

# Geographic Integration and Firm Exports: Evidence from China

Teaching Slides

Lin Tian (INSEAD & CEPR)  
Yue Yu (University of Toronto)

October 2026

# Motivation

Exporting requires **knowledge**: product standards, buyer preferences, regulatory requirements, distribution channels, specialized inputs, ...

- This knowledge is *unevenly distributed* across space
- The resulting **information frictions** are quantitatively important barrier to trade
  - ▶ Often as large as tariffs or transport costs (e.g., Allen, 2014; Startz, 2024)
  - ▶ Particularly severe in developing countries (Atkin and Khandelwal, 2020)
- Yet we know relatively little about *how to reduce them*
  - ▶ Transport infrastructure is a natural lever, but the literature has focused its effects on **shipping costs** (e.g., Donaldson, 2018; Fajgelbaum & Redding, 2022; Fan et al., 2023)
  - ▶ Information frictions are different: knowledge travels with **people**, not goods  $\Rightarrow$  *passenger* infrastructure may matter

# An Illustrative Anecdote (NYT, Sept 24 2013)

Study world's large passenger rail project: China's **high-speed rail (HSR)** network

## • Speedy Trains Transform China

- ▶ Li Qingfu: sales manager of a women's apparel exporter based in *Changsha*
- ▶ Used to visit *Guangzhou*—China's apparel hub—twice a year
- ▶ After HSR connection: goes *almost every month*



*'More frequent [trips] ... have allowed me to more quickly pick up on fashion changes in color and style. My orders have increased by 50 percent.'*

# This Paper

We use the staggered rollout of China's **high-speed rail (HSR)** network to study how domestic connectivity affects firm exports

- **What we do:** link the HSR network to the universe of Chinese exporters
  - ▶ Connectivity: access to *same-sector* export peers in HSR-connected cities
  - ▶ Sector-specificity: same-sector exporters overlap in products, destinations, inputs, and compliance—a proxy for **relevant** contacts
  - ▶ Triple-difference using within-city, cross-sector variation
- **What we find:** HSR-driven connectivity raises firm exports
  - ▶ Higher revenue, driven by larger volumes and lower unit prices
  - ▶ Entry into more destinations and broader product scope
  - ▶ *Within* a firm, exports grow more to those destinations where the firm becomes better connected to same-sector peers *serving that market*
    - ★ a destination-specific effect
- **Why information?**
  - ▶ Rule out non-information channels: no effect on domestic sales, freight costs, employment
  - ▶ Heterogeneity: stronger effects for firms facing greater information barriers
  - ▶ Dimensions: information about *destinations* and *specialized inputs*; less about buyer

# Related Literature

- **Information frictions in trade:** Allen (2014); Fernandes & Tang (2014); Cai and Szeidl (2017); Steinwender (2018); Atkin and Khandelwal (2020); Eaton et al. (2021); Startz (2024)
  - ▶ We show that a specific *policy lever* (passenger infrastructure) reduces these frictions.
- **Transport networks and information exchange:** Giroud (2013); Campante & Yanagizawa-Drott (2017); Bernard, Moxnes & Saito (2019)
  - ▶ We extend these insights to firms' *export performance*.
- **Domestic geography and trade:** Autor et al. (2013); Kovak (2013); Donaldson (2018); Cosar & Demir (2016); Fajgelbaum & Redding (2022)
  - ▶ We focus on how domestic integration affects exports, through *information channel*.
- **Economic effects of China's HSR:** Qin (2017); Lin (2017); Dong et al. (2019); Hanley et al. (2022); Barwick et al. (2022)
  - ▶ We provide firm-level evidence on the *export margin*.

