Logistics Networks and E-Commerce Trade: Theory and Evidence from China

Chunmian Ge[1] **Hanwei Huang**[2] **Kalina Manova**[3] **Daniel Yi Xu**[4]

Abstract

This paper studies how transportation firms configure their logistics networks across space to serve e-commerce trade, and how the network structure and imperfect competition among transportation firms shape delivery pricing, trade costs, and consumer welfare. Using novel data from *Cainiao*, one of China's largest logistics informatics platforms , we document four stylized facts: (i) most routes are served by only a few firms; (ii) routes with higher trade potential attract more firms; (iii) delivery prices decline with the number of competing firms; and (iv) higher-productivity firms enter more routes, but their marginal routes exhibit lower average popularity. We develop a tractable spatial model that embeds an endogenous logistics network within an otherwise standard Krugman (1980) framework. The transportation sector features heterogeneous firms engaging in differentiated Bertrand competition and scale economies in network operation, generating interdependence across route-level entry decisions. The model delivers a pecking order in transportation firm entry and predicts pro-competitive effects of entry on markups and trade costs, as well as network spillovers to neighboring routes. We estimate the model using a combination of IV-GMM and simulated method of moments and find that it fits observed entry patterns and trade flows reasonably well. Counterfactual analysis shows that a reduction in transportation entry costs disproportionately benefits remote provinces, thereby reducing regional inequality in welfare gains from e-commerce trade.

[1] South China University of Technology.
[2] City University of Hong Kong, CEP, CIFER.
[3] UCL, CEPR, CEP.
[4] Duke University, NBER, CEPR, ABFER.

Logistics Networks and E-Commerce Trade: Theory and Evidence from China

Chunmian Ge[a] Hanwei Huang[b] Kalina Manova[c] Daniel Yi Xu[d]

[a]South China University of Technology

[b]City University of Hong Kong, CEP, CIFER

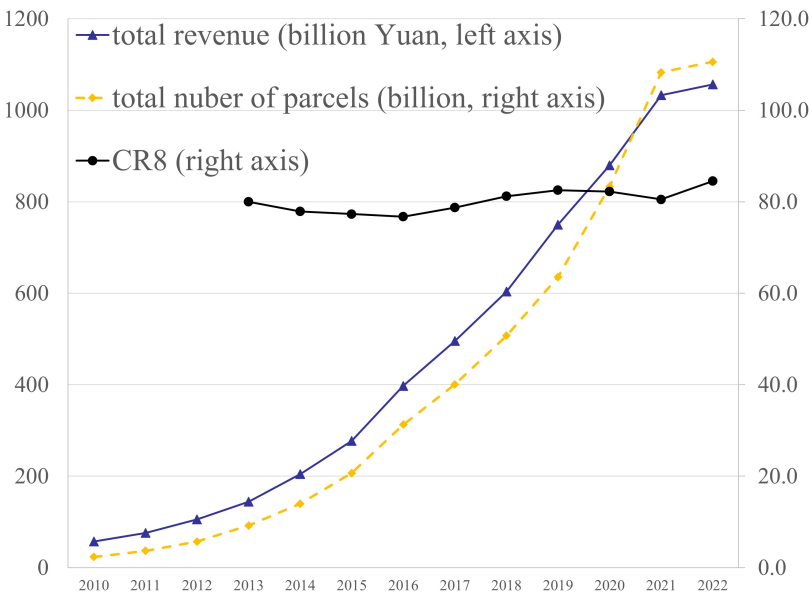
[c]UCL, CEPR, CEP

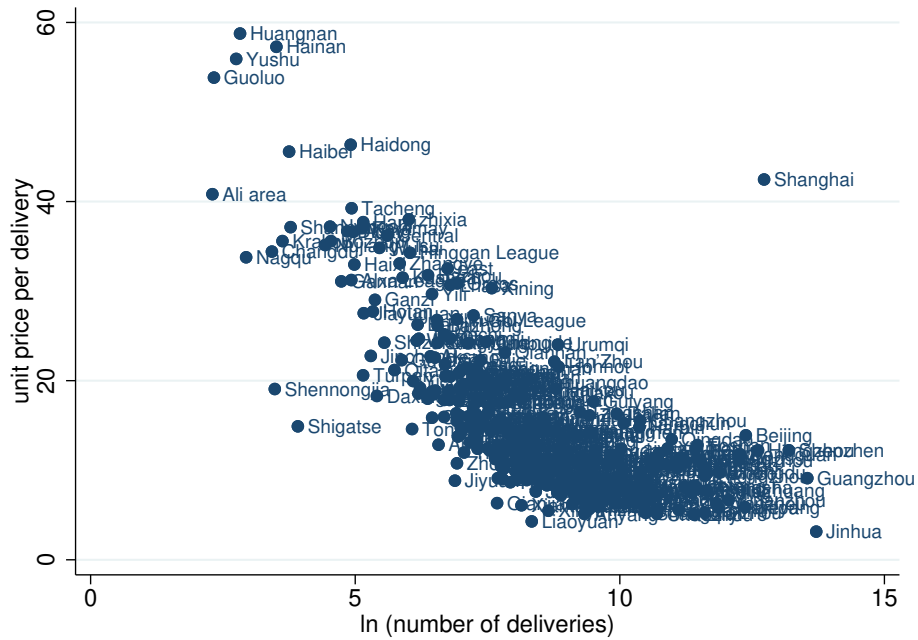
[d]Duke University, NBER, CEPR, ABFER

Preliminary

Background

- China has the world's largest e-commerce and e-commerce delivery markets
  - Share in global online retail: 45.8% in 2017
  - Online retail sales made up 26.2% of total retail sales of consumer goods in 2025
  - Share in global parcel volume: 68.5% in 2022
- E-commerce delivery is dominated by large firms: CR8 $\approx$ 80%
- The cost of parcel delivery exhibit significant spatial variations





Research Question

- How do transportation firms configure their spatial network to serve cross-region e-commerce trade?
- How do transportation firm networks and competition shape e-commerce delivery pricing, trade costs, and consumer welfare?

