets, or incentive problems within the firm.

Counterfactual simulations imply that under present conditions sharing decreases final disposal volumes by 37% and reduces wastewater-hauling tanker-truck miles by 19%. These figures correspond to large private cost savings, particularly for smaller regional firms who would otherwise turn to injection disposal at much higher rates. Moreover, reduced injection disposal and wastewater transportation generate important external benefits, although I find that the magnitude of these benefits is modest. Nevertheless, the benefits of sharing are significantly diminished by the presence of transaction costs. In the absence of transaction costs, final disposal volumes would fall by a further 62 percentage points, and tanker-truck miles by as much as 27 percentage points. Thus, policies aimed at mitigating transaction costs could increase producer surplus while potentially generating positive environmental spillovers.

One implication of the analysis is that regulators seeking to encourage reuse should minimize barriers to sharing. However, the benefits of such efforts are largest in settings in which the cost of reuse is already comparatively low. Although sharing can be an important complement to "internal" reuse, taxes on final disposal or freshwater withdrawals are likely to be more effective instruments for incentivizing reuse outside Pennsylvania. Even in Pennsylvania, the benefits of efforts to mitigate transaction costs may only be modest, particularly if factors orthogonal to shipping distance are important drivers of matching patterns.

The remainder of the paper is structured as follows. Section 2 introduces the data and provides descriptive evidence on reuse, sharing rates, and the existence of transaction costs. Section 3 develops the empirical model. Section 5 presents the main parameters estimates. Section 6 presents the counterfactuals. Section 7 concludes.

1.1 Related literature

This work contributes most directly to the policy literature on the local environmental impacts of fracking. In economics, Hausman and Kellogg (2015) and Black et al. (2021) survey the local environmental impacts of the shale boom, among broader considerations.[2] The potential impacts of wastewater specifically and the applicable legal and regulatory frameworks are described in EPA (2016) and Groundwater Protection Council (2019).

I find that a key force governing the extent of trade is the presence of Coasean transaction costs.[3] Such costs have been studied extensively from a theoretical perspective in transaction

[2] An important but distinct issue is whether the shale boom has increased or decreased global greenhouse gas emissions. See, e.g., Newell and Raimi (2014). For simplicity I do not consider the elasticity of drilling with respect to wastewater management costs, although this is an interesting avenue for future research.

[3] "Coasean" transaction costs can be understood as the sum of all costs incurred when a transaction occurs

cost economics (Williamson, 1971), property rights theory (Grossman and Hart, 1986), and elsewhere. I build on the existing transaction costs literature by embedding firms' "make-vs-buy" decisions in a market-level structural model. There have been relatively few direct quantifications of Coasean transaction costs, in large part because identification generally requires data on firms' internal operations which is rarely available. For instance, Masten et al. (1991) quantified transaction costs using a firm's estimates of its own internal costs of production. Instead, my approach exploits on the availability of detailed spatial data. Atalay et al. (2019) adopted an approach based on spatial variation to quantify the "net benefits of ownership" using Census data. I build on their work by allowing for greater heterogeneity in transaction costs across transactions, which is possible due to the specificity of the setting. Similar approaches are used to estimate trade costs in the trade literature (Anderson and van Wincoop, 2004; Head and Meyer, 2014).

The Coasean view of the firm is complementary to the typical strategic view of the firm in industrial organization (Bresnahan and Levin, 2012). In this way, the analysis in this paper complements recent empirical work at the intersection of industrial organization and environmental economics that highlights the challenges of environmental regulation when firms act strategically (e.g., Mansur, 2007; Fowlie, 2009; Leslie, 2018; Preonas, 2023).

Finally, this paper relates to a variety of recent papers in empirical industrial organization that study interaction between oil and gas producers in the wake of the shale boom (e.g., Kellogg, 2011; Covert, 2015; Covert and Sweeney, 2022). None of these studies explores the issue of wastewater management specifically.

2 Background

In this section, I provide a description of wastewater reuse and sharing in Pennsylvania.

through exchange between firms rather than within a firm. This notion of transaction costs encompasses any explicit costs of market transactions (for example, taxes), but is potentially far more general.

2.1 Data sources

The primary data are derived from monthly disposal records that oil and gas producers file with the Pennsylvania Department of Environmental Protection (DEP). These records indicate the disposal method and destination of all quantities of waste materials leaving each well pad, including every barrel of wastewater. The reports clearly indicate whether a transfer was intended for reuse. Beginning in 2017, further information is provided identifying the destination facility. I therefore focus on the period from 2017 to 2020. I supplement this data with freshwater consumption data provided by the Susquehanna River Basin Commission (SRBC) and completions data from the FracFocus database. I calculate over-the-road shipment distances using the Open Source Routing Machine (Luxen and Vetter, 2011) and data from OpenStreetMaps.[4] The data are described in greater detail in Appendix A.

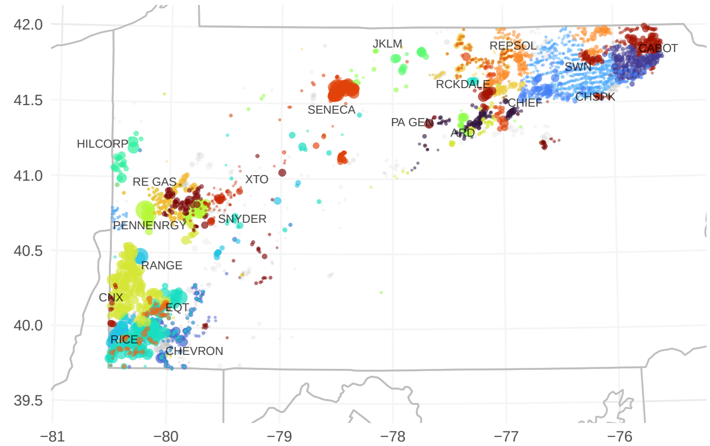
2.2 Reuse vs. injection disposal

Oil and gas extraction in Pennsylvania is conducted by numerous firms ranging from small, independent firms operating only a few wells to the largest global energy firms (Small et al., 2014). Figure 1 shows the locations of well pads operated by each of the twenty largest operators (by disposal volume) in the period that I study. The clustering visible in the figure reflects economies of density in permitting, exploration, drilling, and marketing, as well as in freshwater and wastewater management, which I discuss in this section.

The process of fracking is water intensive. A typical completion requires more than five million gallons of water (over a hundred thousand barrels), and this demand has only increased in recent years as wells have grown longer. During the fracking process, water is blended with sand and various chemicals and injected into a well under pressure. After completion, a large proportion of this fluid returns to the surface as wastewater (also known as flowback or produced water), having been mixed further with minerals and groundwater

[4] I do not account for roadway-specific vehicle weight or hazardous material restrictions that could alter optimal shipment routes for wastewater-hauling trucks in comparison to passenger vehicles.

Figure 1: Well pad locations for the twenty largest firms



in the shale. Wastewater production continues for the life of a well, in steadily diminishing volumes. Much like with hydrocarbons, the amount of wastewater that a given well will produce is difficult to predict, but typically amounts to around 50% of injected volume over the lifetime of a Pennsylvania well.[5]

The two primary methods of managing wastewater in Pennsylvania specifically and in Appalachia generally are *injection disposal* and *reuse*.[6] Wastewater is highly saline and may contain organic compounds, metals, and naturally occurring radioactive materials. Federal regulations require careful handling and specialized disposal (Groundwater Protection Council, 2019). Injection disposal involves using specialized wells to inject wastewater deep below the surface of the earth. Reuse involves transporting wastewater to a new well site, performing some basic treatment such as filtering and pH reduction, and then using this treated water in place of freshwater when completing new wells. In either case, the vast majority of wastewater is transported away from the well pad by tanker truck.[7]

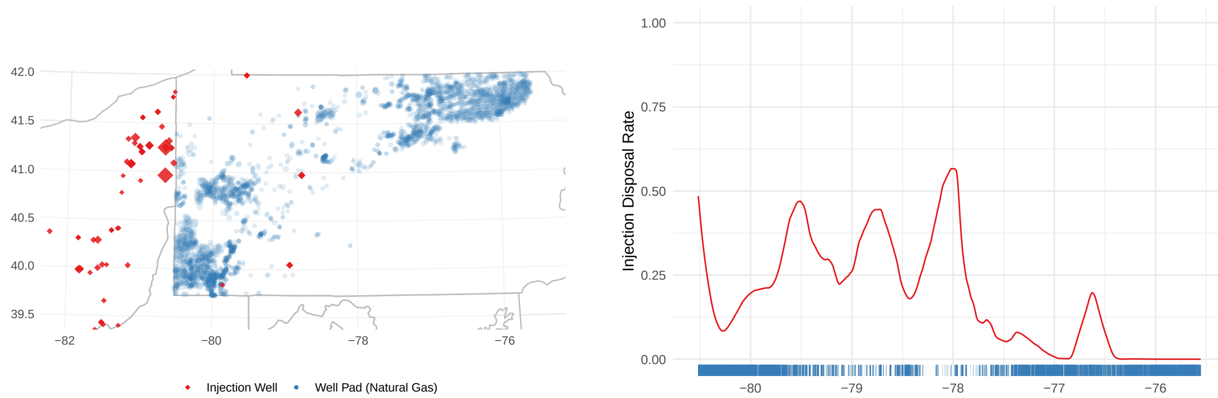
[5]The Marcellus and Utica shales (the main formations underlying Appalachia) are considered "dry" in the sense that relatively little water returns to the surface. In other regions, wastewater generation can be an order of magnitude larger (Kondash et al., 2018), substantially changing the economics of reuse.

[6]Another less commonly used option is desalination. In the western United States, evaporation and agricultural applications are also used (Groundwater Protection Council, 2019). Note that reuse outside the oil and gas industry is extremely limited. This primarily reflects a substantial difference in treatment requirements for reuse in fracking and reuse in other applications, as well as transportation costs.

[7]Approximately 5% of produced water is transported by pipeline (Groundwater Protection Council, 2019). According to the DEP, some operators have also used rail transportation.

Figure 2: Injection well locations and usage

(a) Locations (b) Mean Disposal Rate



In Pennsylvania and West Virginia the underlying geology is not well suited to drilling injection wells (McCurdy, 2011). Injection wells are common in nearby Ohio, but the distance between Pennsylvania gas wells and Ohio injection wells can be significant. This is illustrated in Figure 2a, which shows the location of injection wells relative to gas wells recorded in the data. Given Pennsylvania's rectangular shape, a gas well's longitude represents a crude measure of the accessibility of Ohio injection wells. Figure 2b plots a local linear regression estimate of the expected injection disposal rate per well pad-month conditional on longitude. Injection disposal rates are highest immediately next to the Ohio border, as well as in the center of the state where some disposal facilities exist and drilling activity is less concentrated. The mean injection disposal rate was 12.3%; at longitudes corresponding to the dense region of drilling in the northeast part of the state, the mean rate was less than 1%.