# A Simple Framework to Fix Ideas

- **Setup:** Stylized two-country Melitz (2003)
  - ▶ CES demand across firms with EOS  $\sigma > 1$
  - ▶ Firm  $z$  has: marginal production cost ( $c_z$ ); iceberg trade cost ( $\tau_s$ ); and fixed cost of exporting ( $f_z^x$ )
- Firm has an **export-specific** capacity ( $\varphi_z^x$ )
  - ▶ Ability to overcome frictions specific to foreign markets (search, contracting, compliance, distribution, ...)
- Log export revenue:

$$r_s^x(z) = \text{const} + \log E_s^x + (\sigma - 1) \log P_s^x + (1 - \sigma) \log \tau_s + (1 - \sigma) \log c_z + (\sigma - 1) \log \varphi_z^x$$

- Firm exports if profits cover fixed cost:

$$\pi_s^x(z) > 0 \iff \kappa \left( \frac{E_s^x P_s^x \varphi_z^x}{c_z \tau_s} \right)^{\sigma-1} > f_z^x$$

# Channels and Predictions

**Central hypothesis:** improved domestic connectivity raises firm-level export performance

The framework identifies *three channels*:

- (i)  $\downarrow c_z$ : marginal production cost (better inputs or production techniques)
- (ii)  $\downarrow f_z^x$ : fixed cost of entering foreign markets (learning regulations, distribution channels)
- (iii)  $\uparrow \varphi_z^x$ : export-specific capacity (buyers, standards, quality)

**Predictions:** connectivity raises export revenue (intensive) and the probability of exporting (extensive)

**But the channels differ:** only  $c_z$  moves *domestic* sales;  $\varphi_z^x$  and  $f_z^x$  are export-specific

- $\Rightarrow$  if domestic sales are unaffected, the broad cost channel is unlikely the main driver

# China's HSR Rollout, 2008–2013



(a) 2008



(b) 2010



(c) 2013



(d) 2030 (Expected)

# HSR in China: Four Key Features

## 1. Passenger-only infrastructure

- ▶ Dedicated exclusively to passengers; does not carry freight
- ▶ → improves inter-city *human mobility* without affecting *goods shipping costs*

## 2. Large reductions in travel time

- ▶ Speeds of 250–350 km/h vs. 120–160 on conventional rail.
- ▶ For example, Beijing–Shanghai: 13 hours → under 5 hours
- ▶ National HSR ridership: 300 million (2010) → 830 million (2014); business travelers a substantial share

## 3. Lower information frictions

- ▶ Long-distance day trips become feasible; lowers cost of face-to-face interaction
- ▶ Independent evidence: HSR raises cross-city patent collaboration (Hanley et al., 2022) and academic teamwork (Dong et al., 2020)

## 4. Endogeneity concerns in route planning

- ▶ Criteria for station placement not publicly available; mix of economic and equity considerations (Zheng & Kahn, 2013)
- ▶ → addressed by sector-specific triple-difference

# Data

- **China Customs records (2000-2013):** main outcome data
  - ▶ Universe of Chinese export transactions, regardless of firm/transaction size
  - ▶ Aggregated to Firm  $\times$  HS8  $\times$  destination  $\times$  year
  - ▶ Sample: 338,863 firms exporting at least once before HSR connection
  - ▶ Firm location (prefecture level), firm name, ownership, sector (HS2 level)
- **HSR network data:** China Railway Statistical Yearbook
  - ▶ Opening dates, station locations
  - ▶ Direct trains and one-stop transfers count as connections
  - ▶ Schedules from Chinese National Railway Administration for travel times (robustness)
- **Annual Survey of Industrial Firms (ASIF):** supplementary outcomes
  - ▶ large manufacturing firms (annual sales  $>$  5 million RMB)
  - ▶ Merged with Customs following Yu (2015);  $\sim$  40% of exporters matched
  - ▶ Provides *domestic sales* and *employment* for mechanism analysis

# Measuring Exporter Connectivity

For each city  $c$ , sector  $s$ , and year  $t$ , define:

$$x_{cst} \equiv \sum_{c' \neq c} \underbrace{R_{c'st_0}^x}_{\text{pre-HSR export revenue in same sector, in } c'} \times \underbrace{\mathbb{I}(\text{HSR}_{cc't})}_{\substack{=1 \text{ if } c, c' \\ \text{HSR-connected}}}$$

- **Pre-HSR weights:** avoids reverse causality from current exports
- **Exclude own city:** avoids mechanical correlation with the outcome
- **Sector-specific:** same-sector peers overlap in products, destinations, inputs, logistics, and compliance  $\Rightarrow$  connectivity to them proxies access to *sector-relevant* contacts and information
  - ▶ Also key for identification: same HSR link generates *different* shocks across sectors within a city
- Use inverse travel time for robustness:  $x_{cst} \equiv \sum_{c' \neq c} R_{c'st_0}^x \times (T(\text{Train}_{cc't}))^{-1}$