This paper

- Using data from a leading Chinese logistics platform, we uncovered new facts on e-commerce transportation and trade costs
- We developed a tractable model with endogenous logistics network and endogenous trade costs
- Quantify the impact of network formation costs, imperfect competition, and deregulation on trade costs and the welfare gains from e-commerce trade

Literature

- Endogenous trade costs
  - Infrastructure Network: Fajgelbaum and Schaal (2020), Allen and Arkolakis (2022), Fan et al. (2023), Kreindler et al. (2023)
  - Imperfect competition: Hummels et al. (2009), Atkin and Donaldson (2016), Asturias (2019), Allen et al (2024a)
  - Backhaul problem and Routing: Behrens and Picard (2011), Ishikawa and Tarui (2017), Wong (2022), Allen et al (2024b)
  - Hybrids: Brancaccio et al (2020), Huang et al (2024)
    - ⋆ This paper: the operation network of transportation firms who command market power
- Market size and market structure
  - Sutton (1991), Bresnahan & Reiss (1991), Berry (1992)
  - This paper: endogenous market size and endogenous market structure
- E-commerce trade
  - Bai et al. (2020), Fan et al (2018), Couture et al. (2021), Dolfen et al. (2023), Houde et al. (2023)
    - ⋆ This paper: e-commerce logistics and trade costs

Roadmap

1. Introduction
2. Data
3. Stylized Facts
4. Model
5. Reduced form evidence
6. Quantification

Data

- We have obtained a month of data from *Cainiao* for 2020, 2022 and 2024 from Cainiao (details come next)
- City-level e-commerce parcels throughput (inflows +outflows) and revenues for intra-city delivery, inter-city delivery (in aggregate)
- Travel time and distance via roads by city-pair using Gaode map API service.
- City-level data such as GDP, population, area etc. from the city year book and government reports

The Cainiao Data

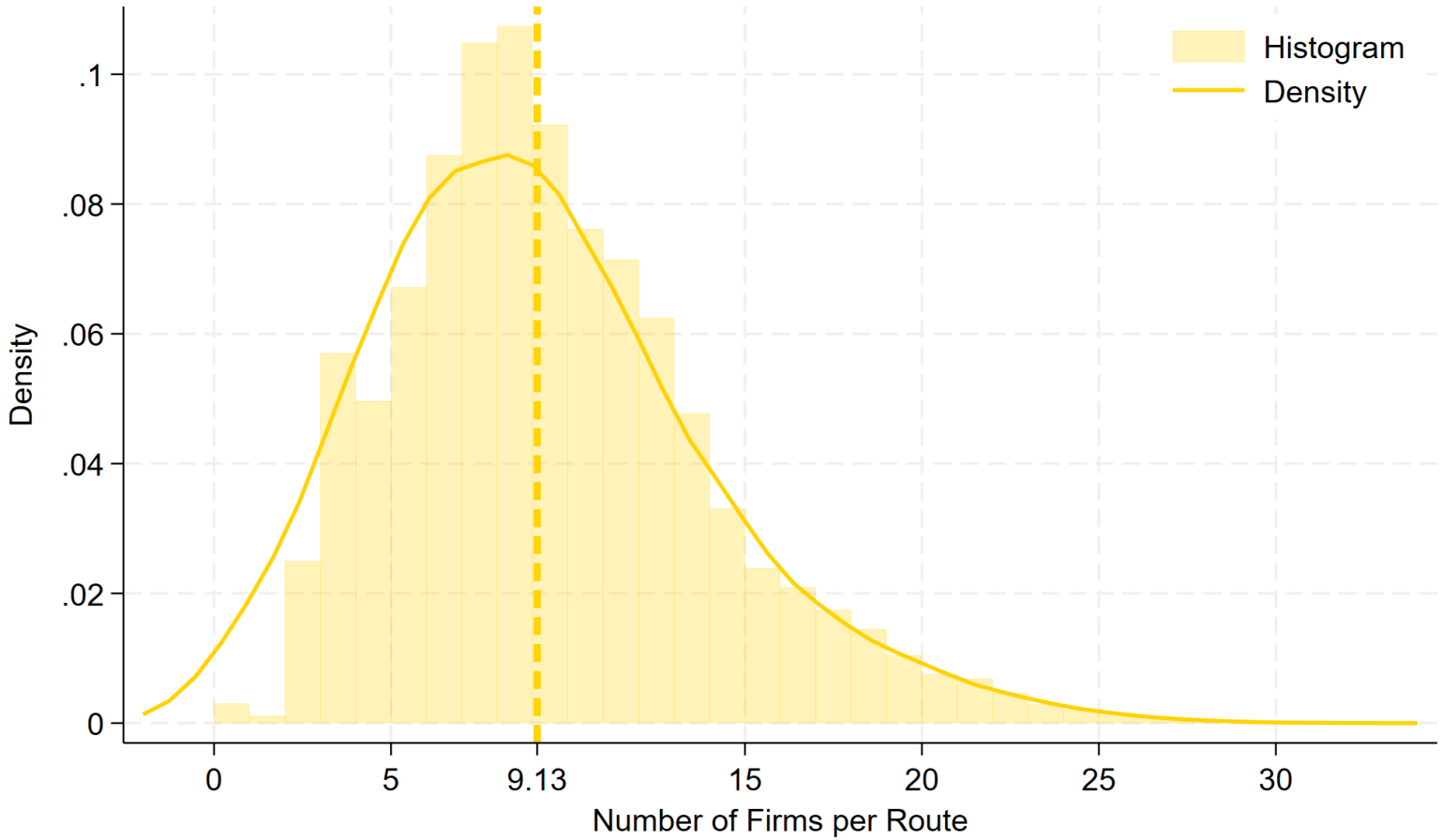
- We ran a web crawler which scraps data from *Cainiao*, the largest platform providing logistics informatics service in China.
- Cainiao was initiated and owned by Alibaba in its effort to build the China Smart Logistic Network.
  - It integrates information flows from logistics firms and allows customers to track deliveries and parcels from multiple companies.
- We crawl two main sets of publicly available information
  - Route level posted price by each company
  - Route level transaction counts