2.2.1 Centralized treatment facilities

Treatment prior to reuse can occur either directly on a well pad or at a centralized treatment facility (CTF). Treatment on a well pad is more common than treatment at a CTF, but both are prevalent (I provide market shares below). Some CTFs are operated by oil and gas producers, while others are operated by third party treatment firms. Producer-affiliated

CTFs are often little more than semi-permanent systems of tanks or impoundments where the same treatments conducted on a well pad can be conducted at a larger scale. Third party CTFs are constructed similarly but may also have technologies that can treat water to higher standards, although these technologies are rarely used in practice.[8]

2.2.2 Limitations of the data

The data have a few limitations worth highlighting. Only the total volume of water transferred between two locations during a month is recorded, rather than the dates, modes, volumes, or circumstances of particular shipments. In particular, the data do not include prices, contract terms, or other details.[9] In the case that wastewater is reused, the data do not indicate locations at which treatment processes occurred, or if these occurred in different stages at different locations. In the case that wastewater is initially transferred to a centralized treatment facility (CTF), the ultimate location of reuse is not indicated.[10]

2.3 Wastewater sharing

Wastewater is typically reused by the firm that generated it. However, in some circumstances, firms may trade ("share") wastewater with one another prior to reuse. In order to assess the extent of sharing, I determine the firm (or, potentially, firms) associated with each observed destination in the disposal data.[11] I assume that all wastewater transferred to a well pad or CTF is ultimately reused. Table 1 presents market shares obtained using this procedure.[12]

[8] In practice, the choice between CTF and on-pad treatment primarily turns on a tradeoff between economies of scale and transportation costs. Regardless of ownership, the use of CTFs can increase transportation costs because wastewater must be transported twice – once to the CTF, and then again to a location where it can be reused. There are also differences in regulatory compliance costs that factor into this decision, such as differences in bonding requirements.

[9] For example, I do not observe whether outsourced transfers are mediated by direct interaction between two rival operators, or through a third party. Incentives might differ in each of these cases.

[10] In contrast, re-transfer of wastewater from one well pad to another is prohibited by the DEP.

[11] In the monthly disposal records, any destination well site is recorded using a unique identifier that can be linked to ownership information in other DEP records. This enables me to infer whether a destination well site was operated by the same firm that generated the wastewater, or by another rival firm. For shipments to CTFs I make a similar inference on the basis of the CTF's location and permit number.

[12] The "Other" category encompasses (for example) shipments for reuse in West Virginia and landfill disposal of unusable sludges produced as a byproduct of treatment.

Table 1: Wastewater disposal market shares

Mode	Facility	% Mode	% Facility
Internal reuse	Own well pad	80.3	46.5
	Own CTF	-	21.9
	3rd party CTF	-	12.0
Rival reuse	Rival well pad	8.3	6.3
	Rival CTF	-	2.0
Injection well		8.1	8.1
Other		3.3	3.3

88.6% of all wastewater is transferred to well pads or CTFs for reuse. In 8.3% of these transfers, the well pad or CTF is exclusively linked to a rival firm.

One complication in assessing the aggregate sharing rate is that wastewater initially transferred to a CTF can later be re-transferred.[13] Thus, I do not observe the ultimate location of reuse for wastewater initially transferred to a CTF: shipments to rival CTFs could be reused internally, while shipments to internal CTFs could be reused by rivals. Third party CTFs could treat water on a contract basis for internal reuse, or could instead serve as de facto brokers between rivals. Allowing for this possibility, the sharing rate can be bounded between 7.1% and 47.6%. If third party CTFs primarily treat water on a contract basis for internal reuse, then a plausible estimate of the sharing rate is 10.6%.

Regardless of the true sharing rate, the patterns in Table 1 reveal the presence of a "sharing market," in which wastewater frequently crosses firm boundaries prior to being reused. Sharing might occur if one firm's wastewater production exceeds its needs for new completions, or if transportation synergies are present. In the remainder of this section, I briefly describe patterns of reuse and sharing in the disposal data.

[13] In contrast, the DEP explicitly precludes firms from accepting water at one well pad and then later transferring it to another. According to the DEP, this regulation is intended to prevent excessive truck traffic.

2.3.1 Reuse as a matching problem

Wastewater reuse requires firms to solve a many-to-few matching problem. Much like the hydrocarbons it accompanies, wastewater is produced in declining volumes over the life of a well. Consequently, the number of well pads generating wastewater (a stock) is large in comparison to the number of new completions (a flow). In the data, 1,712.6 distinct well pads reported wastewater disposals each month, while only 55.0 new wells were completed.[14] The median disposal amount was 415 barrels — less than 0.5% of the fluid volume needed to complete a new well. In contrast, the mean well pad recorded as a destination received 46,814 barrels of wastewater from 31.4 distinct origin well pads. This information is presented in the first two panels of Table 2, which summarize the number of facilities appearing in the data each month and associated shipment volumes (in "truckloads").[15]

One potential motivation for sharing is to realize more efficient matchings of wastewater from old wells to new wells (and CTFs). Because it is costly to transport wastewater, distance appears to be an especially important factor in shaping matching patterns. The last section of Table 2 shows the distribution of shipment distance by destination type. The mean shipment distance was 30.0 miles, roughly 20% of the average distance between all origins and destinations. In Appendix C.1, I show that 15.6% of wastewater was shipped to the nearest recorded destination that received wastewater from any source, and 53.3% of wastewater was shipped to one of the five nearest recorded destinations.

2.3.2 When do firms share wastewater?

During the sample period, 49 out of 75 firms shared wastewater (i.e., transferred wastewater to a rival firm) on at least one occasion, including 9 of the 10 largest firms. On average, a

[14] To obtain the mean number of completions, I take the average number of fracking jobs recorded in FracFocus during the sample period. By comparison, EIA's Drilled but Uncompleted Wells (DUC) data implies 101.5 completions per month for the whole of Appalachia, including Ohio and West Virginia.

[15] The raw data are reported in barrels. I convert barrels to truckloads by dividing by 110 (the modal shipment volume in the data; in practice, water-hauling tanker truck capacity varies from about 80 to 130 barrels).

Table 2: Facility counts and shipment characteristics

	Mean	Std	5%	25%	50%	75%	95%
Facility count per month							
Well pads (origin)	1,712.6	76.6	1,587.2	1,659.8	1,707.5	1,763.0	1,831.0
Well pads (dest)	51.6	14.3	31.4	39.0	50.5	62.8	74.0
Producer CTFs (dest)	11.0	1.5	9.0	10.0	11.0	12.0	13.0
3rd party CTFs (dest)	10.9	1.3	9.0	10.0	11.0	12.0	13.0
Truckloads sent or received by facility-month							
Well pads (origin)	23.8	83.0	0.6	1.4	3.7	11.8	98.4
Well pads (dest)	430.8	915.4	0.9	4.0	29.5	363.0	2,347.7
Producer CTFs (dest)	905.1	1,429.1	2.9	51.4	271.8	1,071.6	4,683.4
3rd party CTFs (dest)	464.5	563.9	4.8	106.0	337.2	649.6	1,349.1
Miles per truckload by destination type							
Own pad or CTF	22.5	20.2	2.8	8.7	17.5	31.5	53.6
Rival pad or CTF	45.0	30.8	10.8	24.2	39.5	57.2	98.7
Injection well	75.5	54.0	18.1	30.1	68.0	88.6	215.9
3rd party CTF	31.4	29.8	4.4	10.6	24.4	44.6	76.0
All destinations	30.0	30.9	3.4	10.5	21.4	37.2	85.1

firm that shared at least once shared in more than half of all sample months. Thus, sharing is widespread and firms that share tend to do so frequently.

To better understand when firms share wastewater, I estimate a series of logit regressions of monthly sharing market participation on firm characteristics. The results are summarized in Table 3. The dependent variable $Shared_{i,t}$ is an indicator for whether firm i transferred wastewater to any rival firm in month t. Across specifications, the most important variable in terms of explained variance is whether i shared wastewater in the previous month (i.e., $Shared_{i,t-1}$). In the data, a firm that shared in the previous month is 73 percentage points more likely to share in the current month than a firm that did not (86.6% vs. 13.5%). Thus, there appears to be persistence in the decision to share. The second most important variable is a measure of i's potential transportation cost savings from participation in the sharing market. This measure is intended to capture the potential difference in i's shipping demand with and without the possibility of sharing.[16] It takes on higher values for firms that operate

[16] In particular, I calculate the difference between the minimum shipment distance for disposal of i's wastewater in two scenarios: (i) if i were restricted to internal reuse and final disposal; and (ii) if i were able ship wastewater to any facilities that accepted wastewater for reuse, including those owned by rivals.

Table 3: Firm-Level Predictors of Sharing Activity

	Dependent Variable: $Shared_{it}$					
	(1)	(2)	(3)	(4)	(5)	(6)
Potential distance reduction	0.027***		0.031***	0.030***		0.032***
	(0.005)		(0.005)	(0.004)		(0.005)
Large regional firm		0.611	1.511**	1.183***	0.340	1.041***
		(0.669)	(0.616)	(0.380)	(0.381)	(0.338)
Small regional firm		0.466	0.497	0.327	0.300	0.513
		(0.544)	(0.535)	(0.357)	(0.328)	(0.338)
Shared previous month				3.531***	3.710***	3.548***
				(0.266)	(0.257)	(0.263)
Concurrent receipt					0.081	0.727***
					(0.254)	(0.262)
McFadden R^2	0.1513	0.0083	0.1835	0.5164	0.4514	0.5215
Observations	1,832	1,832	1,832	1,768	1,768	1,768
Log Likelihood	−1,077.673	−1,259.228	−1,036.752	−614.039	−696.632	−607.574

Notes: Standard errors clustered at the firm level. The "National firm" category is excluded from the firm type dummies.

further from Ohio injection wells, for firms that complete new wells infrequently, and for firms that are located especially close to rivals. As expected, geographic synergies appear to be a key determinant of whether firms share.

I divide firms into three broad categories: "national" firms with substantial operations in multiple oil and gas plays, "large regional" firms that operate exclusively in Appalachia (or nearly so), and small regional firms.[17] Relative to national firms, both large regional firms and small regional firms are more likely to share. After controlling for potential transportation savings, this difference statistically significant for the large regional firms. National firms could be less likely to share on account of having greater liability aversion, greater incentives to protect trade secrets embodied in wastewater, or lower-powered managerial incentives to reduce disposal costs. I discuss these factors further in the context of the main results.

Finally, I find that firms that receive wastewater transfers from rivals in a given month are more likely to share in that month. This suggests that firms do not strictly prefer internal reuse to sharing when both options are available.

2.3.3 Evidence of transaction costs

A merger of two large regional firms that occurred during the sample period provides some evidence that transaction costs in the sharing market may be significant.

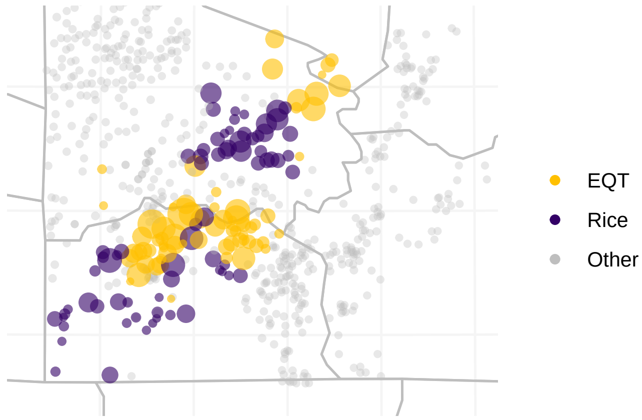
In November 2017 EQT Corporation ("EQT") and Rice Energy Inc. ("Rice") merged, creating one of the largest natural gas producers in the United States. Prior to merging, EQT and Rice both operated large numbers of wells on overlapping acreage in far southwestern Pennsylvania.[18] The locations of EQT and Rice well pads during this period are depicted in Figure 3. In the six months leading up to the merger announcement, 98% of Rice's wastewater volume originated at well pads within 20 miles of an EQT facility that received

In either case, I assume that a facilities capacity corresponds to the observed volume of receipts.