# Empirical Specification

Baseline regression:

$$y_{zt} = \beta x_{c(z)s(z)t} + \alpha_z + \alpha_{s(z)t} + \alpha_{c(z)t} + \iota_{zt}$$

- $y_{zt}$ : outcome for firm  $z$  in year  $t$
- $x_{c(z)s(z)t}$ : city-sector connectivity, normalized by its standard deviation
  - ▶  $\beta$ : effect of a **1 s.d. increase in connectivity**
- Three sets of fixed effects:
  - ▶ **Firm FE**  $\alpha_z$ : absorbs time-invariant firm characteristics
  - ▶ **Sector-year FE**  $\alpha_{s(z)t}$ : absorbs sector-wide shocks each year
  - ▶ **City-year FE**  $\alpha_{c(z)t}$ : absorbs *any* time-varying shock common to all firms in a city
    - ★ Including the *direct effect of receiving an HSR station*, local demand shocks, city-level policies

# Identification

$$y_{zt} = \beta x_{c(z)s(z)t} + \alpha_z + \alpha_{s(z)t} + \alpha_{c(z)t} + \iota_{zt}$$

- **Triple-difference estimation model:**

- (1) Before vs. after HSR
- (2) Cities with more vs. less connectivity change
- (3) Sectors with more vs. less same-sector exposure in newly connected cities

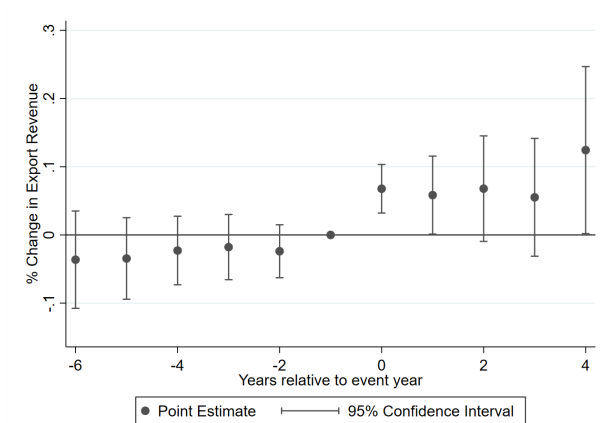
⇒ Remaining identifying variation: *across sectors within the same city-year*

- **Identifying assumption (parallel trends):** absent HSR, export performance would evolve similarly across sectors within the same city

# Event Study: Pre-Trends and Dynamics

Given continuous treatment, we follow Callaway et al. (2024) to compare **high-** vs. **low-exposure** sectors

$$y_{zt} = \sum_k \beta^k \mathbb{I}(\text{HSR}_{c(z)t}^k) \times \mathbb{I}(\text{HighSector}_s(z)) + \alpha_{s(z)t} + \alpha_z + \alpha_{c(z)t} + \iota_{zt}$$



# Baseline: Connectivity Raises Firm Exports

$$y_{zt} = \beta x_{c(z)s(z)t} + \alpha_z + \alpha_{s(z)t} + \alpha_{c(z)t} + \iota_{zt}$$

| Dep. var.    | (1)<br>Revenue<br>$r_{zt}^x$ | (2)<br>Destinations<br>$N_{zt}^x$ | (3)<br>Volume<br>$q_{zt}^x$ | (4)<br>Price<br>$p_{zt}^x$ |
|--------------|------------------------------|-----------------------------------|-----------------------------|----------------------------|
| $x_{cst}$    | 0.042***<br>(0.007)          | 0.202***<br>(0.038)               | 0.087***<br>(0.011)         | -0.045***<br>(0.007)       |
| Observations | 2,032,156                    | 2,032,156                         | 2,028,602                   | 2,028,602                  |
| $R^2$        | 0.738                        | 0.805                             | 0.798                       | 0.852                      |

- A 1 s.d. increase in connectivity raises export revenue by **4.2%**
- $\uparrow$  volume (+8.7%) and  $\downarrow$  unit price (-4.5%)  $\Rightarrow$  consistent with lower cost of serving foreign markets
- Plus entry into **0.20** more destinations  $\Rightarrow$  extensive-margin expansion

Product-level results

# How Do Firms Achieve the Revenue Gain?

The baseline gives the *total* revenue effect. But *how* do firms get there—by selling more where they already are, or by changing *where* they sell?

