Endogenous Spatial Production Networks

Quantitative Implications for Trade and Productivity

Piyush Panigrahi
Yale University

December 31, 2021

Introduction

Heterogeneous Firms, Production Networks, and Trade

- Production is organized in large-scale firm-to-firm networks
  - firms are vastly heterogeneous in size, input sourcing and importance in network
  - firms' outcomes are shaped by those of connected firms – suppliers and customers
  - supply chain networks span across space $\rightarrow$ trade costs affect network formation
  - production networks reorganize endogenously in response to shocks
- Objective
  - Design data generating process for large spatial supply chain networks
    - feasibly estimable weighted directed random graph model
  - Evaluate GE impact of micro- and macro- shocks to spatial network economy
    - e.g. firm-level distortions; market integration; technology improvements

- Production is organized in large-scale firm-to-firm networks
  - firms are vastly heterogeneous in size, input sourcing and importance in network
  - firms' outcomes are shaped by those of connected firms – suppliers and customers
  - supply chain networks span across space $\rightarrow$ trade costs affect network formation
  - production networks reorganize endogenously in response to shocks
- Objective
  - Design data generating process for large spatial supply chain networks
    - feasibly estimable weighted directed random graph model
  - Evaluate GE impact of micro- and macro- shocks to spatial network economy
    - e.g. firm-level distortions; market integration; technology improvements

Introduction

This paper

1. Document importance of endogenous networks in firm size heterogeneity
   - Indian firm network micro-data $\rightarrow$ choice of suppliers & intensity of use explain 80%
2. Develop tractable firm network formation model where firm heterogeneity $\rightleftarrows$ trade
   - rationalizes firm-to-firm network data and accommodates gravity relationships
3. Propose scalable framework for estimation + counterfactual analysis
   - maximum likelihood estimation + no simulation for counterfactuals
4. Evaluate impact of reducing inter-state border frictions by 10%
   - sizable district-level welfare gains [1%,8%]
   - $>$ 1/2 changes in firms' input sales from endogenous network changes

Introduction
This paper

1. Document importance of endogenous networks in firm size heterogeneity
   - Indian firm network micro-data $\rightarrow$ choice of suppliers & intensity of use explain 80%
2. Develop tractable firm network formation model where firm heterogeneity $\rightleftarrows$ trade
   - rationalizes firm-to-firm network data and accommodates gravity relationships
3. Propose scalable framework for estimation + counterfactual analysis
   - maximum likelihood estimation + no simulation for counterfactuals
4. Evaluate impact of reducing inter-state border frictions by 10%
   - sizable district-level welfare gains [1%,8%]
   - $>$ 1/2 changes in firms' input sales from endogenous network changes

Introduction

This paper

1. Document importance of endogenous networks in firm size heterogeneity
   - Indian firm network micro-data → choice of suppliers & intensity of use explain 80%
2. Develop tractable firm network formation model where firm heterogeneity ⇄ trade
   - rationalizes firm-to-firm network data and accommodates gravity relationships
3. Propose scalable framework for estimation + counterfactual analysis
   - maximum likelihood estimation + no simulation for counterfactuals
4. Evaluate impact of reducing inter-state border frictions by 10%
   - sizable district-level welfare gains [1%,8%]
   - $>$ 1/2 changes in firms' input sales from endogenous network changes

Introduction
This paper

1. Document importance of endogenous networks in firm size heterogeneity
   - Indian firm network micro-data $\rightarrow$ choice of suppliers & intensity of use explain 80%
2. Develop tractable firm network formation model where firm heterogeneity $\rightleftarrows$ trade
   - rationalizes firm-to-firm network data and accommodates gravity relationships
3. Propose scalable framework for estimation + counterfactual analysis
   - maximum likelihood estimation + no simulation for counterfactuals
4. Evaluate impact of reducing inter-state border frictions by 10%
   - sizable district-level welfare gains [1%,8%]
   - $>$ 1/2 changes in firms' input sales from endogenous network changes