[17] A complete list is provided in Appendix A. Note that the "large regional" firms are typically greater in size than the national firms.

[18] EQT also had a significant presence in West Virginia, while Rice was present in Ohio. Both were among the ten largest firms in Pennsylvania prior to merging.

Figure 3: EQT and Rice pre-merger well pad locations



wastewater, and 64% of EQT's wastewater volume originated at well pads within 20 miles miles of a Rice facility that received wastewater. Nevertheless, EQT and Rice did not share prior to the merger, even though EQT and Rice both shared with other firms.[19]

After the merger, transfers between formerly-unintegrated EQT and Rice facilities increased dramatically. 22.5% of wastewater generated at former EQT well pads was transferred to former Rice facilities, and 62.4% of wastewater generated at former Rice facilities was transferred to former EQT facilities. This change is summarized in Table 4, which shows pre- and post-merger disposal and origination shares for former EQT and Rice facilities. Thus, the removal of the firm boundary was followed by a significant increase in "sharing," consistent with the elimination of significant ex ante transaction costs.

2.4 Discussion

In Pennsylvania, key regulatory authorities including the DEP and SRBC encourage wastewater reuse and sharing. One potential motivation for this position is the magnitude of avoided

[19]See Table 4, discussed below. Rice did not send wastewater to any rival firms in the pre-merger period, while 2.7% of all shipments to EQT originated at other rivals' well pads. 8.9% of EQT's wastewater was sent to rival E&Ps other than Rice, while 91.1% was reused internally or sent to a third party CTF. This was the case even though Rice received substantial volumes of wastewater from another rival firm, Alpha Shale Resources. Rice and Alpha Shale Resources had formerly been partners in a joint venture.

Table 4: EQT and Rice pre- and post-merger market shares

Share of Wastewater Leaving Well Pad				
	Pre-merger		Post-merger	
Destination	EQT	Rice	EQT	Rice
EQT pad	83.4	0.0	65.1	62.4
Rice pad	0.0	70.6	22.5	31.0
Other rival	8.9	0.0	2.7	0.6
3rd party CTF	7.7	29.3	8.5	4.6
Injection well	0.0	0.0	1.2	1.4

Share of Wastewater Received				
	Pre-merger		Post-merger	
Source	EQT	Rice	EQT	Rice
EQT pad	97.3	0.0	59.1	50.5
Rice pad	0.0	95.9	40.1	49.1
Other rival	2.7	4.1	0.8	0.4

Notes: The pre-merger period spans January to June 2017, while the post-merger figure spans December 2017 to December 2020.

external costs. Absent sharing, most wastewater would be shipped long distances to Ohio injections wells, creating numerous environmental impacts.[20] Sharing appears to play an important role in facilitating reuse, and yet the changes in wastewater reuse associated with the EQT-Rice merger suggest that significant social benefits may go unrealized under the status quo. In the remainder of the paper, I develop and estimate an empirical model of firms' reuse and sharing decisions in the presence of transaction costs. This allows me to uncover the realized and potential effects of sharing on outcomes of interest such as private costs, injection well disposal rates, and wastewater-related truck traffic.

3 Model

This section introduces the model. The goal of the model is to quantify the role of sharing in mediating wastewater reuse. To do so, the model must simultaneously capture how firms balance the costs of internal reuse against the costs of sharing and final disposal.

I adopt the framework of transferrable utility (TU) matching (see, e.g., Choo and Siow,

[20]This would dramatically increase truck traffic on Pennsylvania roadways, resulting in increased greenhouse gas emissions, air pollution, risks of spills and traffic fatalities, congestion disamenities, and wear-and-tear on roadways. Injection disposal has been linked to seismic activity and may pose risks to drinking water resources (Groundwater Protection Council, 2019). Meanwhile, firms' freshwater consumption would increase. Despite being rich in freshwater, Pennsylvania has experienced drought and water shortages in recent years (Wilson, 2023).

2006). Let K_t denote the finite set of well pads generating wastewater in month t, and D_t the finite set of facilities accepting wastewater for reuse. Firm f manages a subset of well pads $K_{tf} \subset K_t$ and a subset of facilities that accept wastewater $D_{tf} \subset D_t$. In month t, well pad $\kappa \in K_t$ generates Q_κ truckloads of wastewater that must be disposed of immediately. Up to C_δ truckloads of wastewater can be accepted at facility $\delta \in D_t$, but no more. All wastewater must be shipped to some facility in D_t or to final disposal, the outside option. For the remainder of this section I suppress the dependence of all objects on t.

Firm f seeks to minimize its total cost of wastewater disposal and water acquisition. Firms may choose to transport wastewater to their own facilities or to facilities owned by other firms. The cost of transporting a truckload of wastewater ij from $\kappa \in K_f$ to $\delta \in D$ for the purpose of reuse is given by $r_{\kappa\delta} - \epsilon_{i\delta} - \eta_{\kappa j}$, where $r_{\kappa\delta}$ denotes the systematic cost of reuse and $\epsilon_{i\delta}$ and $\eta_{\kappa j}$ capture shipment-specific cost shocks for the sending and receiving firms, respectively. If $\delta \in D_f$, firm f incurs the full cost of reuse $r_{\kappa\delta} - \epsilon_{i\delta} - \eta_{\kappa j}$ for truckload ij. On the other hand, if $\delta \in D_{f'}$ for some rival firm f', then firm f incurs a cost $\alpha^K_{\kappa\delta} r_{\kappa\delta} + p_{\kappa\delta} - \epsilon_{i\delta}$ while firm f' incurs a cost $\alpha^D_{\kappa\delta} r_{\kappa\delta} - p_{\kappa\delta} - \eta_{\kappa j}$ where $\alpha^K_{\kappa\delta} + \alpha^D_{\kappa\delta} = 1$ (for either firm, costs may be negative). The cost to firm f of sending i to final disposal is $r_{\kappa 0} - \epsilon_{i0}$, while the cost of not allocating j to wastewater from any well pad in K is $r_{0\delta} - \eta_{0j}$. In this case, $r_{0\delta}$ represents the cost of obtaining freshwater at δ.

I consider a matching game in which firms choose exactly one shipment destination δ for each truckload i and exactly one supply location κ for each unit of capacity j. As the volume of wastewater grows large, the core of this matching game coincides with the set of Walrasian equilibria of an exchange economy (Gretsky et al., 1992).[21] In a Walrasian equilibrium, firm f chooses among potential shipment destinations for the ith truckload from κ to minimize

[21] The core of the matching game consists of the set of all *stable*, *feasible* matchings. A matching is *feasible* if every truckload i is matched to some $\delta \in D_0$, and every j is allocated to a truckload of wastewater or freshwater from some $\kappa \in K_0$. A matching μ is *stable* if no firm would prefer to ship a matched truckload i to final disposal or to allocate a matched unit of capacity j to freshwater, and no two firms (possibly the same) would privately agree to match any i and j not matched under μ.

costs taking as given the equilibrium price matrix p:

$$\min_{\delta \in D_0} \quad \alpha^K_{\kappa\delta} r_{\kappa\delta} + p_{\kappa\delta} - \epsilon_{i\delta} \tag{1}$$

Simultaneously, for the jth truckload of capacity at δ firm f chooses among potential shippers to minimize costs conditional on p:

$$\min_{\kappa \in K_0} \quad \alpha^D_{\kappa\delta} r_{\kappa\delta} + p_{\kappa\delta} - \eta_{\kappa j} \tag{2}$$

Thus, when deciding to ship wastewater within the firm rather than sharing it, a firm simultaneously decides that δ is the least cost destination for i and κ the least cost source for j, taking into account p. For shipments between firms, p can be interpreted as a payment from the sending firm to the receiving firm. For shipments within a firm, p can be interpreted as a shadow cost (or benefit) associated with using the firm's scarce resources.

An equilibrium matching can be summarized by a matrix μ with K rows and D columns, where $\mu_{\kappa\delta}$ represents the probability of observing a shipment of wastewater between κ and δ. In equilibrium, $\mu_{\kappa\delta}$ coincides with the operator of well pad κ's demand for shipments to facility δ and the operator of facility δ's demand for shipments from well pad κ when choices are made according to (1) and (2). Galichon and Salanie (2022) demonstrate that under mild restrictions on the distributions of ϵ and η, an equilibrium matching μ^* can be obtained by maximizing a social surplus function:

$$\min_{\mu \in \mathcal{M}(\mathbf{Q},\mathbf{C})} \sum_{\kappa \in K} \sum_{\delta \in D} \mu_{\kappa\delta} \left\{ r_{\kappa\delta} - r_{\kappa 0} - r_{0\delta} \right\} - \mathcal{E}\left(\mu, \mathbf{Q}, \mathbf{C}\right) \tag{3}$$

where $\mathcal{E}\left(\mu, \mathbf{Q}, \mathbf{C}\right)$ is a *match entropy function* that depends on the distributions of ϵ and η and $\mathcal{M}\left(\mathbf{Q}, \mathbf{C}\right)$ is the set of feasible matchings. I provide further discussion of the match entropy function in Appendix B.1. When taking the model to the data, I assume that $\epsilon_{i\delta}$ is distributed iid across δ for each i according to an extreme value type I distribution with scale parameter σ_ϵ, while $\eta_{\kappa i}$ is distributed iid across κ for each j according to an extreme value type I distribution with scale parameter σ_η (as in Choo and Siow, 2006). In this case, one can show that:

$$-\mathcal{E}\left(\mu, \mathbf{Q}, \mathbf{C}\right) = C\left(\mathbf{Q}, \mathbf{C}\right) + \left(\sigma_\epsilon + \sigma_\eta\right) \sum_{\kappa \in K} \sum_{\delta \in D} \mu_{\kappa\delta} \log \mu_{\kappa\delta} \\ + \sigma_\epsilon \sum_{\kappa \in K} \mu_{\kappa 0} \log \mu_{\kappa 0} + \sigma_\eta \sum_{\delta \in D} \mu_{0\delta} \log \mu_{0\delta} \tag{4}$$

where $\mu_{\kappa 0}$ is the mass of unmatched wastewater originating at κ, $\mu_{0\delta}$ is the mass of unmatched capacity at δ, and $C\left(\mathbf{Q}, \mathbf{C}\right)$ is a constant that does not depend on μ. Formally, (3) is an optimal transport (OT) problem with entropic regularization (see, e.g., Peyre and Cuturi, 2020). If there were no variation in ϵ or η, the match entropy term would disappear and (3) would reduce to a linear program (LP).

As (3) makes clear, equilibrium shipping patterns under TU matching are driven by the structure of the systematic costs of reuse $r_{\kappa\delta}$. The systematic costs of reuse may depend on observables such as shipment distance and the presence of firm boundaries, as well as unobservables such as differences in timing and treatment costs. I assume that:

$$r_{\kappa\delta} - r_{\kappa 0} - r_{0\delta} = x'_{\kappa\delta}\theta + u_{\kappa\delta} \tag{5}$$

where $x_{\kappa\delta}$ is a vector of observables and $u_{\kappa\delta}$ is a scalar unobservable assumed to be iid across κ and δ. I assume that $E\left[u_{\kappa\delta} | x_{\kappa\delta}\right] = 0$ and that $Var\left(u_{\kappa\delta} | x_{\kappa\delta}\right)$ is finite.