To look at the destination margin, take a **long difference** (2007 → 2013):

$$\Delta_t y_z = \theta \Delta_t x_{c(z)s(z)} + \alpha_{c(z)} + \alpha_{s(z)} + \iota_z$$

- $\Delta_t y_z$ : change in firm  $z$ 's export revenue (or a given margin) over the period
- Decompose the total change into four margins that sum exactly to it (entry and exit are defined over the interval)

**Intensive:** within continuing destinations + reallocation across them

**Extensive:** entry into new destinations – exit from dropped ones

# Export Revenue Decomposition

|                   | (1)<br>Total        | (2)<br>Intensive | (3)<br>Across       | (4)<br>Extensive    | (5)<br>Exit         |
|-------------------|---------------------|------------------|---------------------|---------------------|---------------------|
|                   | $\Delta_t r_s^x$    | Within           | Across              | Entry               | Exit                |
| $\Delta_t x_{CS}$ | 0.054***<br>(0.012) | 0.004<br>(0.011) | 0.017***<br>(0.005) | 0.019***<br>(0.006) | 0.014***<br>(0.005) |
| Share of total    | 100%                | 7.4%             | 31.5%               | 35.2%               | 25.9%               |

- Within-market growth is **small and insignificant** (7%)
- The action is in *reconfiguring the destination portfolio*: entry (35%), reallocation (32%), exit (26%)

⇒ firms change *which* markets they serve, not how deeply they serve existing ones

## Within-Firm Cross-Destination: Setup

The decomposition is *across* firms. Now go *within* a firm: does connectivity operate destination-by-destination?

**Add a destination dimension** to the connectivity measure:

$$x_{csnt} \equiv \sum_{c' \neq c} R_{c'snt_0}^x \times \mathbb{I}(\text{HSR}_{cc't})$$

**Within-firm long difference:**

$$\Delta_t y_{zn} = \alpha_z + \alpha_{s(z)n} + \beta \Delta_t x_{c(z)s(z)n} + \iota_{zn}$$

- Firm FE  $\alpha_z$  absorbs *all* firm-level changes; identification is across destinations within a firm
- Assumption: absent HSR, a firm's export growth across destinations would have been parallel

## Within-Firm Cross-Destination: Results

|                    | (1)<br>Intensive ( $\Delta_t r_{z,n}^x$ ) | (2)<br>Extensive ( $\Delta_t \mathbb{I}_{z,n}$ ) |
|--------------------|-------------------------------------------|--------------------------------------------------|
| $\Delta_t x_{csn}$ | 0.021***<br>(0.006)                       | 0.003***<br>(0.001)                              |
| Observations       | 492,591                                   | 22,322,216                                       |
| $R^2$              | 0.280                                     | 0.049                                            |

- Within a firm, exports grow more to destinations with larger connectivity gains
- Firms are also more likely to *enter* those destinations
- Complements the baseline: isolates *destination-specificity*, while the baseline recovers the *aggregate* effect

# Robustness

- **Concurrent policies & infrastructure**

[Details](#)

- ▶ Hukou, minimum wage, college expansion, industrial parks, VAT rebates, broadband

- **Persistent trends:** future connectivity does not predict past export growth

[Details](#)

- **Endogenous HSR placement:** drop each city's largest sector

[Details](#)

- **Alternative connectivity measures:** continuous travel time, IHS transformation, alternative weights, lagged treatment

[Details](#)

- **Staggered treatment**

[Details](#)

- ▶ Borusyak et al. (2022), Sun & Abraham (2021) on a discretized treatment

- **Sample composition**

[Details](#)

- ▶ Balanced panel, drop never-HSR cities, drop SOEs

- **Inference**

[Clustering](#)

- ▶ Alternative clustering; permutation test (200 simulations)

# What Is the Mechanism?

**Leading hypothesis:** improved passenger connectivity helps less well-connected firms access **export-relevant information** concentrated in established hubs

We assess this in three steps:

1. **Rule out non-information channels**

Are domestic sales, freight costs, or employment affected?

