Firm Boundaries and External Costs in Shale Gas Production

Matthew O'Keefe (Northwestern University)

January 2024

When do markets form?

- Coase (1937, 1960): when **transaction costs** are low enough
- Large empirical literature testing specific predictions (e.g., asset specificity)
- Significant empirical challenges $\Rightarrow$ limited work on quantifying transaction costs
  - Wallis and North (1986), Masten et al (1991), Atalay et al (2019), ...
- Little exploration of the ***distribution*** of transaction costs within markets
  - Demsetz (1988): transaction costs might vary across counterparties, across time

This paper: new evidence from wastewater sharing in Pennsylvania

- Fracking a shale gas well requires 10-20M gallons of water at once
- Over time, 25% returns as wastewater (along with the hydrocarbons)
- Firms **reuse** wastewater as a substitute for freshwater to reduce costs
  - Can save up to $1M for a typical well ($0.25M freshwater, $0.75M final disposal)
- Efficiencies $\Rightarrow$ approx. 10% of reuse occurs via **sharing** – trade between rival firms
  - Trade enables more efficient matching from old wells to new wells

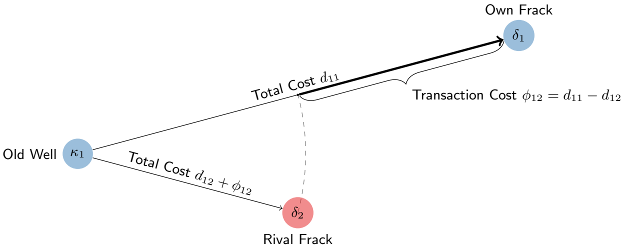
This paper

- Like all market transactions, sharing is subject to **transaction costs**
  - *Transaction costs:* any costs foregone under integration (i.e., if firms merged)
- Three questions:

  1. How large are the transaction costs of sharing?
  2. What are the main sources of transaction costs?
  3. What are the environmental impacts of transaction costs?

Quantifying transaction costs

- Water is heavy $\Rightarrow$ transporting wastewater is costly (typically, trucked)
- Data: wastewater shipments *within* and *between* firms, at high spatial resolution
- Idea: transaction costs $\equiv$ "distance premia" firms incur to avoid sharing



Key findings

1. Transaction costs are large, but heterogenous
    - \$6/bbl mean across transactions, \$2/bbl standard deviation
2. Transaction costs...
    - Are greater for riskier types of wastewater
    - Are greater for counterparties with poor environmental records
    - Vary significantly across firm-pairs
3. In Pennsylvania, environmental impacts are limited
    - Freshwater consumption decreases, but transportation increases

Related literature

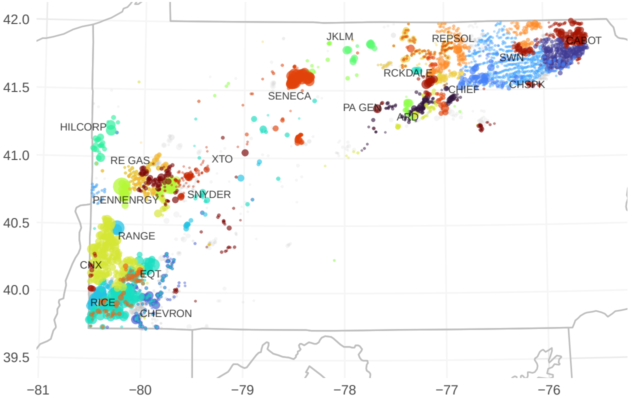
- Quantification of the Coasean transaction costs
  - Masten Meehan Snyder 1991, Atalay Hortacsu Li Syverson 2019, ...
  - Contribution: richer within-market evidence
- Direct environmental impacts of fracking
  - Hausman and Kellogg 2015, Black et al 2021, ...
  - Contribution: novel empirical framework for policy evaluation
- Regulation of environmental externalities in oligopoly
  - Mansur 2007, Fowlie 2009, Ryan 2012, Fowlie et al 2016, Leslie 2018, Preonas 2023, ...
  - Contribution: complementary source of market imperfection

Data: wastewater disposal records

- Monthly disposal records by well pad from Pennsylvania DEP, 2017-20
- What it has:
  - Monthly transfer volumes for **all** well pads / destinations
  - Detailed facility information (precise locations, permit numbers, ...)
- What it doesn't have:
  - Dates, times, or modes of particular shipments
  - Contract terms
  - Prices
- Supplementary data: completion info from FracFocus (incl. fracking inputs)