3.1 Discussion

The model assumes that firms minimize total costs by making truckload-by-truckload decisions, both when sending and receiving wastewater. This assumption is without loss in the context of a matching model: the firm can be viewed as a coalition of individual managers, one located at each facility, and in the core no coalition of managers can achieve lower costs than managers acting independently. The assumption of transferrable utility is reasonable because firms can readily exchange cash, but $p_{\kappa\delta}$ does not necessarily represent a cash transfer.[22] Equilibrium transfers within the firm capture the shadow costs of shipments that crowd out more efficient internal shipments or profitable exchanges in the sharing market.

The primary restriction imposed by the model is that firms have no market power and do not behave strategically. For example, a firm cannot earn more surplus by threatening to abstain from sharing (as recognized by Shapley and Shubik (1971)). Some important forms of strategic behavior could be captured with a multilateral bargaining model such as the well-known "Nash-in-Nash" model associated with Horn and Wolinsky (1988). Apart from the lack of price data, a practical challenge in adapting the Nash-in-Nash approach to this setting is the difficulty of constructing counterfactuals in which the set of trading partners is allowed to adjust, as would likely occur if transaction costs were eliminated.[23]

Another limitation of the model is that systematic costs are incurred in proportion to the number of truckloads sent. Linearity of the transaction costs excludes the possibility that some costs might be amortized over many similar truckloads. In this way, the model differs from trade models that stipulate fixed costs of importing and exporting (Antrás and Chor, 2022). A related shortcoming is the assumption that that all systematic costs are exogenous. For the case of distance, exogeneiety is plausible—there are many factors apart from wastewater management costs that affect where firms choose to complete new wells. For any transaction costs, however, it is less reasonable. From a theoretical perspective, it would be natural to endogenize transaction costs with a model of relationship dynamics. Given the difficulty of estimating any such dynamics, I do not pursue this approach here.

[22] For example, $p_{\kappa\delta}$ could represent a "favor" (as in Samuelson and Stacchetti, 2017).

[23] Ho and Lee (2019) and Ghili (2022) develop models of network formation in Nash-in-Nash environments. These papers exploit institutional differences between upstream and downstream firms to simplify the strategy space (e.g., by assuming that one side of the market can pre-commit to a particular network) that have no clear analogues in this setting.

4 Estimation

I estimate the model using Poisson pseudomaximum likelihood (PPML). Applying the identification result of Galichon and Salanie (2022), it follows from (3) that:

$$\log \mu_{\kappa\delta} = -\left(\sigma_\epsilon + \sigma_\eta\right)^{-1}\left(r_{\kappa\delta} - r_{\kappa 0} - r_{0\delta}\right) + F_\kappa + H_\delta \tag{6}$$

where F_κ and H_δ are sending- and receiving-facility fixed effects that satisfy a set of $|K|+|D|$ market clearing conditions:

$$\begin{aligned} \exp\left\{\gamma^{-1}F_\kappa\right\} + \sum_{\delta\in D}\mu_{\kappa\delta} &= Q_\kappa \quad \forall\, \kappa \in K \\ \exp\left\{(1-\gamma)^{-1}H_\delta\right\} + \sum_{\kappa\in K}\mu_{\kappa\delta} &= C_\delta \quad \forall\, \delta \in D \end{aligned} \tag{7}$$

and $\gamma = \sigma_\eta^{-1}/\left(\sigma_\epsilon^{-1} + \sigma_\eta^{-1}\right)$ is a parameter summarizing the relative dispersion of $\epsilon_{i\cdot}$ and $\eta_{\cdot j}$. Substituting the specification (5) into (6) and exponentiating both sides gives:

$$\mu_{\kappa\delta} = \exp\left\{x'_{\kappa\delta}\tilde{\theta} + F_\kappa + H_\delta\right\}\eta_{\kappa\delta} \tag{8}$$

where $E\left[\eta_{\kappa\delta}|x_{\kappa\delta}\right] = 0$ and $\tilde{\theta} = -\left(\sigma_\epsilon + \sigma_\eta\right)^{-1}\theta$. Given (7), this expression is closely related to the familiar structural gravity from trade (Head and Meyer, 2014). PPML is an (inefficient) weighted least squares estimator based on (8) that is known to exhibit good performance in trade datasets with heteroskedastic errors and a dependent variable that frequently takes a value of zero (Santos Silva and Tenreyro, 2022), as is the case here. In the absence of data on firms' use of the outside options, the PPML estimator solves:

$$\min_{\tilde{\theta},F,H} \quad \sum_{t\in T}\sum_{\kappa\in K_t}\sum_{\delta\in D_t}\left[\hat{\mu}^t_{\kappa\delta}\left\{x^{t\prime}_{\kappa\delta}\tilde{\theta} + F^t_\kappa + H^t_\delta\right\} - \exp\left\{x^{t\prime}_{\kappa\delta}\tilde{\theta} + F^t_\kappa + H^t_\delta\right\}\right] \tag{9}$$

As in the case of the gravity model, components of $\tilde{\theta}$ that are linearly independent of the fixed effects F and H are identified. Although the number of fixed effects is large, estimation of $\tilde{\theta}$ is not biased by the potential incidental parameters problem affecting estimation of the fixed effects as K_t and D_t grow large (Fernandez-Val and Weidner, 2016). However, conventional Eicker-White standard errors for $\tilde{\theta}$ are biased downwards and confidence intervals may be incorrectly centered (Kauermann and Carroll, 2001). To address this issue, I construct two-step bootstrap standard errors and confidence intervals following Zylkin (2024).

For the main estimates, I modify (9) to include data on firms' use of the outside options. This modification is essential for capturing the substitution between reuse and firms' outside options, a key margin considered in the counterfactual analysis. Furthermore, including the outside options enables separate identification of sending- and receiving-facility characteristics apart from F and H. For a fixed γ, the augmented PPML estimator solves:

$$\begin{aligned}\min_{\tilde{\theta},F,H} \quad & \sum_{t\in T}\sum_{\kappa\in K}\sum_{\delta\in D}\left[\hat{\mu}^t_{\kappa\delta}\left\{x^{t\prime}_{\kappa\delta}\tilde{\theta}+F^t_\kappa+H^t_\delta\right\}-\exp\left\{x^{t\prime}_{\kappa\delta}\tilde{\theta}+F^t_\kappa+H^t_\delta\right\}\right] \\ & +\sum_{t\in T}\sum_{\kappa\in K}\left[\left(\hat{\mu}^t_{\kappa 0}\right)^\gamma F^t_\kappa-\exp\left\{F^t_\kappa\right\}\right]+\sum_{t\in T}\sum_{\delta\in D}\left[\left(\hat{\mu}^t_{0\delta}\right)^{1-\gamma}H^t_\delta-\exp\left\{H^t_\delta\right\}\right]\end{aligned} \tag{10}$$

This estimator inherits many of the attractive properties of PPML. To ensure that the fixed effects enter all exponential terms without scaling, the observed final disposal and freshwater consumption volumes are exponentiated γ and $1-\gamma$, respectively. An important drawback of (10) is that γ is assumed to be known. For the main results I fix $\gamma = 0.5$, such that sender- and receiver- cost shocks ϵ and η have the same dispersion.[24]

As is well known in the literature on gravity models, both (9) and (10) can be implemented using an iteratively re-weighted least squares (IRLS) procedure that exploits the Frisch-Waugh-Lovell theorem to avoid inverting large matrices (Correira et al., 2020). After convergence, estimates of F and H conditional on $\tilde{\theta}$ and can be obtained using Sinkhorn's

[24] In principle γ is identified from the relative rates at which sending- and receiving- facilities substitute to the outside option as the sizes of K_t and D_t vary. However, estimation of γ is computationally demanding and complicates inference. In Appendix C, I explore the sensitivity of the results to alternative values of γ.

algorithm (Idel, 2016; Galichon and Salanie, 2022).[25]

4.1 Implementation details

In the main specification, x includes indicators for firm boundaries interacted with time dummies and various firm- and relationship characteristics that may shift transaction costs. x also includes shifters of the costs of firms' outside options.

The estimation sample consists of all observed shipments for reuse, subject to the following exclusions. First, I omit shipments to third party CTFs, since it is ambiguous whether such shipments are qualitatively different from shipments for internal reuse or sharing.[26] Second, I omit shipments from well pads that were concurrently recorded as destinations for wastewater from other well pads. By excluding shipments from these well pad-months, I avoid allowing that the wastewater produced by one fracking event could have been (impossibly) reused as an input for that same fracking event.[27]

Implementation of (10) requires data on final disposal and freshwater consumption at the facility-month level. Final disposal rates are directly observable in the data. Freshwater consumption is not directly observed in the data. Instead, I construct a proxy measure for freshwater consumption rates at the firm-month level by comparing each firms' total water usage as reported in FracFocus against wastewater receipts in the disposal data. I validate this measure against a sample of well-level freshwater consumption records provided by the SRBC. Appendix A.1 presents validation results demonstrating that the FracFocus-derived measure provides a reasonable approximation of firm-level freshwater consumption. Nevertheless, the necessity of relying on this proxy measure introduces a source of measurement

[25]Observe that the first order conditions for the fixed effects in (9) and (10) are closely related to the market clearing conditions (7). If (9) is used to estimate a matching model without outside options, they coincide. In addition to being helpful for recovery of the fixed effects, this approach can be used to obtain good starting values and to accelerate the bootstrap procedure.

[26]Alternatively, one could classify shipments to third party CTFs as internal shipments, or as shipments to the outside option. The former approach would likely understate transaction costs associated with sharing. The latter would complicate the interpretation of substitution to the outside option in counterfactuals as well as the estimated costs of the outside options discussed below.

[27]This restriction results in a loss of 11.4% of wastewater from the sample. An alternative approach would be to assign an infinite cost to these particular shipments.

error that is avoided by (9). Thus, I report several key estimates using both (9) and (10).

4.1.1 Counterfactuals

To construct counterfactuals, I solve the market clearing conditions (7) under the assumption that $u_{\kappa\delta} = 0$ for all κ and δ. Thus, the counterfactuals capture changes in expected shipment patterns relative to the fitted model. In all counterfactuals I hold fixed the set of well pads reporting disposal K, the set of facilities accepting water D, as well as the facility-level demands for disposal $\mathbf{Q}$ and capacities $\mathbf{C}$. Likewise, the costs of the outside options are held fixed at the estimated level, excluding the possibility of equilibrium adjustments in the costs of final disposal and/or freshwater consumption. The shipping distance associated with final disposal from well pad κ is assumed to be the average distance from κ to the set of injection disposal facilities observed in the data. I do not account for shipping associated with freshwater acquisition, which may be substantial.[28]

5 Estimates

I begin by considering a series of preliminary specifications in which transaction costs are assumed to be homogenous transactions. These results are presented in Table 5. The first three columns use the log of over-the-road mileage as a distance measure. The reported pseudo-R^2 statistic is the square of the correlation between the predicted and realized shipment volumes among shipments for reuse. The second column indicates that after controlling for the presence of a firm boundary, the elasticity of shipments for reuse with respect to distance is -1.1, similar to typical estimates in the trade literature.[29] If firm boundaries were not accounted for, the estimated elasticity would be much larger, and the model fit (as measured

[28] I exclude freshwater-related shipments for two reasons. First, the distance to freshwater sources will typically be much smaller than the average distance to injection wells. Second, unlike wastewater, freshwater is frequently transported via pipeline. I do not observe the relative share of trucking vs. pipeline transportation for freshwater.

[29] Head and Meyer (2014) report -1.1 to be the mean distance elasticity of trade across a large number of studies estimating structural gravity models. Looking at manufacturing establishments in the US, Atalay et al. (2019) obtain an estimate of -0.96. See also Chaney (2018).