2. **Heterogeneity by information barriers**

Do firms facing greater information frictions respond more?

3. **Dimensions of information**

What *kind* of information do firms acquire: about destinations, inputs, buyers?

## Step 1: Not Broad Cost Reductions or Labor Pooling

Each alternative channel has a distinct testable implication:

- Broad production-cost cut ( $\downarrow c_z$ )  $\Rightarrow$  *domestic* sales should rise too
- Labor market pooling  $\Rightarrow$  *employment* should rise

|           | Exports<br>$r_{zt}^x$ | Domestic sales<br>$r_{zt}^d$ | Employment<br>$\ln(\text{Workers})$ | Exports<br>(baseline) |
|-----------|-----------------------|------------------------------|-------------------------------------|-----------------------|
| $x_{cst}$ | 0.034***<br>(0.012)   | 0.004<br>(0.010)             | 0.008<br>(0.007)                    | 0.042***<br>(0.007)   |

Domestic sales and employment from ASIF (matched sample).

- Exports rise, but **domestic sales and employment are unaffected**
- $\Rightarrow$  neither broad production cost nor labor pooling can be the main driver

Skill access

## Step 1: Not Cheaper Freight

We test the impact of connectivity on shipping times for outputs and inputs:

- **Outputs:** time to ship firm's finished goods to export ports
- **Inputs:** time to ship supplies in from firm's suppliers, proxied by
  - ▶ **Input–Output (IO):** sectors that supply the firm's sector in national IO tables
  - ▶ **Potential Domestic Suppliers (PDS):** a firm's imports reveal its input needs; suppliers are domestic firms exporting those same products (Kee et al., 2024)

|           | Outputs to ports | Inputs from suppliers |                  |
|-----------|------------------|-----------------------|------------------|
|           |                  | IO-based              | PDS-based        |
| $x_{cst}$ | 0.001<br>(0.002) | −0.001<br>(0.002)     | 0.001<br>(0.002) |

- No effect on shipping costs  $\Rightarrow$  the classic trade-cost channel plays a limited role

Measure details

## Step 2: Larger Effects Where Information Barriers Are Higher

If HSR works by easing information frictions, the effect should be **larger for firms facing greater barriers** to export-relevant knowledge

$$y_{zt} = \alpha_z + \alpha_{s(z)t} + \alpha_{c(z)t} + \beta x_{c(z)s(z)t} + \gamma x_{c(z)s(z)t} \times \mathbb{I}(\text{Firm Type}) + \mathbb{I}(\text{Firm Type}) \times \alpha_t + \iota_{zt}$$

- $\gamma$  captures whether a given firm type responds *more* to connectivity
- $\mathbb{I}(\text{Firm Type}) \times \alpha_t$  allows differential time trends across types

## Step 2: Heterogeneity Analysis Results

- Five proxies for facing greater information barriers:
  - ▶ Firm size: smaller firms respond more [Details](#)
  - ▶ Product type: firms exporting more style-sensitive products respond more [Details](#)
  - ▶ Location: firms locating in more remote regions respond more [Details](#)
  - ▶ Destination: firms exporting to less open countries respond more [Details](#)
  - ▶ Exporter type: firms not in processing trade respond more [Details](#)
- All five point the same way: **firms facing greater information frictions respond more**, consistent with an information story

## Step 3: Dimensions of Information—Quality

*Which dimensions* of export knowledge does connectivity carry?