Related Literature

This paper: Firm-to-Firm Trade in Endogenous Production Networks

Theory		Discrete Choice	Love of Variety (+ Fixed Costs)
Trade	Aggregate	Dornbusch, Fischer & Samuelson (1977) Eaton & Kortum (2002)	Krugman (1980)
	Firm-Level	Bernard, Eaton, Jensen & Kortum (2003)	Melitz (2003), Chaney (2008) Eaton, Kortum & Kramarz (2011)
	Firm-to-Firm	Eaton, Kortum & Kramarz (2016) **This paper**: firm-to-firm predictions	Huneeus (2019) Tintelnot, Kikkawa, Mogstad & Dhyne (2019)
Macro	Endogenous Production Networks	Oberfield (2018) Boehm & Oberfield (2020) Acemoglu & Azar (2020)	Lim (2017) Taschereau-Dumouchel (2017) Bernard et al. (2020)
Estimation & Counterfactuals		Eaton, Kortum & Sotelo (2013) Dingel & Tintelnot (2020) **This paper**: maximum likelihood Menzel (2015)	simulation-based

Notation

- **Locations** indexed by $o, d \in \mathcal{J} \equiv \{1, \cdots, J\}$
  [o for *origin*, d for *destination*]
- **Firms** indexed by $s, b \in \mathcal{M} \equiv \{1, \cdots, M\}$
  [s for *seller*, b for *buyer*]

Data

Sources

- Universe of intra-state firm-to-firm transactions
  [assembled from commercial tax authorities in 5 Indian states]
  - 141 districts:
    Gujarat (25), Maharashtra (35), Tamil Nadu (32), Odisha (30) and West Bengal (19)
  - 5 years: FY 2011-12 to 2015-16
  - 2.6 million firms and 103 million firm-to-firm connections
- Universe of inter-state firm-to-firm transactions
  [from Ministry of Finance, Govt. of India]
  - 5 years: FY 2011-12 to 2015-16

Data

Sources

- Universe of intra-state firm-to-firm transactions
  [assembled from commercial tax authorities in 5 Indian states]
  - 141 districts:
    Gujarat (25), Maharashtra (35), Tamil Nadu (32), Odisha (30) and West Bengal (19)
  - 5 years: FY 2011-12 to 2015-16
  - 2.6 million firms and 103 million firm-to-firm connections

- Universe of inter-state firm-to-firm transactions
  [from Ministry of Finance, Govt. of India]
  - 5 years: FY 2011-12 to 2015-16

- **Data** [value of goods sold by s to b]

$$\text{sales}_{od}(s,b)$$

- **Cost Shares** [b's intensity of use of s]

$$\pi_{od}(s,b) = \frac{\text{sales}_{od}(s,b)}{\text{input costs}_d(b)}$$

$$\text{input costs}_d(b) = \sum_s \text{sales}_{od}(s,b)$$

- **Intensity of Use**

$$\text{intensity of use}_o(s) = \sum_b \pi_{od}(s,b)$$

Empirical Regularities

Margins of Firms' Sales

$$
\begin{aligned}
\text{input sales}_o(s) &= N_o(s) && \text{[\# Customers]} \\
&\times \frac{\sum_b \pi_{od}(s,b)}{N_o(s)} && \text{[Intensity per Customer]} \\
&\times \frac{\sum_b \pi_{od}(s,b) \times \text{input costs}_d(b)}{\sum_b \pi_{od}(s,b)} && \text{[Average Customer Size]}
\end{aligned}
$$

- Larger Indian firms (higher input sales)
  - tend to have more customers [35%]
  - tend to be used more intensively by customers [46%]
  - tend to have larger customers [19%]

Empirical Regularities

Upstream & Downstream Margins of Firms' Sales

$$\overbrace{\text{\#Customers} \times \text{Intensity per Customer}}^{\text{upstream margin}\approx 81\%} \times \underbrace{\text{Average Customer Size}}_{\text{downstream margin}\approx 19\%}$$