Table 5: Transaction Cost Estimates

	Dependent Variable: # of Truckloads					
	(1)	(2)	(3)	(4)	(5)	(6)
Distance	−1.896	−1.120	−1.073	−0.070	−0.045	−0.039
	(0.005)	(0.003)	(0.006)	(0.0004)	(0.0003)	(0.0004)
Firm Boundary		−5.674	−4.788		−5.804	−4.846
		(0.038)	(0.038)		(0.039)	(0.038)
EQT-Rice Post-Merger		0.005	0.016		−0.171	−0.164
		(0.022)	(0.022)		(0.021)	(0.023)
Distance Measure	Log	Log	Log	Linear	Linear	Linear
κ-month FEs	Yes	Yes	Yes	Yes	Yes	Yes
δ-month FEs	Yes	Yes	Yes	Yes	Yes	Yes
Outside Options	No	No	Yes	No	No	Yes
Pseudo R^2	0.670	0.883	0.873	0.672	0.895	0.886
Observations	3,775,553	3,775,553	3,843,919	3,775,553	3,775,553	3,843,919

by the pseudo-R^2 value) would be comparatively poor. In the third column, shipment data for the outside options are included. Despite the potential for measurement error, I obtain a similar distance elasticity and similar model fit, although transaction costs are estimated to be somewhat smaller in magnitude.

The last three columns report results for the same specifications using over-the-road mileage itself ("linear distance") rather than log distance. While the coefficient on log distance has a convenient interpretation, linear distance may be a more suitable proxy for firms' costs in the context of wastewater management. Distances are short relative to the distances between foreign countries. Nearly all shipments occur via trucking, and trucking contracts often have a per-mile component. Given this, it is perhaps unsurprising that the specifications using linear distance have slightly higher pseudo-R^2 values than the corresponding log specifications. Motivated by this, I henceforth assume linear distance. In Appendix C.2, I report alternative results obtained using driving time.

Turning to the magnitude of transaction costs, the point estimate for firm boundaries in

Table 6: Cost Savings from Reuse vs. Transaction Costs

			$/Barrel	
	Miles	SE	Low	High
Cost Savings from Reuse (SW)	22.0	1.10	0.60	1.20
Cost Savings from Reuse (NE)	80.0	9.57	2.18	4.37
Firm Boundary	123.3	20.82	3.36	6.73
EQT-Rice Post-Merger	4.2	0.40	0.11	0.23

Notes: The "low" column assumes marginal trucking costs are $3/mile. The "high" column assumes marginal trucking costs are $6/mile. SEs obtained from the bootstrap covariance matrix via the delta method.

Column (6) of Table 5 indicates that crossing the firm boundary (in other words, sharing) raises the total costs of reuse by an amount equivalent to the cost of shipping a truckload of wastewater $\frac{4.846}{0.039} = 123.3$ additional miles, roughly four times the mean shipment distance observed in the data. For trucking costs on the order of $3 per mile, this implies a transaction cost of a bit more than $3 per barrel of wastewater. Table 6 reports these figures along with the estimated for the "cost savings" for Column (6) – the foregone costs of injection disposal and freshwater acquisition net of treatment costs. I report one estimate for shipments within southwestern Pennsylvania, and another for shipments within northeastern Pennsylvania. Cost savings are $0.50-2.00 per barrel under the same trucking cost assumptions, somewhat smaller than industry reports.[30] Thus, transaction costs are estimated to be economically large and of similar magnitude to the surplus created by reuse. This finding is consistent with the descriptive evidence presented in Section 2 suggesting that high transaction costs coincide with frequent trade, a possibility highlighted by (Demsetz, 1988).

If the estimated transaction costs represent the costs of crossing the firm boundary, then they can be eliminated by bringing a transaction inside the firm. To validate this implication

[30] According to industry insiders, the cost of treatment is roughly $0.25 per barrel. In comparison, the cost of injection disposal from northeastern PA is $8-12 per barrel before injection fees of $2-4 per barrel. Freshwater costs are typically less than $1 per barrel. Thus, the cost savings from reuse are on the order of $10-15 per barrel in northeastern PA, and less in southwestern PA on account of closer proximity to injection wells. I thank the SRBC for this information.

of the model, the specifications that include firm boundaries also include a dummy for post-merger transactions across the pre-merger boundary between EQT and Rice. If mergers eliminate transaction costs, then the estimated coefficient should be equal to zero. For the linear distance specifications, the estimates are small and negative, implying the existence of a small but positive cost of crossing the pre-merger firm boundary. This cost is estimated to be only 3.4% as large as the cost of sharing. For the log distance specifications, the post-merger dummy is statistically insignificant. These results lend support to the interpretation of the estimated transaction costs as "Coasean" in the sense of primarily capturing managerial frictions rather than unobserved technological factors that might inhibit sharing.

5.1 Sources of transaction costs

For the main results, I use a richer baseline specification that allows for greater heterogeneity in transaction costs across transactions. Detailed parameter estimates for this model are provided in Appendix C.3. The key implications are summarized in Table 7, which reports the mean estimated transaction cost for different categories of transactions. The average estimated transaction cost for a potential transaction is equivalent to the cost of shipping a truckload of wastewater 156.7 miles, about the same as the average distance between origins and destinations.[31] However, there is considerable variation across potential transactions.

There are two main sets of results. First, I find that transaction costs are nearly 50% lower on average between counterparties that traded in the previous month compared with counterparties that never previously traded. This suggests that transaction costs are persistently lower within some relationships, either for exogenous reasons or because trade itself reduces transaction costs. The latter force could arise if sharing entails relationship-specific fixed costs or learning. If this latter mechanism were important, one might expect to find lower transaction costs in northeastern Pennsylvania, where the gains from trade tend to be

[31] I report a weighted average in which the weights correspond to expected shipment volumes if matching occurred completely at random. Using the same weights, the mean shipment distance between origins and destinations was 168.6 miles.

Table 7: Mean transaction costs (in miles)

	Mean	SE
All transactions	156.7	7.03
Region		
Southwestern PA	145.6	5.88
Northeastern PA	164.4	8.39
Prior sharing status		
Last Month	94.7	3.17
Last Year	141.5	6.94
Never	175.4	10.99

	Mean	SE
Sending firm type		
Large regional	160.6	7.21
National	146.8	13.63
Small regional	141.4	8.53
Receiving firm type		
Large regional	147.5	6.53
National	200.8	13.75
Small regional	177.2	16.21

larger and trade occurs more frequently. However, I find that transaction costs are slightly higher on average in northeastern Pennsylvania ($p < 0.01$), suggesting that heterogeneity in transaction costs across relationships arises from other factors.

Second, I consider how transaction costs vary with firm characteristics. I find that transaction cost are higher when the sender is a large regional firm than a national firm or small regional firm ($p < 0.05$), despite the finding in Section 2 that large regional firms are the most likely to share. One potential explanation is that large regional firms, who are generally larger and possibly more productive than national firms, may have stronger incentives to protect any trade secrets embodied in wastewater. Nevertheless, being larger implies that these firms have greater total demand for disposal, and thus a greater probability of sharing in a given month despite elevated transaction costs. Characteristics of the receiving firm may affect transaction costs differently. I find that transaction costs are higher when the receiving firm is a "national" firm ($p < 0.01$). Managers within national firms may have weaker incentives to accept wastewater, either because incentives to reduce completion costs are low-powered or because of differing attitudes towards risks to well productivity. On the other hand, transaction costs tend to be higher when the receiving firm is a small regional firm rather than a large regional firm ($p < 0.01$) even though large regional firms share many characteristics with national firms. If smaller firms spill wastewater at higher rates, this could be consistent with a mechanism in which larger and more sophisticated firms avoid sending wastewater to smaller rivals because of perceived liability risks.

6 The value of sharing

Table 8 compares key outcomes in the estimated model to those obtained under changes in the cost of sharing in order to quantify the effects of sharing and transaction costs.

In the second row, transaction costs are infinite, and no sharing occurs. Instead, firms either reuse wastewater internally or ship it injection wells. Freshwater consumption rises to meet firms' completion needs. The estimates indicate under the status quo, sharing reduces injection disposal volumes by 37%. This reduction in disposal rates, along with possibility of realizing geographic synergies, reduces total truck-miles by 19%, from 37.1 to 29.9 miles per truckload. This reduction is concentrated among shipments originating from small regional firms, who turn to injection disposal at much higher rates in the absence of a sharing market.

In the third row, transaction costs are zero. In the absence of transaction costs, the final disposal rate falls to less than 1%. In this case, the mean shipment distance increases by 12% rather than decreasing as might have been expected. This occurs because the likelihood of matches based on unobserved cost shocks rather than distance increases rapidly as transaction costs fall. Because this conclusion may be sensitive to the assumption that cost shocks are drawn independently across origins and destinations, the fourth row reports a modified counterfactual in which transaction costs are eliminated, and distance is the only factor affecting shipments for reuse.[32] Interpreting the result as a lower bound indicates that the elimination of transaction costs could reduce shipping distances by up to 34%, with much of this potential reduction attributable to reduced final disposal. In either case, sharing rates increase by the largest margins for national firms.

[32]Final disposal and freshwater consumption rates are held fixed at the level implied by the elimination of transaction costs.

Table 8: Counterfactuals

	Mean Distance (mi)				Sharing Rate (%)				Disposal Rate (%)			
	Large	Nat'l	Small	All	Large	Nat'l	Small	All	Large	Nat'l	Small	All
Model	26.4	31.6	51.6	29.9	2.8	13.0	32.4	7.6	3.9	21.5	28.7	9.3
Transaction costs $\to \infty$	27.8	45.2	89.5	37.1	-	-	-	-	4.8	31.8	57.2	14.6
Transaction costs $\to 0$	31.3	36.6	42.9	33.3	56.5	80.1	88.1	63.6	0.3	0.2	0.3	0.3
Transaction costs $\to 0$, $\mathcal{E} \to 0$	18.1	22.5	27.1	19.7	38.5	61.0	74.4	45.9	0.3	0.2	0.3	0.3
No firm-size heterogeneity	26.3	31.3	51.3	29.8	2.9	14.5	32.9	7.9	3.8	20.8	28.4	9.2

6.1 Environmental spillovers

Taken together, the results imply that wastewater sharing meaningfully reduces injection disposal rates, while reducing the average length of trips by wastewater-hauling trucks. In this section, I explore the magnitude of external benefits associated with these impacts.

First, I consider the benefits of reduced wastewater trucking. A full water-hauling truck weighs about 40 tons. Using typical tanker-truck emissions rates, it follows that a 19% reduction in truck-miles saves roughly 7,000 metric tons of carbon emissions annually, which can be valued at \$0.3M using the EPA Social Cost of Carbon.[33] A similar calculation implies that the sharing market saves roughly 0.1 metric tons of PM2.5 emissions and 14 metric tons of NOx emissions annually, which can be valued at around \$0.3M per year.[34] Thus, sharing generates roughly \$0.6M per year in external benefits from reduced wastewater trucking.

Next I consider the benefits of reduced final disposal. Water that is sent to final disposal is effectively "destroyed" in the sense of being permanently removed from the hydrological cycle. This distinguishes the use of water in fracking from the use of water in other applications such as agriculture, where most water is ultimately recovered via evaporation. One way

[33] I use the average emissions factors for tanker trucks from EPA SmartWay Carrier data. This data is self-reported and may not be representative for the wastewater-hauling market specifically. I use the Social Cost of Carbon for 2020, assuming a 3% discount rate.