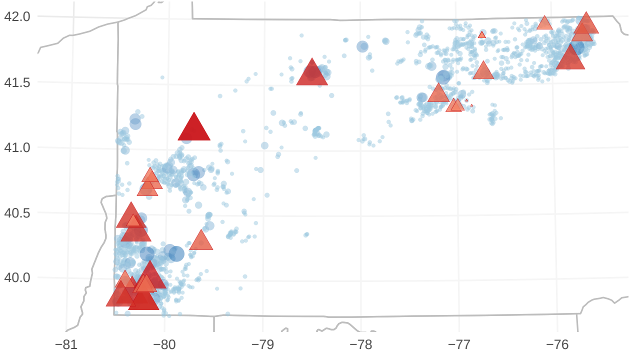
Fact 1: decentralized production

- Wastewater disposal HHI: 1,090
- Locations of twenty largest firms (by disposal volume):



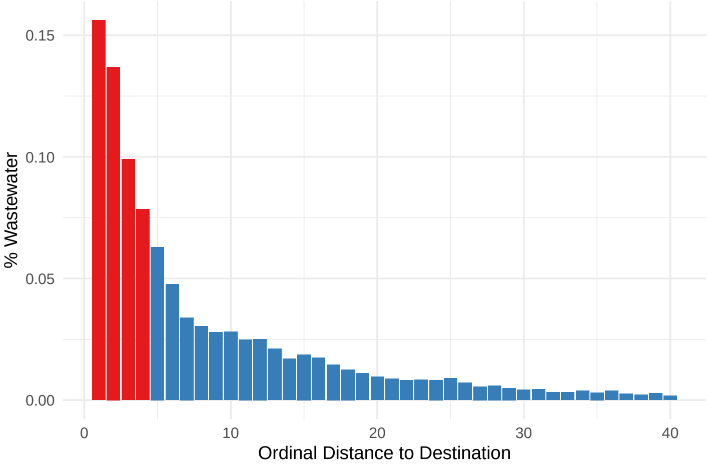
Fact 2: more old wells than new wells

- Average month: 1,721 well pads reporting disposal vs. 55 completions
- June 2018 (well pads reporting disposal in blue, completions in red):



Fact 3: most reuse occurs locally

- 47% of wastewater is shipped to one of the four nearest destinations:



Prevalence of wastewater sharing

- Disposal market shares:

Outcome	Facility	% Vol.
Reuse	Own well/CTF	70.6
	Independent CTF	12.5
	Rival well/CTF	**8.6**
Final Disposal	Injection Well	8.4

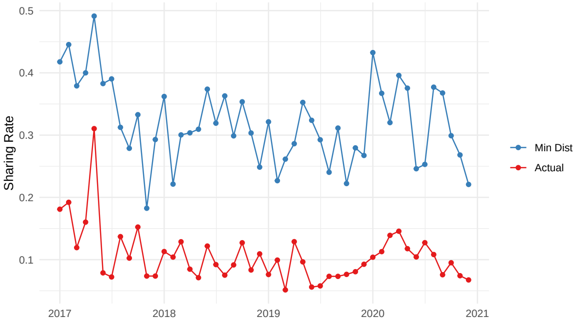
- Three reasons for sharing:

1. Temporal mismatches
2. Geographic synergies
3. Non-geographic synergies

Final Disposal

Is there enough sharing?

- Actual sharing rate (10.5%) is lower than distance-minimizing rate (31.9%):

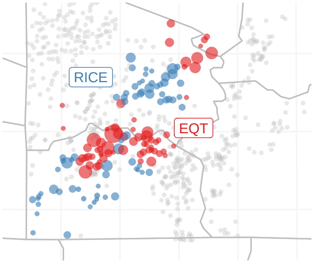


- Why? Either (1) transaction costs; or (2) technological incompatibility

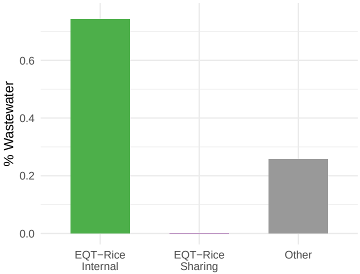
Evidence of transaction costs: 2017 EQT-Rice merger