Roadmap

1. Introduction
2. Data
3. Stylized Facts
4. Model
5. Reduced form evidence
6. Quantification

Stylized facts

Fact 1: most of routes are served by only a few firms



Stylized facts

Fact 2: There are more firms serving routes with higher trade potential

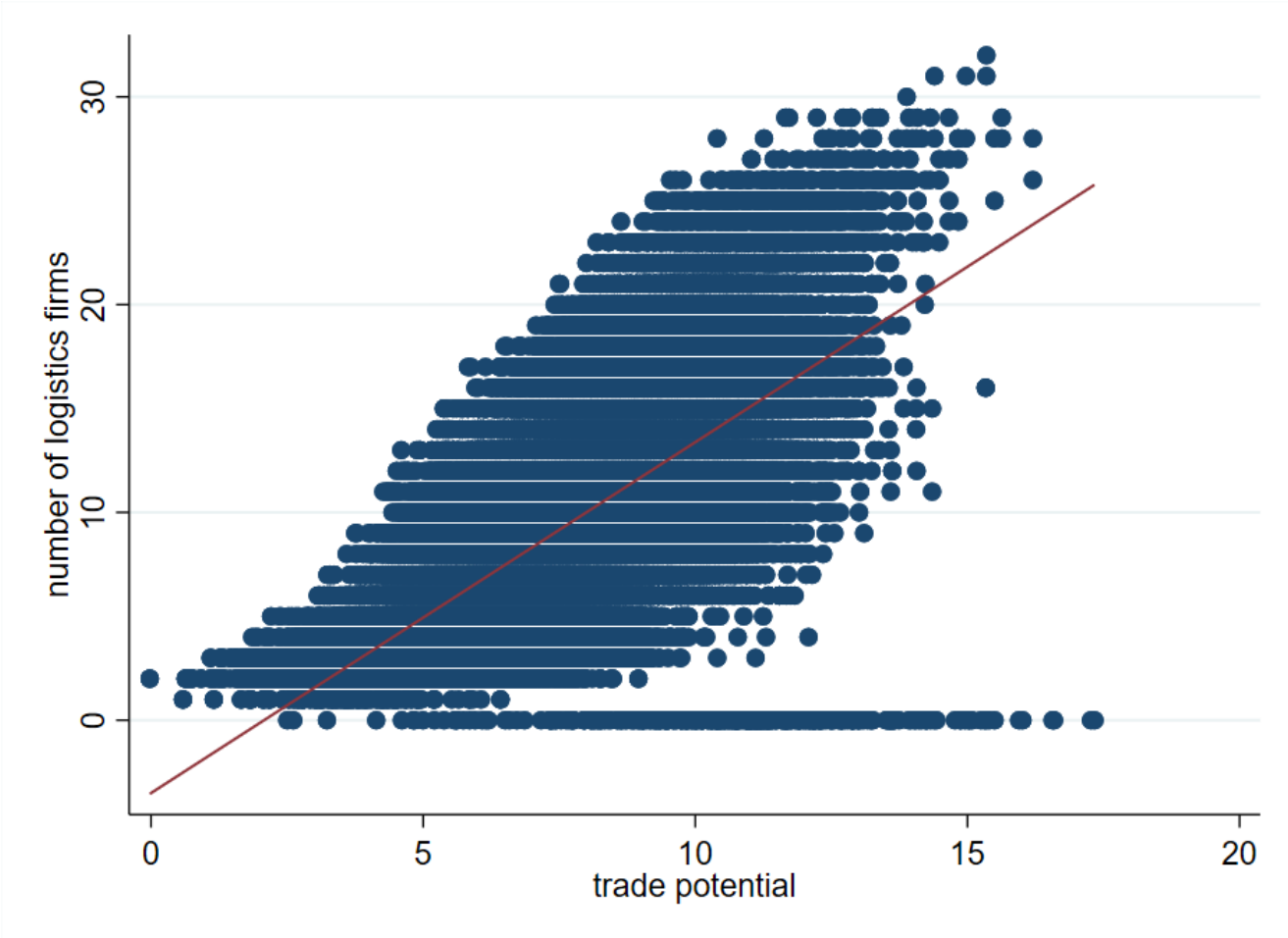
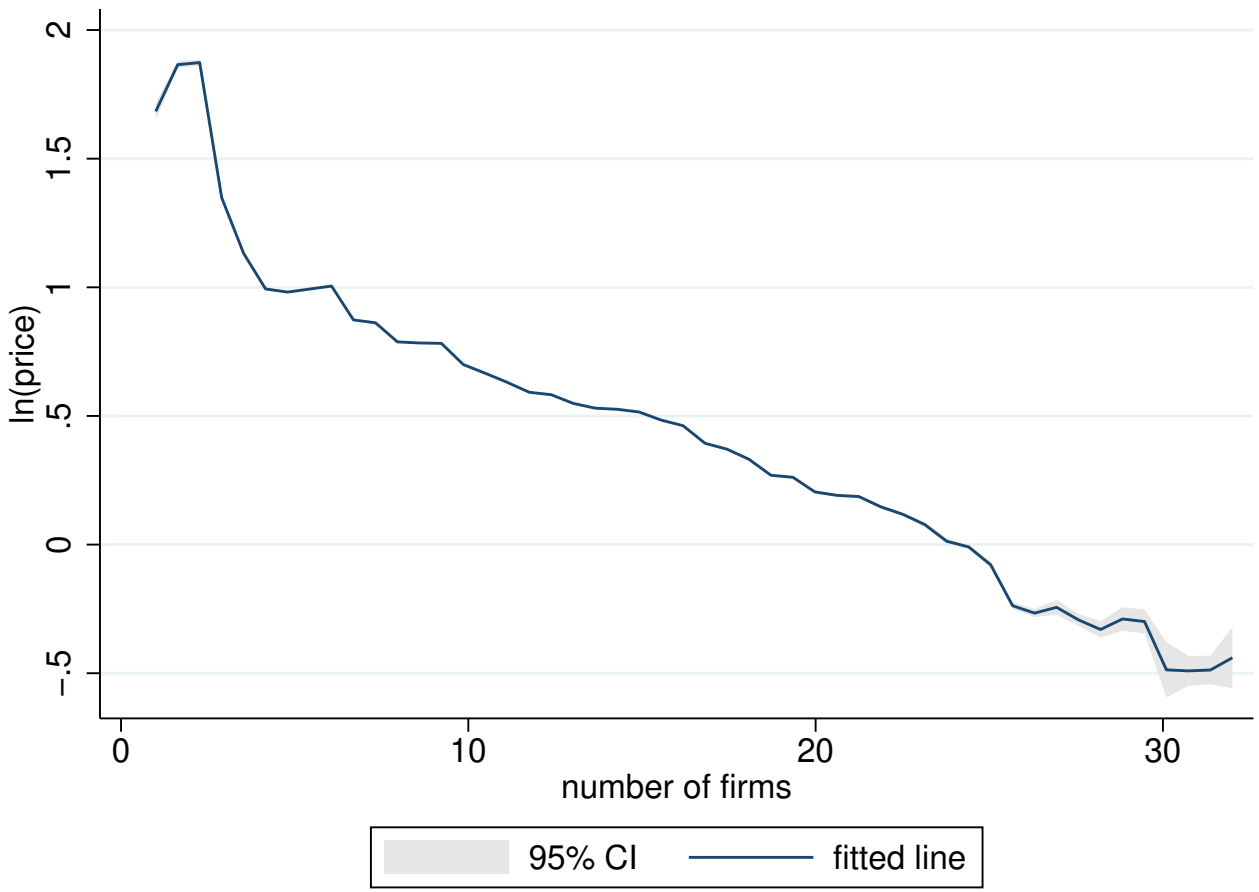