[34] This calculation is based on the EASIUR air quality model (Heo et al., 2016).To compute air pollution, I assume that all trucking-related air pollution occurs at the well site from which the wastewater originated, rather than along the trucking route. This likely results in an underestimate of air pollution damages because well pads are often located in remote areas, whereas trucking routes pass through more populous areas.

to approximate the social cost of final disposal is to consider the cost of desalinating an equivalent quantity seawater. Typical desalination costs imply that foregone final disposal from sharing generates \$0.2-0.6M per year in external benefits.[35]

A potentially larger source of external benefits from reduced final disposal relates to the risk of seismic activity. Final disposal in Ohio injection wells has been linked to notable earthquakes in Youngstown, OH and Poland Township, OH (Schultz et al., 2020). Over the period from 2017 to 2020, there were seven earthquakes in Ohio significant enough to cause disturbances with peak ground velocities of 0.5cm/s at a distance of one mile.[36] Wastewater from Pennsylvania accounted for approximately 10% of all injection disposals in Ohio during this period. As one indication of the external costs of seismic activity from final disposal, Koster and van Ommeren (2015) found that natural gas-related earthquakes generating peak ground velocities of 0.5cm/s decreased house prices by 1.9% in the Netherlands.

Although the benefits associated with reduced seismicity are difficult to quantify, the sum of all plausible external benefits appears to be relatively modest in comparison to private costs. Private transportation cost savings can be valued at \$10M per year under conservative trucking cost assumptions. Foregone disposal fees can be valued at \$5-10M per year, several million dollars larger than increased treatment costs when evaluated at industry estimates.

6.2 Policy implications

The primary implication of the results is that regulators who wish to encourage reuse should minimize barriers to sharing. The estimates described in Section 5.1 suggest a few potential channels for policy interventions to improve the sharing market in Pennsylvania specifically. First, indemnification against legal liability for firms that send wastewater to rivals could reduce the transaction costs of sharing, particularly when the recipient is a smaller regional

[35] I use the cost range for seawater reverse osmosis reported in Curto et al. (2021). This calculation ignores transportation costs as well as the indirect external costs of energy-intensive desalination processes. The difficulty of transporting water implies that water desalinated near the sea does not immediately replace lost freshwater in inland Pennsylvania.

[36] Earthquake data obtained from the Ohio Department of Natural Resources. Calculations by the author.

firms. Second, policies that increase transparency regarding the composition of fracking fluids could reduce the value of trade secrets, encouraging sharing.[37] Some observers have suggested that reducing documentation requirements for wastewater shipments could also be beneficial (Groundwater Protection Council, 2019); however, lack of transparency regarding aggregate reuse and sharing rates could exacerbate potential incentive problems within larger firms.

Given the predicted impacts of fully eliminating transaction costs shown in Table 8, policy interventions that marginally reduce transaction costs may only generate limited external benefits. Indeed, the results suggest that there could external harms from increased wastewater transportation depending on the relative importance of distance in determining matching patterns. In the fifth row of Table 8, I consider the effect of eliminating all heterogeneity in transaction costs associated with firm size.[38] This counterfactual approximates the maximum effect of policy interventions that primarily aim to ease trade between smaller and larger firms, such liability shields for wastewater-sending firms and enhanced disclosure requirements regarding the composition of fracking fluids.[39] Addressing these issues alone is predicted to have minimal effects on the rates of reuse and sharing.

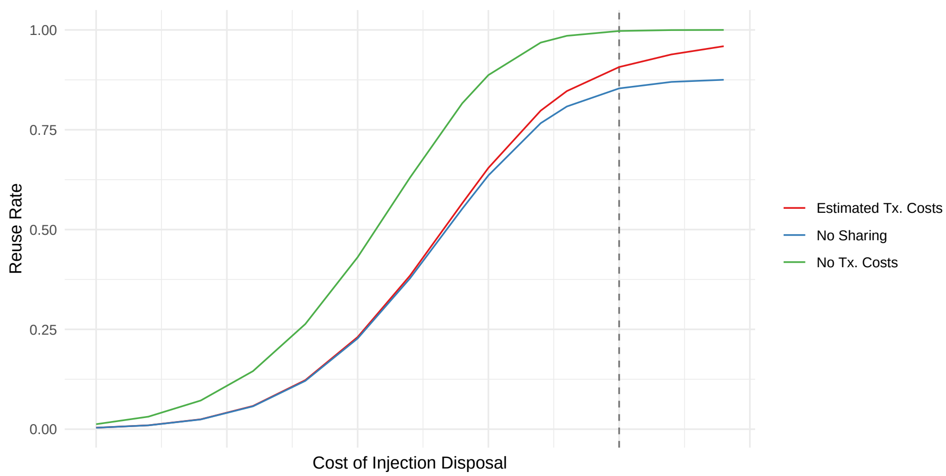
An important question is whether the magnitude of external benefits from sharing would be larger or smaller in regions where reuse rates are far lower than in Pennsylvania. Simply adopting the sharing institutions observed in Pennsylvania could be beneficial in other regions, even if transaction costs primarily arise from sources that are difficult to address with policy. To assess this possibility, Figure 4 explores the effect of changing the cost of final disposal (or, equivalently, the cost of freshwater withdrawal). The dashed vertical line indicates the estimated level. When the cost of final disposal is sufficiently low that the reuse rate is 12.5% (the national average), the estimated transaction costs exceed the cost of

[37] Currently, firms are required to report the composition of their fracking fluids in the FracFocus. Thus, there may be few trade secrets to protect. However, certain fracking fluid constitutions may be designated as "proprietary" in FracFocus.

[38] I use the smallest estimated coefficients for sending firm-type and receiving firm-type.

[39] The remaining sources of heterogeneity in transaction costs are those associated with the observed frequency of trade, firm location, and the time fixed effects.

Figure 4: Effects of changing the cost of final disposal



final disposal, making sharing unlikely. The existence of a sharing market with the estimated level of transaction costs increases the reuse rate by only 0.2 percentage points, whereas fully eliminating transaction costs would increases the reuse rate by 14 percentage points.[40] Thus, adopting institutions similar to those in Pennsylvania is likely to generate limited external benefits outside contexts in which the costs of final disposal and freshwater acquisition are already high. In such contexts, taxes on final disposal or freshwater withdrawals may be more effective policy instruments. The figure also shows that small shifts in the cost of final disposal can generate large increases in reuse rates even when sharing is not possible.

7 Conclusion

Wastewater generated by oil and gas wells creates significant local environmental challenges, which have only become more severe in recent years. Reuse mitigates these impacts. Onshore oil and gas production is characterized by fragmented market structures with many diverse firms. In this context, wastewater sharing can reduce the cost of reuse and improve

[40] An important limitation of this exercise is that I do not consider differences in wastewater production rates across regions. Pennsylvania wells produce similar volumes of wastewater as wells in most other US shale basins, with the important exception of the Permian where wells are substantially more "wet" (often producing wastewater volumes that exceed the volume of injected fracking fluid).

efficiency. I show that the effects of sharing on aggregate reuse rates and wastewater trucking intensity in Pennsylvania are large but limited by the presence of substantial transaction costs. Thus, the realized benefits of sharing are smaller in magnitude than the potential social benefits, suggesting that policy intervention to reduce transaction costs could enhance efficiency. However, the potential external benefits of increased sharing are likely to be modest despite Pennsylvania's high rates of reuse, and may be even smaller in regions where rates of reuse are currently lower. Efforts to increase all forms of reuse, such as through taxes on injection disposal and freshwater consumption, may generate greater external benefits in such regions.

References

Anderson, J.E. and E. van Wincoop (2004) "Trade Costs," *Journal of Economic Literature*, 42 (3), pp. 691–751.

Antrás, P. and D. Chor (2022) "Global Value Chains," *Handbook of International Economics*, pp. 297–376, eds. G. Gopinath, E. Helpman, and K. Rogoff. Elsevier.

Atalay, E., A. Hortaçsu, C. Syverson, and M.J. Li (2019) "How Wide is the Firm Border?" *Quarterly Journal of Economics*, 134 (4), pp. 1845–1882.

Black, K.J., A.J. Boslett, E.L. Hill, L. Ma, and S.J. McCoy (2021) "Economic, Environmental, and Health Impacts of the Fracking Boom," *Annual Review of Resource Economics*, 13, pp. 311–334.

Boomhower, J. (2019) "Drilling like there's no tomorrow: Bankruptcy, insurance, and environmental risk," *American Economic Review*, 109 (2), 391–426.

Bresnahan, T.F. and J.D. Levin (2012) "Vertical Integration and Market Structure," NBER Working Paper.

Chaney, T. (2018) "The Gravity Equation in International Trade:An Explanation," *Journal of Political Economy*, 126 (1), 150–177.

Choo, E. and A. Siow (2006) "Who Marries Whom and Why," *Journal of Political Economy*, 114 (1), pp. 175–201.

Coase, R.H. (1937) "The Nature of the Firm," *Economica*, 4 (16), pp. 386–405.

Correira, S., P. Guimaraes, and T. Zylkin (2020) "Fast Poisson estimation with high-dimensional fixed effects," *The Stata Journal* (20), 95–115.

Covert, T.R. (2015) "Experiential and Social Learning in Firms: The Case of Hydraulic Fracturing in the Bakken Shale," Working Paper.

Covert, T.R. and R.L. Sweeney (2022) "Secrecy Rules and Exploratory Investment: Theory and Evidence from the Shale Boom," Working Paper.

Curto, D., V. Franzitta, and A. Guercio (2021) "A Review of the Water Desalination Technologies," *Applied Sciences*, 11, 670.

Demsetz, H. (1988) "The Theory of the Firm Revisited," *Journal of Law, Economics, & Organization*, 4 (1), pp. 141–161.

EPA (2016) "Hydraulic Fracturing for Oil and Gas: Impacts from the Hydraulic Fracturing Water Cycle on Drinking Water Resources in the United States (Final Report)," Report of the Environmental Protection Agenecy.

Fernandez-Val, I. and M. Weidner (2016) "Individual and time effects in nonlinear panel models with large N , T," *Journal of Econometrics* (192), 291–312.

Fowlie, M.L. (2009) "Incomplete Environmental Regulation, Imperfect Competition, and Emissions Leakage," *American Economic Journal: Economic Policy*, 1 (2), pp. 71–112.

Galichon, A. and B. Salanie (2022) "Cupid's Invisible Hand: Social Surplus and Identification in Matching Models," *Review of Economic Studies*, 89 (5), pp. 2600–2629.

Ghili, S. (2022) "Network Formation and Bargaining in Vertical Markets: The Case of Narrow Networks in Health Insurance," *Marketing Science*, 41 (3), pp. 501–527.

Gretsky, N.E., J.M. Ostroy, and W.R. Zame (1992) "The Nonatomic Assignment Model," *Economic Theory*, 2 (1), pp. 103–127.

Grossman, S.J. and O.D. Hart (1986) "The Costs and Benefits of Ownership: A Theory of Vertical and Lateral Integration," *Journal of Political Economy*, 94 (4), pp. 691–719.

Groundwater Protection Council (2019) "Produced Water Report: Regulations, Current Practices, and Research Needs," Report.

GWPC, Ground Water Protection Council and Interstate Oil and Gas Compact Commission IOGCC (2021) "FracFocus," https://www.fracfocusdata.org (accessed January 23, 2021).

Hausman, C. and R. Kellogg (2015) "Welfare and Distributional Implications of Shale Gas," *Brookings Papers on Economic Activity*, pp. 71–125.