- EQT and Rice merger created largest gas producer in US

Pre-merger well pads



Pre-merger disposal shares

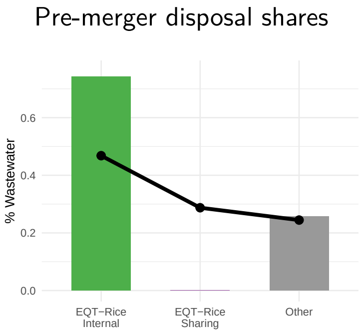


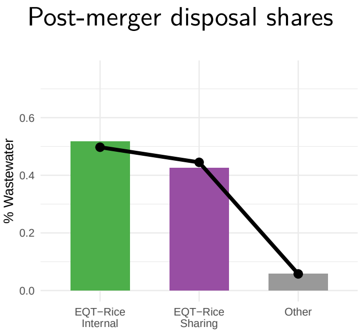
- Large pre-merger transportation synergies, but no pre-merger sharing

Evidence of transaction costs: 2017 EQT-Rice merger (cont)

- Post-merger "sharing" rate 43%

Evidence of transaction costs: 2017 EQT-Rice merger (cont)





- Post-merger "sharing" rate 43%, close to **model-implied optimum**
  - Suggests *firm boundaries* matter, not technological incompatibilities

How large are transaction costs?

- K is the set of old well pads in month t
- D is the set of new wells in month t
- A **transaction** is a shipment from an old well pad to a new well
  - One transaction = one *truckload* of wastewater (110 barrels)
- $r^K_{\kappa\delta}$ and $r^D_{\kappa\delta}$ are the costs of reusing wastewater from κ at δ, where:

$$r_{\kappa\delta} = r^K_{\kappa\delta} + r^D_{\kappa\delta} = \begin{cases} \text{Technological Cost} & \kappa \text{ and } \delta \text{ in same firm} \\ \text{Technological Cost} + \text{Transaction Cost} & \kappa \text{ and } \delta \text{ in rival firms} \end{cases}$$

  - Technological costs: transportation, treatment, labor, ...

Empirical strategy: matching with transferable utility

1. Truckloads of wastewater matched from old well pads K to new wells D

2. Shipment/receipt decisions are made **locally** in response to **prices:**

$$r_{\kappa\delta} = \overbrace{r^K_{\kappa\delta} + \tau_{\kappa\delta}}^{\text{Sender's Costs}} + \overbrace{r^D_{\kappa\delta} - \tau_{\kappa\delta}}^{\text{Reciver's Costs}}$$

   - τ: intra-firm transfer prices, inter-firm market prices, determined **in equilibrium**

3. $r_{\kappa\delta}$ identified from observed match (Galichon and Salanie 2022)
   - Identification in terms of **shipment distances**
   - Within-firm shipments $\Rightarrow$ technological costs
   - Between-firm shipments $\Rightarrow$ transaction costs

Supply: wastewater disposal as a discrete choice problem

- Q_κ truckloads of wastewater are generated at κ
- The operator at κ ships ith truckload to the least cost destination δ:

$$\delta^* = \arg\min_{\delta \in D_0} \quad r^K_{\kappa\delta} + \tau_{\kappa\delta} - \epsilon_{i\delta}$$

  - $D_0 = D \cup \{0\}$ includes all new wells and the outside option (final disposal)
  - $r^K_{\kappa\delta} + \tau_{\kappa\delta}$ is the sender's share of the joint costs of reuse at δ
  - $\epsilon_{i\delta}$ is a truckload-specific, non-systematic latent cost (EV Type 1, dispersion σ_K)

Demand: water acquisition as a discrete choice problem

- C_δ truckloads of water (wastewater or freshwater) are needed at δ
- The operator at δ accepts jth truckload from the least cost source:

$$\kappa^* = \arg\min_{\kappa \in K_0} \quad r^D_{\kappa\delta} - \tau_{\kappa\delta} - \eta_{\kappa j}$$

  - $K_0 = K \cup \{0\}$ includes all producing wells and the outside option (freshwater)
  - $r^D_{\kappa\delta} - \tau_{\kappa\delta}$ is the receiver's share of the joint costs of reusing wastewater from κ
  - $\eta_{\kappa j}$ is a truckload-specific, non-systematic latent cost (EV Type 1, dispersion σ_D)