Figure: Predicted Trade and the Number of Logistics Firms

Note: predicted trade from i to j $= \frac{GDP_i GDP_j}{Distance_{ij}}$

Stylized facts

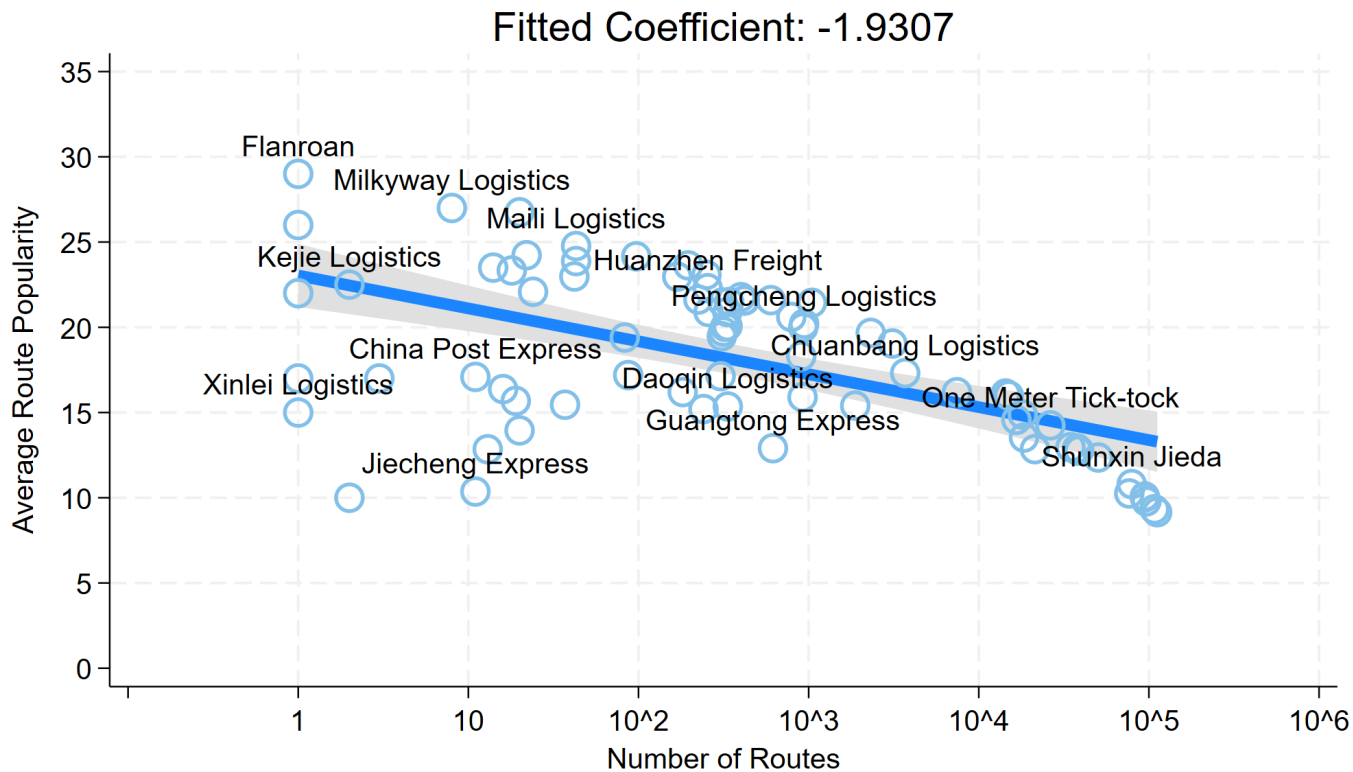
Fact 3: There exists a negative relationship between delivery prices and number of transportation firms



Stylized facts

Fact 4: There is a pecking order in the entry of transportation firms across routes, with better firms entering more routes.

- but their marginal routes are less attractive...



Notes: Average route popularity against total routes for each transportation firms. Route popularity is the number of transportation firms operating this route.