Head, K. and T. Meyer (2014) "Gravity Equations: Workhorse, Toolkit, and Cookbook," *Handbook of International Economics*, 4, pp. 131–195, eds. G. Gopinath, E. Helpman, and K. Rogoff. Elsevier.

Heo, J., P.J. Adams, and H.O. Gao (2016) "Reduced-form modeling of public health impacts of inorganic PM2.5 and precursor emissions," *Atmospheric Environment*, 137, pp. 80–89.

Ho, K. and R.S. Lee (2019) "Equilibrium Provider Networks: Bargaining and Exclusion in Health Care Markets," *American Economic Review*, 109 (2), pp. 473–522.

Horn, H. and A. Wolinsky (1988) "Bilateral Monopolies and Incentives for Merger," *The RAND Journal of Economics*, 19 (3), pp. 408–419.

Idel, M. (2016) "A review of matrix scaling and Sinkhorn's normal form for matrices and positive maps," arXiv:1609.06349v1.

Kauermann, G. and R.J. Carroll (2001) "A Note on the Efficiency of Sandwich Covariance Matrix Estimation," *Journal of the American Statistical Association*, 96 (456), 1387–1396.

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

Kondash, A.J., N.E. Lauer, and A.V. Vengosh (2018) "The intensification of the water footprint of hydraulic fracturing," *Science Advances*, 4 (8).

Koster, H.R.A. and J. van Ommeren (2015) "A shaky business: Natural gas extraction, earthquakesand house prices," *European Economic Review* (80), 120–139.

Leslie, G. (2018) "Tax induced emissions? Estimating short-run emission impacts from carbon taxation under different market structures," *Journal of Public Economics*, 167, pp. 220–239.

Luxen, D. and C. Vetter (2011) "Real-time routing with OpenStreetMap data," pp. 513–516, in Proceedings of the 19th ACM SIGSPATIAL International Conference on Advances in Geographic Information Systems. Association for Computing Machinery.

Mansur, E.T. (2007) "Do Oligopolists Pollute Less? Evidence from a Restructured Electricity Market," *The Journal of Industrial Economics*, 55 (4), pp. 661–689.

Masten, S.E., J.W. Meehan, and E.A. Snyder (1991) "The Costs of Organization," *Journal of Law, Economics, & Organization*, 7 (1), pp. 937– 976.

McCurdy, R. (2011) "Underground Injection Wells For Produced Water Disposal," in Proceedings of the Technical Workshops for the Hydraulic Fracturing Study: Water Resources Management. EPA.

Newell, R.G. and D. Raimi (2014) "Implications of Shale Gas Development for Climate Change," *Environmental Science and Technology*, 48 (15), pp. 8360–8368.

Ohio DNR, Ohio Department of Natural Resources (2024) "Ohio Earthquake Database," https://ohiodnr.gov/discover-and-learn/land-water/earthquakes/ohio-earthquake-database?=undefined (accessed November 17, 2024).

PA DEP, Pennsylvania Department of Environmental Protection (2021) "Oil and Gas Waste Reports," https://greenport.pa.gov/ReportExtracts/OG/OilGasWellWasteReport (accessed May 11, 2021).

Peyre, G. and M. Cuturi (2020) "Computational Optimal Transport," arXiv:1803.00567, https://arxiv.org/abs/1803.00567.

Preonas, L. (2023) "Market Power in Coal Shipping and Implications for U.S. Climate Policy," *The Review of Economic Studies*, rdad090.

Samuelson, L. and E. Stacchetti (2017) "Even up: Maintaining relationships," *Journal of Economic Theory*, 169, pp. 170–217.

Santos Silva, J.M.C. and S. Tenreyro (2022) "The Log of Gravity at 15," *Portuguese Economic Journal* (21), 423–437.

Schultz, R., R.J. Skoumal, M.R. Brudzinski, D. Eaton, B. Baptie, and W. Ellsworth (2020) "Hydraulic fracturingâinduced seismicity," *Reviews of Geophysics*, 58 (3).

Shapley, L.S. and M. Shubik (1971) "The assignment game I: The core," *International Journal of Game Theory*, 1, pp. 111–130.

Small, M.J. et al. (2014) "Risks and Risk Governance in Unconventional Shale Gas Development," *Environmental Science and Technology*, 48 (15), pp. 8289–8297.

SRBC, Susquehanna River Basin Commission (2021) "Post-Hydro Frac Data."

Williamson, O.E. (1971) "The Vertical Integration of Production: Market Failure Considerations," *American Economic Review*, 61 (1971), pp. 112–123.

Wilson, K. (2023) "Pennsylvania Watersheds at Risk: Drought and Fracking," *FracTracker Alliance*, https://www.fractracker.org/2023/09/pennsylvania-watersheds-at-risk-drought-and-fracking/.

Wunz, E.C. (2015) "Summary of Shale Gas Wastewater Treatment and Disposal In Pennsylvania: 2014," Report, dated August 20.

Zylkin, T. (2024) "Bootstrap for Gravity Models," *Working Paper*.

A Data preparation

The main dataset consists of Oil and Gas Well Waste Reports collected from the Pennsylvania Department of Environmental Protection web site.[41] Operators are required to report the method of disposal for various waste products, including solids such as drill cuttings and shredded containment liners. I rely on the classifications from Wunz (2015) as well as internet research on the functions performed at different waste facilities (e.g., landfills vs. injection wells) to identify presumably reusable wastewater. This procedure is inevitably imperfect. Reporting errors are possible, and not all liquid waste necessarily represents reusable wastewater.[42]

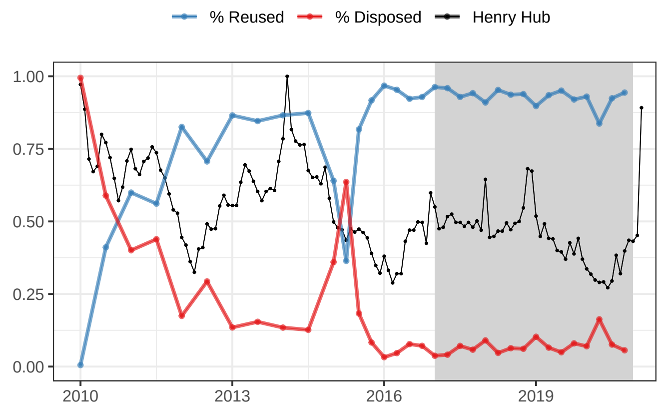
For the main analysis, I consider waste reports for all unconventional wells and for all production periods between January 2017 and December 2020. This choice of analysis period reflects the fact that the waste reporting format was modified in January 2017 to consistently indicate the location of reuse.[43] Prior to January 2017, it is possible to determine whether reuse occurred, but not whether reuse occurred internally or via sharing. Figure 5 presents a longer time series indicating the relative frequency of final disposal and reuse, along with the spot price of natural gas. The sample period corresponds to the shaded region.

[41] PA DEP (2021). Other referenced data sources include GWPC and IOGCC (2021), SRBC (2021), Ohio DNR (2024).

[42] For instance, sludges produced as a byproduct of treatment for reuse are often sent to injection wells.

[43] I choose to retain data from the COVID pandemic period. Although drilling rates in general fell during this period, the demand for disposal did not, and overall reuse rates remained relatively stable, as evidenced by Figure 5.

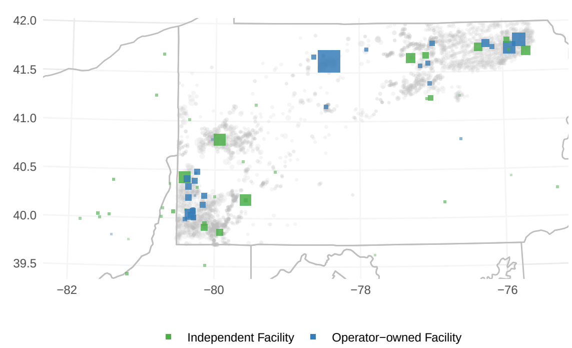
Figure 5: Pennsylvania wastewater reuse over time



Notes: The red and blue lines indicate the share of wastewater shipments in the data for which the reported destination was a site at which only disposal could have occurred (primarily injection wells), or a site at which reuse could have occurred. The black line indicates the spot price of natural gas. The sample period for this analysis is highlighted in gray.

As described in the main text, the waste reports do not report the dates or quantities associated with specific transfer events, but rather the aggregate quantities of different types of waste transferred from a given well to a given disposal location during a month. Wastewater intended for reuse can be transferred either to a CTF prior to reuse or directly to another well pad for reuse. These cases appear differently in the data. In the former case, it is not possible to identify the ultimate location of reuse. However, whether the treatment facility is operated by the reporting firm or by a third party can be inferred from the reported permit information and facility names (although in some cases this requires consulting separate DEP resources). If the destination is a well pad located in Pennsylvania, as occurs most often, a numeric identifier associated with the destination well pad is provided. I use this numeric identifier to determine whether a given amount of wastewater was transferred for internal or external reuse. In particular, I classify reuse locations as internal or external depending on whether the reporting firm is currently listed as an operator for any well at the destination well pad (in a separate DEP data source). If the destination well pad is located outside of Pennsylvania (primarily in West Virginia), no such identifier is provided, and I do not attempt to infer the ownership of the destination well pad. I identify firms by their DEP OGO Number (where OGO is an acronym for "Oil and Gas Operator"). I rely on press releases and changes in the reporting operator over time to track changes in corporate ownership (the Rice-EQT merger was the most significant but not the sole merger during the sample period). It is rare for multiple operators to be associated with the same well pad, but when this is the case I treat the well pad as "internal" for both parties.

Figure 6: Centralized treatment facility locations

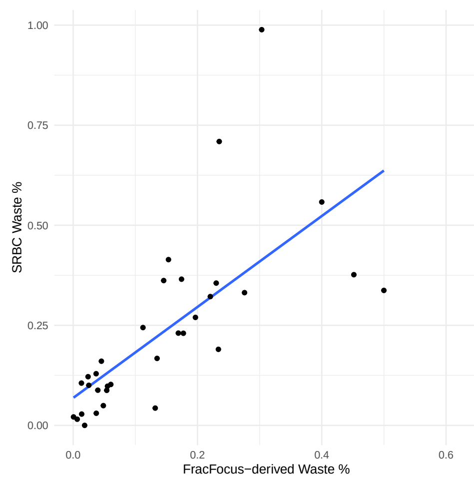


Typically several wells are located at a single well pad, which encompasses common infrastructure such as access roads and storage tanks. Technically operators are required to report waste information on a well-by-well basis, but because wastewater is often stored in a single location on the pad many simply report well pad-level averages. Therefore I focus on the well pad rather than the well as my primary unit of analysis. I infer the number of shipments in a month by dividing the total volume by the capacity of a typical water hauling truck.[44] To mitigate the impact of data reporting errors, I winsorize shipment volumes at

[44]I assume that this is 110 barrels (the modal volume), although truck capacities range from around 80 to

the 99.9%-tile (about 77,000 barrels, or 600-700 truckloads in a month).

Figure 7: Share of Wastewater Shipments by Ordinal Distance



Classification of large regional, national, and small firms I define "Large regional" firms to include Range, EQT, Seneca, Rice, Pennenergy, CNX, and SWN. These firms were each among the largest firms by wastewater disposal volume during the sample period, and none has significant operations outside Appalachai. Many, but not all, are privately held. I define "national" firms to include Cabot, Chief, XTO (an Exxon subsidiary), Chevron, Repsol, Hilcorp, SWEPI (a Shell subsidiary), Chesapeake, BKV, EOG, and Noble Energy. Each of these firms has a significant presence outside Appalachia. Many, but not all, are publicly held. A classify all remaining firms as "Small regional" firms. Note that this designation includes some affiliates of national firms with particularly small operations.