Equilibrium: local supply = local demand

- Equilibrium is characterized by utility transfers τ and shipments μ
  - $\mu_{\kappa\delta}$ is the number of truckloads expected to be shipped from κ to δ
- Markets clear **in expectation** for each $\kappa \in K$ and $\delta \in D$:

$$\begin{aligned}
\mu^*_{\kappa\delta} &= Q_\kappa \times P\left(\delta = \arg\min_{\delta' \in D_0} r^K_{\kappa\delta'} + \tau_{\kappa\delta'} - \epsilon_{i\delta'}\right) \quad \text{Supply of } \kappa\text{-trucks to } \delta \\
&= C_\delta \times P\left(\kappa = \arg\min_{\kappa' \in K_0} r^D_{\kappa'\delta} - \tau_{\kappa'\delta} - \eta_{\kappa' j}\right) \quad \text{Demand for } \kappa\text{-trucks at } \delta
\end{aligned}$$

Equilibrium as a convex program

- Galichon and Salanie (2022): the unique equilibrium μ^* satisfies:

$$\begin{aligned} \min_{\mu \geq 0} \quad & \sum_{\kappa \in K} \sum_{\delta \in D} \mu_{\kappa\delta} r_{\kappa\delta} + \mathcal{E}\left(\mu, \mathbf{Q}, \mathbf{C}\right) \\ \text{s.t.} \quad & \sum_{\delta \in D} \mu_{\kappa\delta} \leq Q_\delta \quad \forall\, \kappa \in K \\ & \sum_{\kappa \in K} \mu_{\kappa\delta} \leq C_\delta \quad \forall\, \delta \in D \end{aligned}$$

  - $\mathcal{E}$ is a convex ***match entropy*** function that depends on distributions of ϵ and η

Details

Parameterization

$$r_{\kappa\delta} = \begin{cases} \overbrace{d_{\kappa\delta} + x'_{\kappa\delta}\beta + \xi_\kappa^{\mathcal{I}} + \xi_\delta^{\mathcal{I}}}^{\text{Technological Cost}} & \kappa \text{ and } \delta \text{ in same firm} \\ d_{\kappa\delta} + x'_{\kappa\delta}\beta + \xi_\kappa^{\mathcal{I}} + \xi_\delta^{\mathcal{I}} + \underbrace{z'_{\kappa\delta}\alpha + \pi_b}_{\text{Transaction Cost}} & \kappa \text{ and } \delta \text{ in rival firms} \end{cases}$$

- $d_{\kappa\delta}$ represents distance-related costs (over-the-road shipping distance)
- $x_{\kappa\delta}$ is a vector of observable transaction characteristics (e.g., fluid composition)
- $\xi_\kappa^{\mathcal{I}}$ and $\xi_\delta^{\mathcal{I}}$ are unobserved, additively separable costs of reuse within the firm
- $z_{\kappa\delta}$ is a vector of transaction characteristics (e.g., facility-type interactions)
- π_b is an unobserved friction for firm pair b

Identification

- If $\mathcal{E}$ is known, then r is identified by its gradient at the equilibrium match μ^*:

$$r_{\kappa\delta} - \frac{\partial \mathcal{E}\left(\mu^*, \mathbf{Q}, \mathbf{C}\right)}{\partial \mu_{\kappa\delta}} = 0$$

- Data reveals $\mu^* \Rightarrow$ system of $|K| \times |D|$ linear equations:

$$d_{\kappa\delta} + x'_{\kappa\delta}\beta + \xi^{\mathcal{I}}_{\kappa} + \xi^{\mathcal{I}}_{\delta} + z'_{\kappa\delta}\alpha + \pi_b - \sigma_K \log\left\{\frac{\mu_{\kappa 0}}{\mu_{\kappa\delta}}\right\} - \sigma_D \log\left\{\frac{\mu_{0\delta}}{\mu_{\kappa\delta}}\right\} = 0$$

- β, ξ, α, π, σ are identified if system is invertible

- In practice: $\mu_{0\delta}$ is poorly observed $\Rightarrow$ partial identification of ξ and σ