Roadmap

1. Introduction
2. Data
3. Stylized Facts
4. Model
5. Reduced form evidence
6. Quantification

Model: an Overview

- We endogenize trade costs in an otherwise standard Krugman (1980) model
- The transportation sector features differentiated Bertrand competition among transportation firms
- Scale economy in the transportation sector: per-route fixed cost decreases with the number of routes a firm operate
- Transportation firms are heterogeneous and enters sequentially in increasing order of marginal cost
- The model pins down trade flows, trade costs, and the market structure of the transportation sector simultaneously

The Goods Sector

- Consider an economy with I regions, each indexed by i and a measure L_i of homogeneous households
- CES Demand for final goods:

$$q_{ji}(\omega) = \left(\frac{p_{ji}(\omega)}{P_i}\right)^{-\sigma} \frac{w_i}{P_i}, \tag{1}$$

  where w_i and P_i are household income and price index of region i.
- Production: firms with homogeneous productivity φ_i:

$$l_i = f_i + q_i/\varphi_i \tag{2}$$

- Zero profit condition implies that outputs per firm is

$$q_i = \varphi_i f_i(\sigma - 1) \tag{3}$$

- Free entry condition implies that the number active firms in region i is

$$n_i = \frac{L_i}{f_i \sigma} \tag{4}$$

The Transportation Sector

- There are M transportation firms. The marginal cost of a firm m to provide transportation service from region i to j is c_{ijm}
- We indicate whether a firm enters a route or not using

$$I_{ijm} = \begin{cases} 1, \text{ if firm } m \text{ operates on route } i \rightarrow j \text{ (i} \neq \text{j)} \\ 0, \text{ otherwise.} \end{cases}$$

  - Free intra-region trade: $\tau_{ii} = 1$ and $I_{iim} = 0$ for any i and m

- To provide services for route $i \rightarrow j$, a firm m has to pay a fixed cost of:

$$F_m(i,j) = F_{ij0} \cdot e^{-\alpha(N_{im}+N_{jm})}, \tag{5}$$

  - F_{ij0}: initial entry cost without operation in either city i or city j;
  - N_{im} and N_{jm}: number of routes at node j and j respectively:

$$N_{im} = \sum_{k=1}^{I} I_{ikm} + \sum_{k=1}^{I} I_{kim}, \text{ and } N_{jm} = \sum_{k=1}^{I} I_{kjm} + \sum_{k=1}^{I} I_{jkm}. \tag{6}$$

  - $\alpha > 0$: captures the scale economy of firm operation

The Transportation Sector

- Transportation firms engage in Bertrand competition for shipping services, with firm m posting t_{ijm} on route $i \rightarrow j$
- Final-good buyers face Weibull-distributed quality shocks in the transportation service for each variety ω

$$\varepsilon(\omega) \sim Prob(\varepsilon(\omega) \geq t) = e^{-t^{\theta}}. \tag{7}$$

and minimize the cost of buying each variety ω:

$$\tau_{ij}(\omega) = \min_{m=1,\cdots,m_{ij}} \{t_{ijm}\varepsilon(\omega)\}. \tag{8}$$

- With a continuum of varieties, the law of large numbers implies buyers face an expected iceberg trade costs of

$$\tau_{ij} \equiv E[\tau_{ij}(\omega)] = \gamma(\sum_{m=1}^{m_{ij}} t_{ijm}^{-\theta})^{-\frac{1}{\theta}}.$$

The Transportation Sector

- The market share of firm m along route ij

$$\chi_{ijm} = \frac{t_{ijm}^{-\theta}}{\sum_{m=1}^{m_{ij}} t_{ijm}^{-\theta}}. \tag{9}$$

- The problem that a transportation firm m faces is

$$\max_{t_{ijm}} \pi_{ijm} = Q_{ij}\chi_{ijm}(t_{ijm} - c_{ijm}),$$

  - $Q_{ij}\chi_{ijm}$: residual demand for transportation service of firm m
  - Q_{ij} the aggregate trade flow (quantity) from i to j

$$Q_{ij} = \int_{\omega \in \Omega_{ij}} q_{ij}(\omega) L_j d\omega = n_i (p_{ij}/P_j)^{-\sigma} \frac{w_j L_j}{P_j}, \tag{10}$$

Transportation Pricing

Assumption

Transportation firms take consumer price indexes as given during pricing.

- "Large in the small, small in the large" (Neary, 2016)
  - it holds if there are many locations

Proposition 1

There exists a unique Nash Equilibrium in which the transportation firms post the following price

$$t_{ijm} = \frac{\varepsilon_{ijm}}{\varepsilon_{ijm} - 1} c_{ijm},$$

where $\varepsilon_{ijm} = (1 - \chi_{ijm})\theta + \chi_{ijm}\sigma$ is the price elasticity of residual demand of the transportation service.

Transportation Pricing

- The *Pro-competitive effect* a rising number of transportation firms m_{ij} from i to j

Corollary

An increase in m_{ij} reduces market shares, markups, and prices charged by incumbent firms, and the expected trade costs from i to j, τ_{ij}.

Transportation Firms' Network Strategy

- Denote a firm m's **network strategy** $\mathbf{l}_m = \{l_{ijm}\}_{i=1,j=1,i\neq j}^{I,I}$ which is a $0-1$ hollow matrix, and all other firms' network strategies as $\mathbf{l}_{-m}$
- A firm m faces following problem:

$$\max_{\mathbf{l}_m} \pi_m(\mathbf{l}_m, \mathbf{l}_{-m}) = \sum_{i=1}^{I}\sum_{j=1}^{I} l_{ijm}(\pi_{ijm} - F_m(i,j)). \tag{11}$$

- Network strategies features interdependence
  - $F_m(i,j)$ depends on operation on other routes due to scale economy
  - π_{ijm} depends on other firms' pricing due to strategic interaction

Transportation Firms' Network Strategy

Assumption

Suppose transportation firms' marginal cost satisfies $c_{ijm} = c_{ij}/A_m$ and firms enter the routes sequentially in decreasing order of A_m.

- A_m: firm-level productivity; c_{ij}: route-specific marginal cost
- The assumption allows us to get unique equilibrium in firm entry

Transportation Firms' Network Strategy

Assumption

Suppose transportation firms' marginal cost satisfies $c_{ijm} = c_{ij}/A_m$ and firms enter the routes sequentially in decreasing order of A_m.