- **Downstream Margin** $\Longrightarrow$ role of exogenous network linkages
  - choice of quantity to sell $\equiv$ downstream decision
  - downstream decision affects upstream firms $\rightarrow$ demand shocks propagate upstream
- **Upstream Margin** [Intensity of Use] $\Longrightarrow$ role of endogenous network formation
  - choice of suppliers and intensity of use $\equiv$ upstream decision
  - upstream decision affects downstream firms $\rightarrow$ cost savings propagate downstream

Empirical Regularities

Upstream & Downstream Margins of Firms' Sales

$$\overbrace{\#\text{Customers} \times \text{Intensity per Customer}}^{\text{upstream margin}\approx 81\%} \times \underbrace{\text{Average Customer Size}}_{\text{downstream margin}\approx 19\%}$$

- **Downstream Margin** $\implies$ role of exogenous network linkages
  - choice of quantity to sell $\equiv$ downstream decision
  - downstream decision affects upstream firms $\rightarrow$ demand shocks propagate upstream
- **Upstream Margin** [Intensity of Use] $\implies$ role of endogenous network formation
  - choice of suppliers and intensity of use $\equiv$ upstream decision
  - upstream decision affects downstream firms $\rightarrow$ cost savings propagate downstream

Model

Overview

- Develop GE model of network formation between spatially distant firms
  - firms have multiple input requirements
  - randomly encounter potential input suppliers
  - select most cost-effective supplier for each requirement
- Low production cost firms end up larger because
  - find more customers
  - used more intensively by their customers
  - customers use cheaper inputs intensively $\rightarrow$ lower production costs
  - lower production costs $\rightarrow$ customers become larger themselves

Model

Overview

- Develop GE model of network formation between spatially distant firms
  - firms have multiple input requirements
  - randomly encounter potential input suppliers
  - select most cost-effective supplier for each requirement
- Low production cost firms end up larger because
  - find more customers
  - used more intensively by their customers
  - customers use cheaper inputs intensively $\rightarrow$ lower production costs
  - lower production costs $\rightarrow$ customers become larger themselves

- **Production Function**

$$y_d(b) = z_d(b) \underbrace{\left(\frac{l_d(b)}{1-\alpha_d}\right)^{1-\alpha_d}}_{\text{labor}} \underbrace{\left(\frac{\prod_{k=1}^{K_d(b)} \overbrace{m_d(b,k)^{1/K_d(b)}}^{\text{symmetric}}}{\alpha_d}\right)^{\alpha_d}}_{\text{materials}}$$

$$m_d(b,k) = \underbrace{\sum_{s \in \mathcal{S}_d(b)} m_{od}(s,b,k)}_{\text{substitutes}}$$

  - α_d, materials share at d
  - $K_d(b)$, # tasks of b
  - $\mathcal{S}_d(b)$, set of potential suppliers for b

Model

Technology

- **Marginal Cost**

$$\overbrace{c_d(b)}^{\text{buyer MC}} = \frac{w_d^{1-\alpha_d}}{z_d(b)} \times \prod_{k=1}^{K_d(b)} \left(\underbrace{p_d(b,k)}_{\text{effective price of task } k \text{ for } b} \right)^{\frac{\overbrace{\alpha_d}^{\text{cost share of task } k}}{K_d(b)}}$$

- **Effective Price**

$$p_d(b,k) = \min_{s \in \mathcal{S}_d(b)} \left\{ \frac{\overbrace{\bar{m}_{od}(s,b,k)}^{\text{markup}} \overbrace{\tau_{od}}^{\text{trade cost}}}{\underbrace{a_{od}(s,b,k)}_{\text{match productivity}}} \times \overbrace{c_o(s)}^{\text{seller MC}} \right\}$$

Model

Functional Form Assumptions

$$\mathbb{P}\left(s \text{ meets } b\right) = \frac{\lambda}{M} \qquad \text{Bernoulli Encounters}$$

$$\mathbb{P}\left(a_{od}(s,b,k) \leq a\right) = \left(1 - \left({}^{a}/_{a_0}\right)^{-\zeta}\right) \mathbf{1}\left\{a > a_0\right\} \qquad \text{Pareto Match Productivities}$$

$$\bar{m}_{od}(s,b,k) \sim \text{Limit Pricing} \qquad \text{Bertrand Competition}$$

$$\mathbb{P}\left(z_d(b) \leq z\right) = \exp\left(-T_d z^{-\theta}\right) \mathbf{1}\left\{z > 0\right\} \quad \theta > \zeta \qquad \text{Fréchet Productivities}$$