Estimation

- For a given $\theta = (\beta, \xi, \alpha, \pi, \sigma)$, can compute the equilibrium match $\mu(\theta)$
  - $\mu_{\kappa\delta}(\theta) \propto$ equilibrium prob. of observing a shipment from κ to δ
- Maximum likelihood estimator (without outside options):

$$\hat{\theta} = \arg\max_{\theta \in \Theta} \sum_{\kappa \in K} \sum_{\delta \in D} \hat{\mu}_{\kappa\delta} \log \left(\frac{\mu_{\kappa\delta}(\theta)}{\sum_{\kappa\delta} \mu_{\kappa\delta}(\theta)} \right)$$

  - Implemented similarly to BLP (Conlon and Gortmaker 2020)
  - Standard MLE inference (one observation = one truckload)
- In practice: pool data from many markets (one month = one market)
  - Assumption: β, α, π, σ are fixed across months, while ξ adjusts (with facility age)

Estimated transaction costs

- Summary stats (μ-weighted, $N \approx 1.3$M):

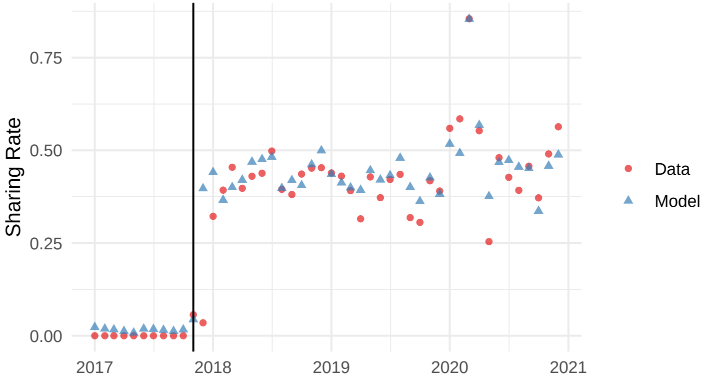
	Est (miles)	SE	\$/bbl
Mean $z'_{\kappa\delta}\alpha + \pi_b$	125.7	0.07	5.71
Std. Dev.	48.1	0.09	2.18

- The mean transaction cost is equivalent to shipping a truck 125.7 extra miles
- \$5/mile trucking costs $\Rightarrow$ \$5.71/bbl or \$26.3M/year
  - Approx. 67% of typical conventional disposal costs (roughly \$9/bbl)
  - Approx. 5% of "all in" water costs (from sourcing to disposal, roughly \$500M/year)

Model Fit Dispersion Inverse CDF Full Estimates

Model validation: EQT-Rice merger

- EQT-Rice “sharing” rate pre- and post-merger



Sources of transaction costs

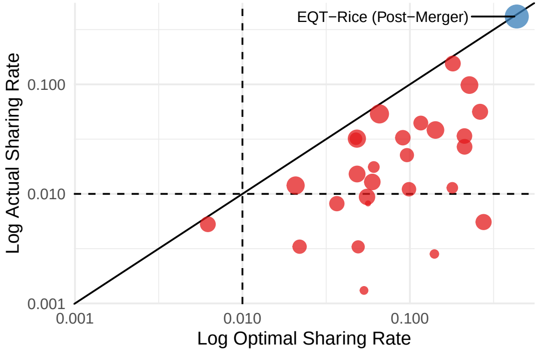
	Est (miles)	SE	\$/bbl
Sharing market cost shifters α			
rival $\times$ poor $\rightarrow$ good env record	-	-	-
rival $\times$ good $\rightarrow$ poor env record	8.5	0.11	0.39
rival $\times$ gel $\rightarrow$ slickwater	-28.6	0.10	-1.30
rival $\times$ slickwater $\rightarrow$ gel	85.3	3.00	3.88
...			
Relationship fixed effects π_b			
mean	117.9	0.07	5.36
std dev	49.2	0.09	2.23

- Interpretation: evidence of *contracting frictions...*
  - Inter-operator environmental liability, information frictions, relationship dynamics

Full Estimates

Limited trade within relationships

- Actual vs. no-friction bilateral sharing rates:



- Evidence of dynamic frictions? Ex ante coordination vs. ex post opportunism
  - Difficult to communicate future fracking plans, commit to delivery schedules

Policy implications

1. To encourage sharing, improve the contracting environment
   - Liability rules / shields (e.g., Oklahoma)
   - Disclosure of wastewater composition
   - Public pre-registration of fracking activity
   - ...
2. Interventions that ignore contracting fundamentals may fail
   - Digital platforms (i.e., Uber for wastewater)