- A_m: firm-level productivity; c_{ij}: route-specific marginal cost
- The assumption allows us to get unique equilibrium in firm entry

Proposition 2

As long as $1 < \sigma < \theta$, there is a pecking order in transportation firms' network strategy, such that firms with higher productivity entering more routes, that is

$$A_m > A_n \Rightarrow \mathbf{l}_m \geq \mathbf{l}_n. \tag{12}$$

for any two firms m and n.

- High-productivity firms entering more routes than low-productivity firms
  - This is consistent with Stylized Fact 4

Comparative Statics

Proposition 3

As long as $1 < \sigma < \theta$, a reduction in route $i \to j$'s fixed cost F_{ij0} or variable cost c_{ij} tends to

(a) increase the number of firms operating along route $i \to j$

(b) increase the number of firms operating along route connecting to i or j.

- Result (a): increasing differences between I_{ijm} and π_{ijm} and -F_{ijm}
- Result (b): due to scale economy, $F_{ij'm}$ and $F_{i'jm}$ decline as N_{im} and N_{jm} increases for all $m \to$ ***network effect***

Roadmap

1. Introduction
2. Data
3. Stylized Facts
4. Model
5. Reduced form evidence
6. Quantification

Testable Predictions

- The **Pro-competitive effect** of firm entry on pricing

$$\ln(p_{ijnt}) = \alpha \ln(m_{ijt}) + \delta_t + \delta_i + \delta_j + \delta_n + \varepsilon_{ijt} \tag{13}$$

  - $ln(m_{ijt})$: number of firms serving route i to j
  - where δ_t, δ_i, δ_j, and δ_n, are time, origin, destination and firm fixed effects, respectively

- The **Network effect** of firm entry

$$\ln(M_{-ijt}) = \alpha \ln(m_{ijt}) + \delta_t + \delta_{ij} + \varepsilon_{ijt} \tag{14}$$

  - $\ln(M_{-ijt})$: average number of firms (other than i) connecting j that *did not* experience significant change in the driving distance

- IV for $ln(m_{ijt})$: change in average driving distance to j except from i

Pro-competitive Effect

- Does a rise in the number of competing transportation firms lower delivery prices?

	OLS	FE	IV
	(1)	(2)	(3)
$\ln(m)$	-0.584***	-0.123***	-1.261***
	(0.0881)	(0.0266)	(0.174)
Constant	1.928***		
	(0.149)		
Origin FE	N	Y	Y
Destination FE	N	Y	Y
Company FE	N	Y	Y
Year FE	N	Y	Y
R^2	0.131	0.778	-0.388
KP F-stat			78.31
No. of Observations	2411103	2411103	2411103

Notes: 3-year samples with observed prices are considered. STATA command used for FE is reghdfe; for IV is ivreghdfe. Standard errors are clustered at the origin-destination-firm level. The IV is $\ln(ddist_{-ijt})$, the average log(driving distance) from all other origins to the same destination j. The self driving distance IV generates an overly large estimate which is not a good IV in this case. * 0.1 ** 0.05 *** 0.01

Network Effect

- Does a rise in the number of firms raise the number of firms in neighboring routes?

	OLS	FE		IV
	(1)	(2)	(3)	(4)
$\ln(m)$	0.213***	0.153***	0.151***	0.498***
	(0.00195)	(0.00136)	(0.00132)	(0.0184)
Constant	1.650***			
	(0.00454)			
Route FE	N	Y	Y	Y
Year FE	N	N	Y	Y
KP F-stat				872.9
No. of Observations	223778	223778	223778	223778

Notes: Intra-city routes are included. Only routes from 2020 and 2024 with a positive number of transportation firms are considered. The dependent variable is $\ln(M_{jt}) = \frac{1}{|\mathcal{A}_j|}\sum_{i' \in \mathcal{A}_j} \ln(m_{i'jt})$, where $\mathcal{A}_j = \{i' \in \mathcal{I} : |\Delta \ln(ddist_{i'j})| \leq \infty\}$. The specifications use full samples. The IV is $\ln(ddist_{ijt})$. STATA command used for FE is reghdfe; for IV is ivreghdfe. Standard errors are clustered at the route level. * 0.1 ** 0.05 *** 0.01

Roadmap

1. Introduction
2. Data
3. Stylized Facts
4. Model
5. Reduced form evidence
6. Quantification

Identification: Production and Demand

We estimate/calibrate the model to the 2020 Chinese economy. At the provincial level, we measure:

1. Population L_i: the year-end population
2. Production firms number n_i: the number of legal entities
3. Output per firm q_i: the real regional GDP divided by the number of legal entities
4. Wage w_i: the nominal average wage

Given the demand elasticity σ, we use the following Krugman model results to calibrate production entry cost f_i and productivity φ_i:

$$\widehat{f}_i = \frac{L_i}{n_i \sigma} \tag{15}$$

$$\widehat{\varphi}_i = \frac{q_i}{\widehat{f}_i(\sigma - 1)} \tag{16}$$

Identification: Transportation Sector

- We parametrize the route-firm-level marginal cost as:

$$\ln c_{ijm} = \underbrace{\beta_c \ln dist_{ij} + \eta_c \ln hour_{ij}}_{=\ln c_{ij}} + \underbrace{a_m}_{=\ln(\frac{1}{A_m})} + \xi_{ijm}$$

  where $dist_{ij}$ is driving distance, $hour_{ij}$ the reference delivery time, a_m the firm cost shifter, ξ_{ijm} the error term.

- Plugging it into the pricing equation $t_{ijm} = \frac{\varepsilon_{ijm}}{\varepsilon_{ijm}-1} c_{ijm}$, we have

$$\ln t_{ijm} = \ln \frac{\chi_{ijm}\sigma + (1-\chi_{ijm})\theta}{\chi_{ijm}\sigma + (1-\chi_{ijm})\theta - 1} + \beta_c \ln dist_{ij} + \eta_c \ln hour_{ij} + a_m + \xi_{ijm}$$

  allowing for estimation of θ, β_c, η_c and a_m for observed χ and t

- χ is endogenous, instrumented by neighboring routes competitors

SMM Estimation of Entry Costs

- Suppose fixed entry costs structure follows

$$\ln F_{ijm} = F + \beta_F \ln dist_{ij} - \alpha(N_{im} + N_{jm})$$

- Let $\Phi = \{F, \beta_F, \alpha\}$ and $\tilde{m}(\Phi) = [m_1(\Phi), ..., m_S(\Phi)]$ be the model moments. We estimate Φ by minimizing the distance between the model and data moments:

$$y_k = (\widetilde{\mathrm{m}}(\Phi_k) - m)\mathrm{W}(\widetilde{\mathrm{m}}(\Phi_k) - m)'$$

  for a given weighting matrix W.