- **Recursive Problem**

$$\underbrace{c_d(b)}_{\text{value function}} = \frac{w_d^{1-\alpha_d}}{z_d(b)} \times \prod_{k=1}^{K_d(b)} \min_{s \in \mathcal{S}_d(b)} \left\{ \frac{\bar{m}_{od}(s,b,k)\tau_{od}}{a_{od}(s,b,k)} \times \underbrace{c_o(s)}_{\text{upstream value function}} \right\}^{\frac{\overbrace{\alpha_d}^{\text{discount factor}}}{K_d(b)}}$$

- **Estimands** [exogenous: $\tau_{od}|$ endogenous: $c_d(b)$]
  - very high-dimensional → full solution methods infeasible
  - interdependence in link formation → simulation burdensome

[Rust (1987), Anderson & van Wincoop (2003), Antràs & de Gortari (2020)]

- **Conditional Choice Probabilities**
  [conditional on $c_o(s)$, probability that s gets chosen for any task of any firm at d]

$$\pi^0_{od}(s,-) = \frac{c_o(s)^{-\zeta}\tau_{od}^{-\zeta}}{\sum_{s'\in\mathcal{M}} c_{o'}(s')^{-\zeta}\tau_{o'd}^{-\zeta}}$$

- CCPs which depend on endogenous state $\mapsto$ sample analogs
  [Hotz & Miller (1993)→ Menzel (2015)]

- **Conditional Choice Probabilities**
  [conditional on $c_o(s)$, probability that s gets chosen for any task of any firm at d]

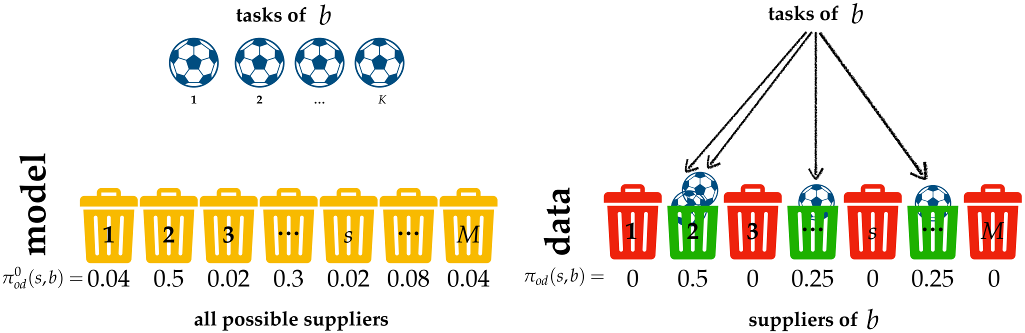
$$\pi^0_{od}(s, -) = \frac{c_o(s)^{-\zeta} \tau_{od}^{-\zeta}}{\sum_{s' \in \mathcal{M}} c_{o'}(s')^{-\zeta} \tau_{o'd}^{-\zeta}}$$

- CCPs which depend on endogenous state $\mapsto$ sample analogs
  [Hotz & Miller (1993)→ Menzel (2015)]

Taking Model to Data

Conditional Choice Probabilities → Balls-and-Bins Model

symmetric + Cobb-Douglas tasks $\implies$ task proportions $=$ cost shares



discrete # tasks $\implies \overbrace{\text{success probabilities [CCPs]}}^{\text{model}} = \mathbb{E}\left[\text{task proportions}\right] = \mathbb{E}\left[\overbrace{\text{cost shares}}^{\text{data}}\right]$