Setting

Model

Estimates

Externalities

Conclusion

Environmental externalities

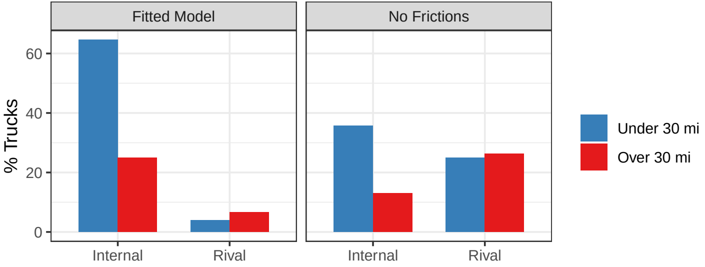
- Potential benefits from improved sharing markets:
  1. Less freshwater usage, less final disposal
  2. Less transportation
- 90% reuse $\Rightarrow$ limited scope for improvement on 1
  - Max reduction in freshwater consumption is approx. 580 acre-feet/year
    - Social value on the order of \$1M/year, valued at desalination costs
  - Difficult to quantify social costs of final disposal

External costs of transportation

- In PA: nearly all wastewater transported via heavy water-hauling trucks
  - 500,000 truck trips each year, 30.0 miles per truckload
- Unpriced transportation externalities are roughly \$7M per year
  - \$3.4M CO2 (EPA Social Cost of Carbon); \$3.3M NOx, PM2.5 (EASIUR)
  - Not included: at least 1-2 trucking-related wastewater spills per year
- In comparison, private transportation costs are roughly \$100M per year

Transaction costs and equilibrium transportation

- Fitted model vs. counterfactual with no transaction costs:



1. Sharing rate increases from approx. 10% to approx. 50%
2. Mean shipment distance increases by 15%
    - $\Rightarrow$ transportation externalities **increase** by roughly \$1M/year

Illustration Details

Implications for optimal regulation

- Net environmental spillovers from transaction costs are likely modest (in PA...)
  - Potentially even positive! If transportation externalities are large
- Nevertheless, Pigouvian intervention could be justified by private cost savings
  - Social costs = external costs + private costs
  - Roughly \$50M/year in excess private costs (direct + indirect)
- In paper: show Pigouvian program can entail large sharing subsidies
  - Depending on interpretation of transaction costs (Pareto relevance?)

Details

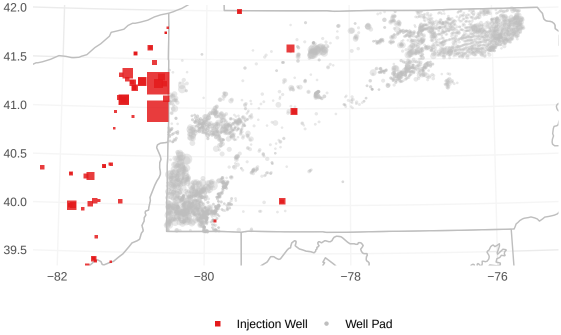
Conclusion

- New evidence on transaction costs from a unique setting
  1. Transaction costs are large, but heterogeneous
  2. Significant variation across counterparties
  3. Contracting frictions appear to play an outsized role
- Generic empirical framework for wastewater policy evaluation
- Environmental spillovers from imperfect sharing markets are modest in PA, but...
  - Rates of reuse significantly lower in Permian, Bakken, ...
  - Largest shale plays globally are not yet well developed (Vaca Muerta, Sichuan Basin)

- Thank you!
- Questions/comments: mfokeefe@u.northwestern.edu

Conventional disposal and reuse

- In Pennsylvania, injection well capacity severely limited by geology, regulation



- Due to high transportation costs, 89% of wastewater is **reused** in subsequent fracking
  - Minimal treatment required $\Rightarrow$ cost of reuse $\approx$ cost of transportation