Estimated Parameters

Table: Summary of Parameter Estimates

σ	θ	β_c	η_c	α	F	β_F
2.000	2.856^{***}	0.318^{***}	-0.084^{***}	0.017	0.0003	0.305

Note: θ, β_c, η_c are the IV-GMM estimates. α, F, β_F are point estimates from the single-sample SMM. For SA parameters, $T_0 = 1000$, $c = 0.995$, $K = 2000$, and we use 10 parallel multi-starts. Initial guesses are $\theta_0 = 1$, $\beta_{c0} = 0.345$, $\eta_{c0} = -0.1$, $\alpha_0 = 0.015$, $F = 0.0003$, $\beta_{F0} = 0.360$. Robust standard error is used. * 0.1 ** 0.05 *** 0.01.

Targeted Moments

- The model fits the entry patterns of firms across the top, middle and bottom tiers of productivity relatively well.

Table: Targeted Moments: Data vs Model

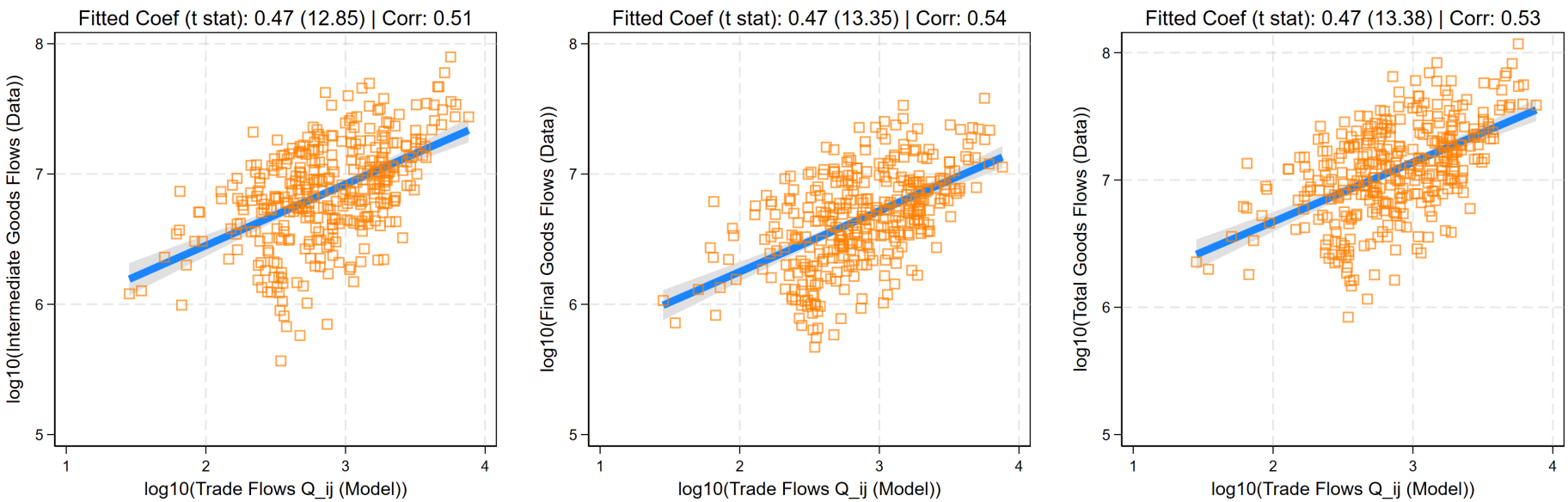
	Data	Model
Top 1/3 Tier	0.480	0.373
Middle 1/3 Tier	0.296	0.319
Bottom 1/3 Tier	0.030	0.041
Middle 3/5 Tier	0.316	0.269
Bottom 1/5 Tier	0.029	0.024

Note: moments are average route shares (out of total $31 * 30 = 930$ provincial routes) among firms in different tiers. Tiers are ranked by descending order of estimated productivity A_m.

Model Performance

- The model-predicted trade flows are positively correlated with their data counterparts (aggregated from the 2020 Input-Output table).

Figure: Trade Flows: Model vs Data



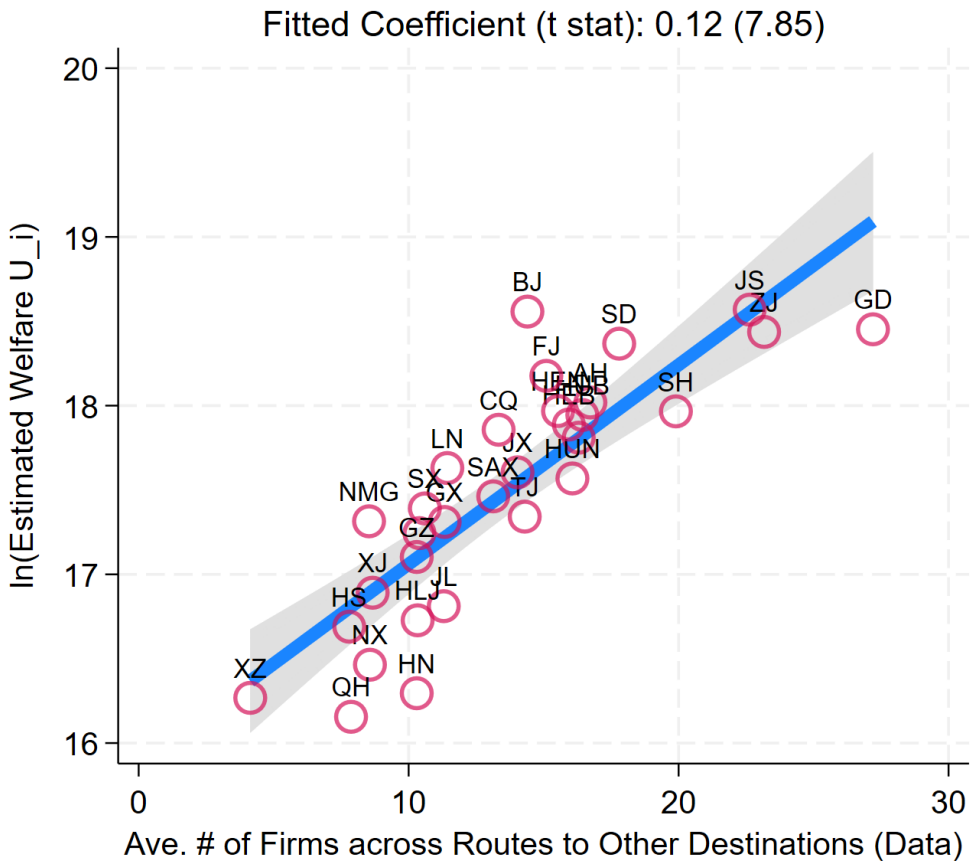
(a) Intermediate Goods (b) Final Goods (c) Total Goods