Estimation

Balls-and-Bins Model → Multinomial Logit

- **Estimation Equation**

$$\mathbb{E}\left[\pi_{od}(s,b)\right] = \pi^0_{od}(s,b)$$
$$= \frac{c_o(s)^{-\zeta}\tau_{od}^{-\zeta}}{\sum_{s'\in\mathcal{M}} c_{o'}(s')^{-\zeta}\tau_{o'd}^{-\zeta}}$$

- **Estimands**
  - marginal costs $c_o(s)^{-\zeta} \equiv$ firm fixed effects
  - trade frictions $\tau_{od}^{-\zeta} \equiv \exp\left(\boldsymbol{X}'_{od}\boldsymbol{\beta}\right)$ $[\boldsymbol{X}_{od} \equiv$ distance, borders etc.$]$

- natural choice since probability of sourcing adds to unity
  [Gourieroux, Monfort & Trognon (1984) → Eaton, Kortum & Sotelo (2013)]

Estimation

Multinomial Logit: Computational Issues

- generalized linear model + millions of fixed effects $\implies$
  - high-dimensional non-linear optimization $\rightarrow$ infeasible by Newton methods
  - incidental parameters bias in $\boldsymbol{\beta}$
- not a problem!
  - multinomial likelihood score equations coincide with Poisson likelihood [Baker (1994) $\rightarrow$ Taddy (2015)]
  - Poisson likelihood automatically satisfies adding up constraints [Fally (2015)]
  - Poisson likelihood $\implies$ no bias + fixed effects in closed-form [Hausman, Hall & Griliches (1984)]

Estimation

Multinomial Logit: Computational Issues

- generalized linear model + millions of fixed effects $\implies$
  - high-dimensional non-linear optimization $\rightarrow$ infeasible by Newton methods
  - incidental parameters bias in $\boldsymbol{\beta}$
- not a problem!
  - multinomial likelihood score equations coincide with Poisson likelihood [Baker (1994) $\rightarrow$ Taddy (2015)]
  - Poisson likelihood automatically satisfies adding up constraints [Fally (2015)]
  - Poisson likelihood $\implies$ no bias + fixed effects in closed-form [Hausman, Hall & Griliches (1984)]

- **Firm Fixed Effects** [low marginal costs $\iff$ high intensity of use]

$$\left(c_o(s)^{-\zeta}\right)^* = \overbrace{\sum_{b \in \mathcal{M}} \pi_{od}\left(s,b\right)}^{\text{intensity of use}}$$

- **Origin-Destination Fixed Effects → Structural Gravity Specification**

$$\left(\frac{\exp\left(\ln\left(c_o^{-\zeta}\right) + \boldsymbol{X}'_{od}\boldsymbol{\beta}\right)}{\sum_{o'} \exp\left(\ln\left(c_{o'}^{-\zeta}\right) + \boldsymbol{X}'_{o'd}\boldsymbol{\beta}\right)} \right)^{*} = \frac{1}{M_d} \sum_{b \in \mathcal{M}_d} \left(\underbrace{\sum_{s \in \mathcal{M}_o} \pi_{od}\left(s, b\right)}_{\text{total cost share of } b \text{ from } o} \right)$$

- **Aggregate Trade Models + Exact Hat Algebra**

$$\text{model degeneracy} \implies \text{model prediction} = \text{observed data}$$

- **Models with Large Networks and Granularity**

$$\text{model non-degeneracy} \implies \text{model prediction(s)} \neq \text{observed data}$$

  - observed data $\rightarrow$ estimated model $\rightarrow \mathbb{E}\left[\text{model predictions} \mid \text{initial state}\right]$
  - counterfactual evaluation:

$$\mathbb{E}\left[\widehat{\text{model predictions}}\right] = \frac{\mathbb{E}\left[\text{model predictions} \mid \text{counterfactual state}\right]}{\mathbb{E}\left[\text{model predictions} \mid \text{initial state}\right]}$$

[Head & Mayer (2019), Dingel & Tintelnot (2020)]

Details

- **Aggregate Trade Models + Exact Hat Algebra**

$$\text{model degeneracy} \implies \text{model prediction} = \text{observed data}$$

- **Models with Large Networks and Granularity**

$$\text{model non-degeneracy} \implies \text{model prediction(s)} \neq \text{observed data}$$