Back

Full estimates

	Est	SE	\$/bbl
Mean $\phi_{\kappa\delta}$			
weighted by data	125.7	0.072	5.71
weighted by benchmark	154.2	0.081	7.01
Sharing market cost shifters α			
rival $\times$ poor $\rightarrow$ good env record	-	-	-
rival $\times$ good $\rightarrow$ poor env record	8.5	0.110	0.39
rival $\times$ gel $\rightarrow$ slickwater	-28.6	0.103	-1.30
rival $\times$ slickwater $\rightarrow$ gel	85.3	2.996	3.88
rival $\times$ large κ $\rightarrow$ well pad	-	-	-
rival $\times$ large κ $\rightarrow$ CTF	25.2	0.044	1.15
rival $\times$ small κ $\rightarrow$ well pad	4.4	0.151	0.20
rival $\times$ small κ $\rightarrow$ CTF	29.6	0.261	1.35
Within-firm cost shifters β			
gel $\rightarrow$ slickwater	6.7	0.092	0.31
slickwater $\rightarrow$ gel	-8.7	0.046	-0.39
small κ $\rightarrow$ CTF	-5.7	0.129	-0.26
$\sigma_\kappa + \sigma_\delta$	22.5	0.006	1.02

Sources Mean

Match entropy function

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

- $G^*(\mu, n)$ is the generalized entropy of choice for disposal

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

- $H^*(\mu, m)$ is the generalized entropy of choice for reuse

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

Logit Back

Match entropy function (cont)

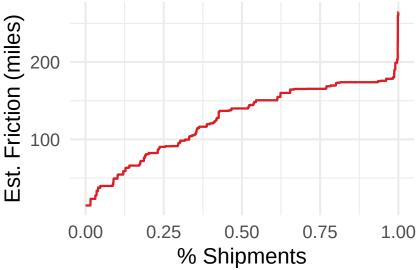
- For ϵ and η EV type 1,

$$\mathcal{E}\left(\mu, \mathbf{Q}, \mathbf{C}\right) = -\sum_{\kappa,\delta} \mu_{\kappa\delta} \left\{ \sigma_K \log\left(\frac{\mu_{\kappa\delta}}{Q_\kappa}\right) + \sigma_D \log\left(\frac{\mu_{\kappa\delta}}{C_\delta}\right) \right\}$$
$$- \sigma_K \sum_k \mu_{\kappa 0} \log\left(\frac{\mu_{\kappa 0}}{Q_\kappa}\right) - \sigma_D \sum_\delta \mu_{0\delta} \log\left(\frac{\mu_{0\delta}}{C_\delta}\right)$$

Back

Estimated transaction cost distribution

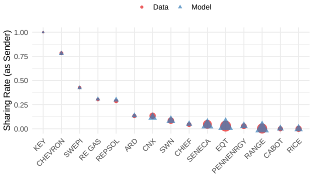
- Inverse CDF (μ-weighted):



Back

Model fit

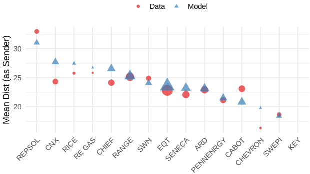
Sharing Rate (as sender, by firm)



Sharing Rate (as receiver, by firm)

Mean Dist (as sender, by firm)

Mean Dist (as receiver, by firm)



Back

Dispersion estimates

- Point estimate for dispersion:

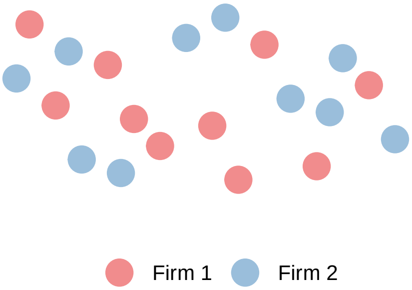
	Est (miles)	SE	\$/bbl
$\sigma_K + \sigma_D$	22.5	0.01	1.02

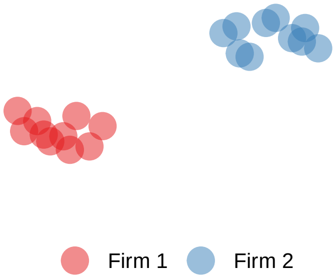
- Counterfactuals:

	Mean Dist (mi)	Share %
Data	24.86	10.60
Fitted model	24.86	10.58
$\sigma_K + \sigma_D \to 0$	21.61	9.72
$\sigma_K + \sigma_D \to \infty$	146.99	84.37