Note: the plot excludes intra-city routes and inter-city routes with zero logistics firms/trade flows predicted.

Welfare

- Welfare of more central provinces tends to be higher.

Figure: Welfare and Network Centrality

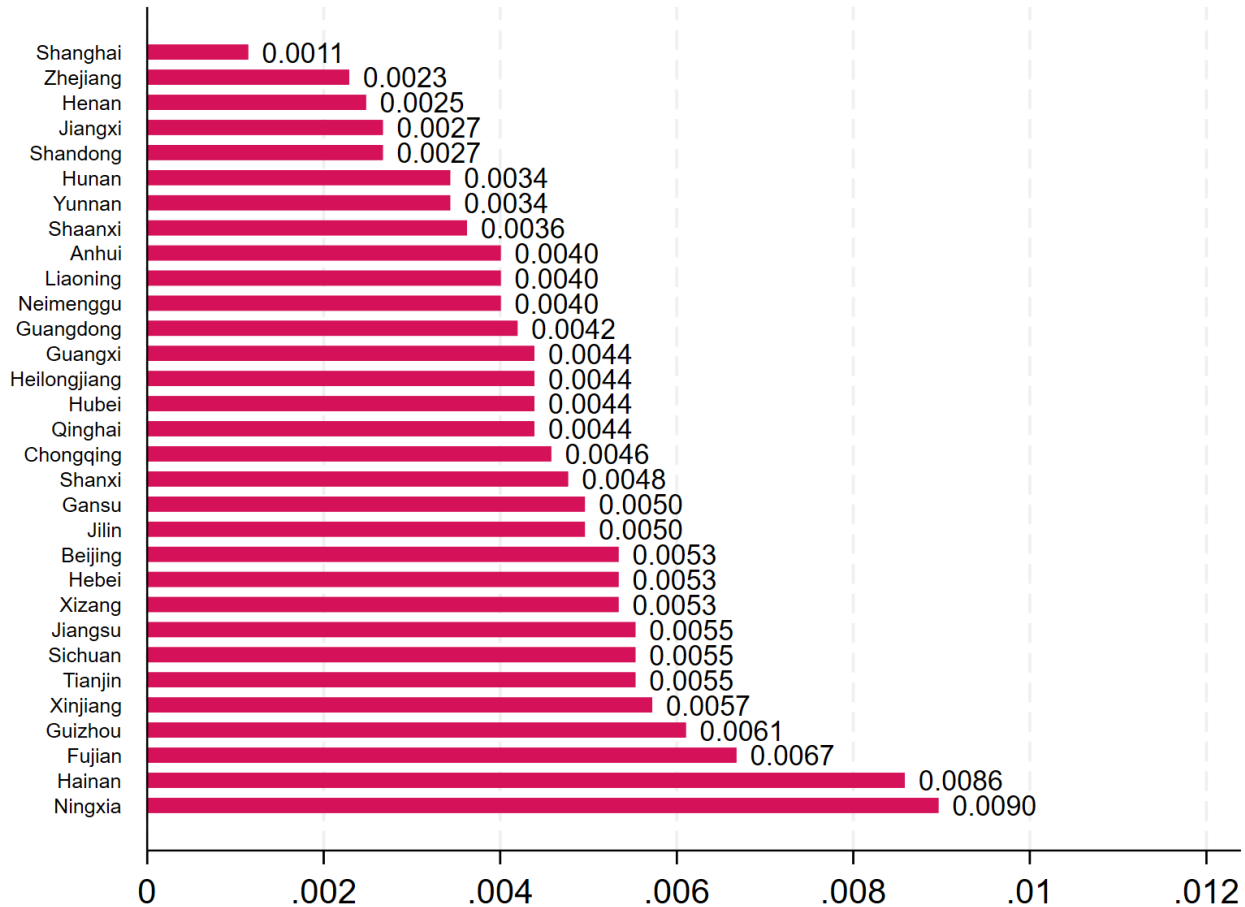


Note: Centrality is the degree centrality measured by the average number of logistics firms across outbound routes.

Counterfactual: 10% decrease in F_{ij}

- Remote provinces like Ningxia, Xinjiang, Guizhou tend to gain more

Figure: Welfare Effect of Reduction in Fixed Transportation Entry Cost



Conclusion

- **Facts:** E-commerce delivery markets exhibit spatial heterogeneity in pricing and firm entry; larger routes attract more firms at lower prices.
- **Model:** Endogenous logistics networks with heterogeneous firms and scale economies generate pro-competitive pricing and a pecking order in entry.
- **Evidence:** IV estimates confirm competition lowers prices and generates network spillovers to neighboring routes.
- **Policy:** Reduction in the entry costs faced by transportation firms disproportionately benefits remote provinces, reducing regional inequality.

Thank you!