  - observed data $\rightarrow$ estimated model $\rightarrow \mathbb{E}\left[\text{model predictions} \mid \text{initial state}\right]$
  - counterfactual evaluation:

$$\mathbb{E}\left[\widehat{\text{model predictions}}\right] = \frac{\mathbb{E}\left[\text{model predictions} \mid \text{counterfactual state}\right]}{\mathbb{E}\left[\text{model predictions} \mid \text{initial state}\right]}$$

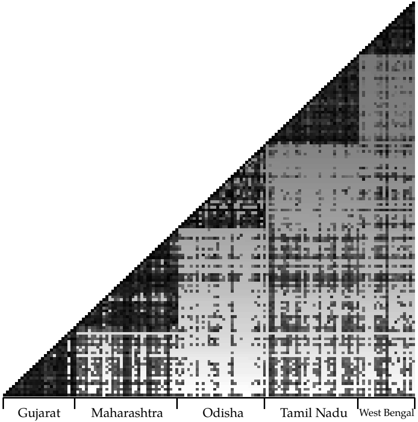
[Head & Mayer (2019), Dingel & Tintelnot (2020)]

Details

Decline in Border Frictions

Counterfactual Experiment

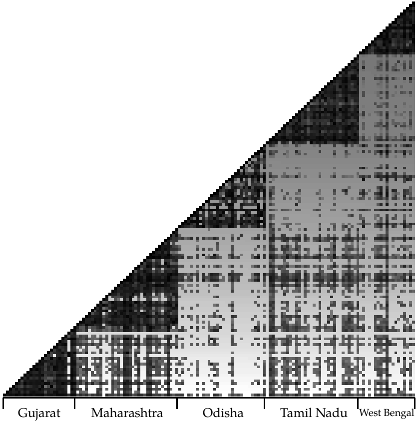
- Trade across state borders subject to frictions
  - significant border effects in gravity regressions
  - sales taxes, border inspections, logistical delays etc.
  - 141×141 symmetric matrix of inter-district Head-Ries indices, $\sqrt{\frac{\textbf{sales}_{od}\textbf{sales}_{do}}{\textbf{sales}_{oo}\textbf{sales}_{dd}}} \Longrightarrow$
- 10% decline in trade costs between inter-state district pairs



Decline in Border Frictions

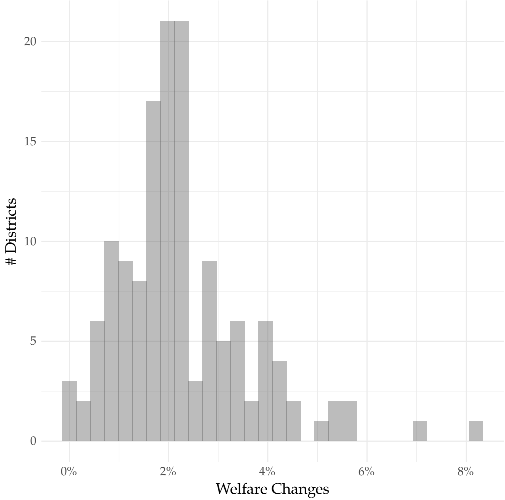
Counterfactual Experiment

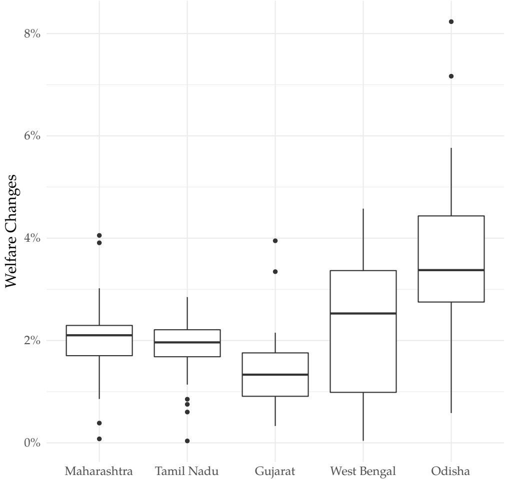
- Trade across state borders subject to frictions
  - significant border effects in gravity regressions
  - sales taxes, border inspections, logistical delays etc.
  - 141×141 symmetric matrix of inter-district Head-Ries indices, $\sqrt{\frac{\textbf{sales}_{od}\textbf{sales}_{do}}{\textbf{sales}_{oo}\textbf{sales}_{dd}}} \implies$
- 10% decline in trade costs between inter-state district pairs