Back

Ambiguous effects of reducing transaction costs

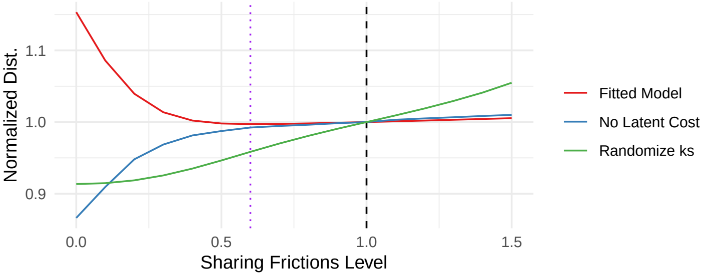




Back

Transaction costs and transportation

- Proportional re-scaling of transaction costs:



- In fitted model, removing transaction costs *increases* shipment distance by 15%
- Why? Non-random distribution of firms + matching on non-transport costs
  - As ϕ shrinks, marginal matches tend to be further away

Back

Pigouvian regulation

- Socially optimal (Pigouvian) shipment plan μ^* solves:

$$\min_{\mu \in \mathcal{M}(Q,C)} \Gamma(\mu) + C(\mu)$$

  - $\Gamma(\mu)$ represents external costs under shipment plan μ
  - $C(\mu)$ represents private costs under μ

- **Question:** should sharing frictions ϕ count towards $C(\mu)$?

Are sharing frictions welfare-relevant?

- Familiar distinction from consumer markets (e.g., switching costs)
  - Some "costs" may be relevant to decisionmakers, but not the social planner
- Examples of welfare-relevant sharing frictions:
  - Wages expended in finding out about sharing opportunities
  - Wages expended in haggling / bargaining / price discovery
  - Quantifiable risks to future profits (e.g., risk of lawsuits)
  - ...
- Examples of welfare-irrelevant sharing frictions:
  - Managerial inattention / status quo bias, loss aversion, excessive secrecy, ...

Pigouvian regulation (cont)

- Let $s \in [0,1]$ index the welfare-relevance of sharing frictions:
- $s\phi$ is welfare-relevant and $(1-s)\,\phi$ is not
  - $s = 0$ if sharing frictions are entirely welfare-irrelevant
  - $s = 1$ if sharing frictions are entirely welfare-relevant
- Socially optimal (Pigouvian) shipment plan μ_s^* solves:

$$\min_{\mu \in \mathcal{M}(Q,C)} \Gamma(\mu) + C_s(\mu)$$

  - $\Gamma(\mu)$ represents external costs under shipment plan μ
  - $C_s(\mu)$ represents welfare-relevant component of private costs under μ

Pigouvian tax rates

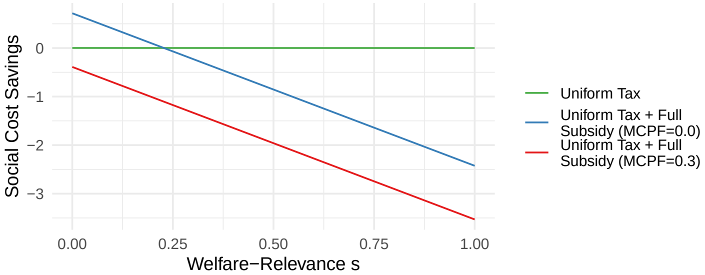
- Hypothetical policy response: Pigouvian tax on truck-miles
- Holding volume fixed, μ^* can be implemented with a tax on truck-miles:

$$tax_{\kappa\delta}^{(s)} = \gamma - (1 - s)\, d_{\kappa\delta}^{-1} \phi_{\kappa\delta}$$

  - γ is the marginal external cost of trucking (calibrate to 7%)
  - If $s < 1$, uniform tax + sharing subsidies is optimal
  - If $s = 1$, uniform tax is optimal
- Two inference problems: for optimal tax, regulator needs to know $\phi_{\kappa\delta}$ **and** s
  - In many settings $\phi_{\kappa\delta}$ (or an equivalent parameter) is identified, but s is not
    - Standard practice: argue $s = 0$ or $s = 1$ is more correct, check robustness
  - Even if firms knew s, would have incentives to shade (for larger subsidies)

Social cost savings and regret

- Change in social costs vs. status quo (\$/bbl):



- Sharing subsidies can reduce social costs by \$0.72 per barrel vs. uniform tax, **but:**
  1. Unnecessary subsidies can increase social costs by \$2.43 per barrel (before MCPF)
  2. Not cost-effective for reasonable MCPF values
  3. External costs are increased by 13.5% (not shown)

Back