A.1 Validation of Freshwater Usage Rates

A well's freshwater usage is equal to one minus its wastewater usage. For firm-years appearing in the SRBC data, Figure 7 plots (i) the calculated wastewater usage rate based on the FracFocus "TotalBaseFluid" and disposal records against on the x-axis and (ii) the true

around 130 barrels. Line items in the data are frequently reported in integer multiples of a truck capacity in this range.

wastewater usage rate in the SRBC data on the y-axis. A linear regression yields a coefficient 1.13 (SE 0.2) and an R^2 value of 0.49.

B Model details

B.1 Match entropy function

In general, the match entropy function is defined to be:

$$\mathcal{E}(\mu, \mathbf{Q}, \mathbf{C}) = -G^*(\mu, \mathbf{Q}) - H^*(\mu, \mathbf{C}) \tag{11}$$

where $G^*(\mu, n)$ is the generalized entropy of choice for disposal and $H^*(\mu, m)$ is the generalized entropy of choice for reuse. In particular,

$$G^*(\mu, \mathbf{Q}) = \sup_{U \in \mathbb{R}^{K \times D}} \left(\sum_{\kappa \in K} \sum_{\delta \in D} \mu_{\kappa\delta} U_{\kappa\delta} - \sum_{\kappa \in K} Q_\kappa E \left[\max_{\delta \in D_0} U_{\kappa\delta} + \epsilon_{i\delta} \right] \right)$$

and

$$H^*(\mu, \mathbf{C}) = \sup_{V \in \mathbb{R}^{K \times D}} \left(\sum_{\kappa \in K} \sum_{\delta \in D} \mu_{\kappa\delta} V_{\kappa\delta} - \sum_{\delta \in D} C_\delta E \left[\max_{\kappa \in K_0} V_{\kappa\delta} + \eta_{\kappa j} \right] \right)$$

Intuitively, $G^*(\mu, \mathbf{Q})$ and $H^*(\mu, \mathbf{C})$ quantify the amount of latent cost heterogeneity required to rationalize a given match μ conditional on the distributions of P_K and P_D. Galichon and Salanie (2022) provides an interpretation of these objects.

In the case that P_K and P_D are extreme value type I distributions with scale parameters

σ_ϵ and σ_η, (11) is equivalent to:

$$\begin{aligned}\mathcal{E}\left(\mu,Q,C\right) = &\overbrace{\sigma_{\epsilon_{\sum^{\kappa\in K}}} Q_\kappa\, log\, Q_\kappa + \sigma_{\eta_{\sum^{\delta\in D}}} C_\delta\, log\, C_\delta}^{-C(Q,C)} - (\sigma_\epsilon+\sigma_\eta)\sum_{\kappa\in K}\sum_{\delta\in D}\mu_{\kappa\delta}\log\mu_{\kappa\delta}\\ &-\sigma_\epsilon\sum_{\kappa\in K}\mu_{\kappa 0}\log\mu_{\kappa 0} - \sigma_\eta\sum_{\delta\in D}\mu_{0\delta}\log\mu_{0\delta}\end{aligned}$$

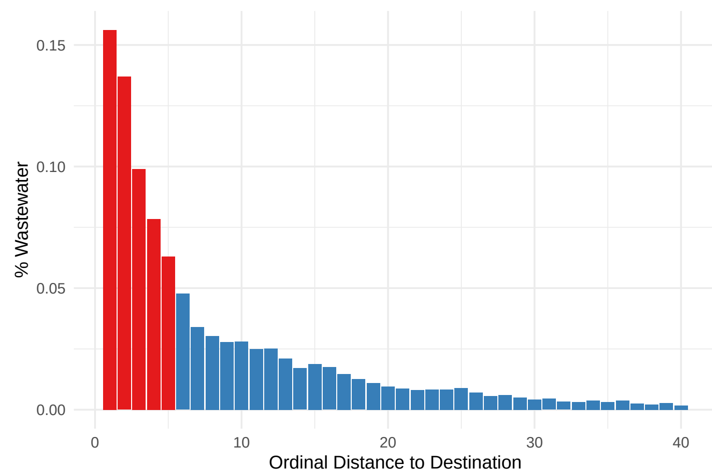
which is the expression in (4).

C Additional Results (for online publication)

C.1 Sharing probabilities by ordinal distance

Figure 8 shows the share of wastewater shipped by ordinal distance to the destination. Thus, an ordinal distance of 1 corresponds to the nearest facility that accepted any wastewater (in terms of over-the-road shipment distance). An ordinal distance of 2 corresponds to the second nearest facility that accepted any wastewater.

Figure 8: Share of Wastewater Shipments by Ordinal Distance



C.2 Alternative Distance Measures

Table 9 reports results from the same specifications considered in Table 5, except using a measure of driving time (in hours) rather than over-the-road distance. The point estimates in Column (6) imply that crossing the firm boundary (in other words, sharing) raises the total costs of reuse by an amount equivalent to the cost of shipping a truckload of wastewater $\frac{4.863}{1.512} = 3.2$ additional hours.

Table 9: Transaction Cost Estimates

	Dependent Variable: # of Truckloads					
	(1)	(2)	(3)	(4)	(5)	(6)
Duration	−2.484	−1.494	−1.414	−2.734	−1.769	−1.512
	(0.008)	(0.004)	(0.009)	(0.013)	(0.007)	(0.014)
Firm Boundary		−5.665	−4.812		−5.771	−4.863
		(0.036)	(0.036)		(0.033)	(0.036)
EQT-Rice Post-Merger		−0.005	0.003		−0.135	−0.133
		(0.020)	(0.022)		(0.021)	(0.022)
Distance Measure	Log	Log	Log	Linear	Linear	Linear
κ-month FEs	Yes	Yes	Yes	Yes	Yes	Yes
δ-month FEs	Yes	Yes	Yes	Yes	Yes	Yes
Outside Options	No	No	Yes	No	No	Yes
Pseudo R^2	0.664	0.884	0.874	0.676	0.897	0.887
Observations	3,775,553	3,775,553	3,843,919	3,775,553	3,775,553	3,843,919

C.3 Main specification

Table 10 reports parameter estimates for the full model. The distance elasticity is equal to the one reported in Column (6) of Table 5. The signs of the estimated coefficients are consistent with the discussion in Section 5.1. Estimates for the firm boundary-quarter fixed effects are presented separately in Figure 9. The excluded category is the first quarter of 2017. Less negative values of the fixed effect correspond to greater transaction costs. Hence, the figure indicates that transaction costs increased sharply beginning in 2018.

C.4 Robustness to Alternative γ Values

Table 11 presents the main counterfactuals from Table 8 obtained under three alternative assumptions for the value of γ. Recall that the main estimates are obtained under the assumption that $\gamma = 0.5$, reproduced for convenience in Sub-Table 11b. For greater values of γ, estimated disposal rates are higher and shipping distances are lower.

Figure 9: Firm Boundary × Quarter Fixed Effects

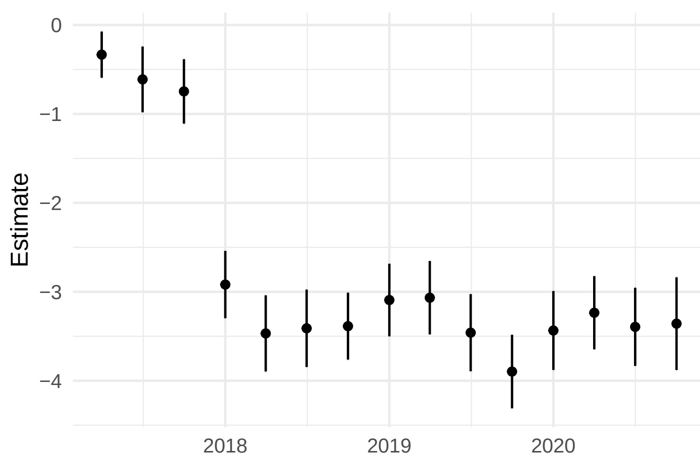


Table 10: Full Model

	Dependent Variable: # of Truckloads
Distance	-0.039
	(0.0004)
Firm Boundary	-4.524
	(0.109)
EQT-Rice Post-Merger	-0.174
	(0.023)
Firm Boundary:From National Firm	0.727
	(0.087)
Firm Boundary:From Small Firm	0.898
	(0.078)
Firm Boundary:To National Firm	0.024
	(0.109)
Firm Boundary:To Small Firm	0.404
	(0.105)
Firm Boundary:From Northwest	0.658
	(0.086)
Firm Boundary:To Northwest	-1.152
	(0.120)
Firm Boundary:Shared Last Month	4.179
	(0.141)
Firm Boundary:Shared Last Year	2.399
	(0.149)
Distance Measure	Linear
Firm Boundary $\times$ Quarter FEs	Yes
κ-month FEs	Yes
δ-month FEs	Yes
Outside Options	Yes
Pseudo R^2	0.892
Observations	3,843,919

Table 11: Counterfactuals with Alternative γ Values

(a) $\gamma = 0.35$

	Mean Distance (mi)				Sharing Rate (%)				Disposal Rate (%)			
	Large	Nat'l	Small	All	Large	Nat'l	Small	All	Large	Nat'l	Small	All
Model	28.5	32.7	53.6	31.9	3.0	14.9	34.3	8.2	6.3	22.1	29.4	11.3
Transaction costs $\to \infty$	29.8	46.5	90.8	38.9	-	-	-	-	7.3	32.8	58.1	16.8
Transaction costs $\to 0$	32.2	37.6	44.6	34.4	56.9	79.9	88.0	63.8	0.5	0.4	0.5	0.5
Transaction costs $\to 0$, $\mathcal{E} \to 0$	18.8	23.8	28.6	20.6	38.6	61.6	74.5	46.1	0.5	0.4	0.5	0.5

(b) $\gamma = 0.50$

	Mean Distance (mi)				Sharing Rate (%)				Disposal Rate (%)			
	Large	Nat'l	Small	All	Large	Nat'l	Small	All	Large	Nat'l	Small	All
Model	26.4	31.6	51.6	29.9	2.8	13.0	32.4	7.6	3.9	21.5	28.7	9.3
Transaction costs $\to \infty$	27.8	45.2	89.5	37.1	-	-	-	-	4.8	31.8	57.2	14.6
Transaction costs $\to 0$	31.3	36.6	42.9	33.3	56.5	80.1	88.1	63.6	0.3	0.2	0.3	0.3
Transaction costs $\to 0$, $\mathcal{E} \to 0$	18.1	22.5	27.1	19.7	38.5	61.0	74.4	45.9	0.3	0.2	0.3	0.3

(c) $\gamma = 0.65$

	Mean Distance (mi)				Sharing Rate (%)				Disposal Rate (%)			
	Large	Nat'l	Small	All	Large	Nat'l	Small	All	Large	Nat'l	Small	All
Model	25.2	31.1	50.3	28.8	2.7	11.7	31.2	7.1	2.6	21.1	28.0	8.2
Transaction costs $\to \infty$	26.7	44.6	88.9	36.1	-	-	-	-	3.5	31.2	56.8	13.5
Transaction costs $\to 0$	30.4	35.6	41.1	32.4	56.3	80.4	88.0	63.5	0.2	0.2	0.2	0.2
Transaction costs $\to 0$, $\mathcal{E} \to 0$	17.5	21.8	25.6	19.1	38.4	62.0	74.6	46.0	0.2	0.2	0.2	0.2