Decline in Border Frictions

Macro Outcomes: Aggregate Welfare Changes





Decline in Border Frictions

Micro Outcomes: Changes in Margins of Firms' Sales, Shapley Decomposition

State	Maharashtra (1)	Tamil Nadu (2)	Gujarat (3)	West Bengal (4)	Odisha (5)	All (6)
Δ% upstream margin	40.76%	40.81%	36.49%	39.44%	38.06%	55.69%
Δ% downstream margin	29.37%	34.14%	45.74%	31.44%	43.02%	33.45%
second order term	29.86%	25.04%	17.76%	29.14%	18.91%	10.85%

$$\frac{\Delta\text{Sales}}{\text{Sales}} \approx \overbrace{\frac{\Delta\text{Intensity of Use}}{\text{Intensity of Use}}}^{\text{upstream margin}} + \underbrace{\frac{\Delta\text{Average Customer Size}}{\text{Average Customer Size}}}_{\text{downstream margin}}$$

$$+ \frac{\Delta\text{Intensity of Use}}{\text{Intensity of Use}} \times \frac{\Delta\text{Average Customer Size}}{\text{Average Customer Size}}$$

Conclusion

- Documented importance of endogenous networks towards firm heterogeneity
- Developed tractable model of endogenous spatial production networks
- Proposed scalable framework for structural estimation + counterfactual analysis
- Reducing border frictions
  - improves welfare across Indian districts in the range [1%,8%]
  - $> 1/2$ firm-level changes from endogenous network changes
- Extensions:
  Supply Chain Dynamics, Search Frictions, Innovation Spillovers, Factor Market Frictions, Industry Dynamics

Conclusion

- Documented importance of endogenous networks towards firm heterogeneity
- Developed tractable model of endogenous spatial production networks
- Proposed scalable framework for structural estimation + counterfactual analysis
- Reducing border frictions
  - improves welfare across Indian districts in the range [1%,8%]
  - $>$ 1/2 firm-level changes from endogenous network changes
- Extensions:
  Supply Chain Dynamics, Search Frictions, Innovation Spillovers, Factor Market Frictions, Industry Dynamics

Conclusion

- Documented importance of endogenous networks towards firm heterogeneity
- Developed tractable model of endogenous spatial production networks
- Proposed scalable framework for structural estimation + counterfactual analysis
- Reducing border frictions
  - improves welfare across Indian districts in the range [1%,8%]
  - $>$ 1/2 firm-level changes from endogenous network changes
- Extensions:
  Supply Chain Dynamics, Search Frictions, Innovation Spillovers, Factor Market Frictions, Industry Dynamics

Conclusion

- Documented importance of endogenous networks towards firm heterogeneity
- Developed tractable model of endogenous spatial production networks
- Proposed scalable framework for structural estimation + counterfactual analysis
- Reducing border frictions
  - improves welfare across Indian districts in the range [1%,8%]
  - $>$ 1/2 firm-level changes from endogenous network changes
- Extensions:
  Supply Chain Dynamics, Search Frictions, Innovation Spillovers, Factor Market Frictions, Industry Dynamics

Conclusion

- Documented importance of endogenous networks towards firm heterogeneity
- Developed tractable model of endogenous spatial production networks
- Proposed scalable framework for structural estimation + counterfactual analysis
- Reducing border frictions
  - improves welfare across Indian districts in the range $[1\%, 8\%]$
  - $> 1/2$ firm-level changes from endogenous network changes
- Extensions:
  Supply Chain Dynamics, Search Frictions, Innovation Spillovers, Factor Market Frictions, Industry Dynamics