First: **product quality**. If firms learn foreign standards and what buyers want, they should upgrade what they export

|                 | Overall             | Intensive           | Extensive         |
|-----------------|---------------------|---------------------|-------------------|
| $\Delta x_{cs}$ | 0.058***<br>(0.016) | 0.042***<br>(0.010) | 0.016*<br>(0.008) |

Quality inferred from price-quantity data (Khandelwal et al., 2013). [Details](#)

- Connectivity raises the quality of exported products
- $\sim 2/3$  from **upgrading within** continuing destinations;  $\sim 1/3$  from **entering higher-quality** markets

[Full decomposition](#)

## Step 3: Dimensions of Information—Destination Convergence

- If firms learn *which markets to serve* from their connected peers, their export destinations should converge toward those of the peers
  - ▶ Compare each firm's change in destination mix (2007–2013) with that of same-sector exporters in newly connected cities
  - ▶ Similarity: dot product, cosine, covariance

|                          | Dot product         | Cosine              | Covariance          |
|--------------------------|---------------------|---------------------|---------------------|
| $\mathbb{I}(\text{HSR})$ | 0.393***<br>(0.041) | 2.303***<br>(0.295) | 0.530***<br>(0.036) |

- Connected firms' destinations **converge** toward those of their same-sector peers
- But convergence alone could reflect common shocks, not learning (Manski 1993)

## Step 3: Convergence Is Directional

- Genuine learning or just reflection?
- If firms *learn* from stronger peers, weak firms should move toward strong cities, not the reverse

| $\mathbb{I}(\text{HSR}) \times$ | Dot product         | Cosine               | Covariance          |
|---------------------------------|---------------------|----------------------|---------------------|
| Low-firm to High-city           | 1.706***<br>(0.061) | 17.186***<br>(1.348) | 0.738***<br>(0.071) |
| High-firm to Low-city           | 0.221***<br>(0.060) | -0.742*<br>(0.445)   | 0.293***<br>(0.053) |
| Same level                      | 0.277***<br>(0.055) | 2.682***<br>(0.357)  | 0.474***<br>(0.049) |

- Weak firms converge sharply toward strong cities; convergence is **much weaker** the other way
- Asymmetric, not symmetric  $\Rightarrow$  rules out reflection (Manski 1993); consistent with *learning*

## Step 3: Dimensions of Information—Supplier Access

Does connectivity to *supplier-rich* cities raise exports, beyond same-sector connectivity (Bernard et al., 2019)?

- Build HSR-enhanced supplier access using sector- and firm-based supplier weights

Construction

|                      | Sector-based        |                     | Firm-based          |                     |
|----------------------|---------------------|---------------------|---------------------|---------------------|
|                      | Revenue             | Dest.               | Revenue             | Dest.               |
| $x_{cst}$            | 0.044***<br>(0.009) | 0.188***<br>(0.040) | 0.051***<br>(0.010) | 0.294***<br>(0.048) |
| $x_{cst}^{Supplier}$ | -0.013<br>(0.015)   | 0.053<br>(0.081)    | 0.022***<br>(0.007) | 0.080**<br>(0.036)  |

- Firm-based supplier access matters; sector-based too coarse
- Baseline connectivity stays large  $\Rightarrow$  supplier access **complements**, doesn't replace

## Step 3: Dimensions of Information—Buyer Access

Does connectivity work through exposure to foreign *buyers* and intermediaries?

- Split connectivity by connections to trade-fair cities, and to trading firms vs. regular exporters

|                                    | Revenue             | Dest.               |
|------------------------------------|---------------------|---------------------|
| Connections to non-fair cities     | 0.052***<br>(0.020) | 0.309***<br>(0.059) |
| Connections to fair-hosting cities | 0.028<br>(0.028)    | 0.040<br>(0.077)    |
| Connections to regular exporters   | 0.057***<br>(0.013) | 0.236***<br>(0.070) |
| Connections to trading firms       | -0.185<br>(0.137)   | -0.343<br>(0.825)   |

- Gains concentrate in non-fair cities and regular exporters, not fairs or intermediaries
- Buyer exposure is **not** a primary driver

# Conclusion

- **Question:** can passenger infrastructure help firms overcome *information* barriers to exporting?
- **Findings:** using China's HSR rollout, connectivity to same-sector exporters raises firm exports—revenue, destinations, product scope—by reshaping *which* markets firms serve
- **Mechanism:** the pattern points to information, not goods movement
  - ▶ No effect on domestic sales, freight costs, employment
  - ▶ Stronger effects where information barriers are greater
  - ▶ Firms upgrade quality, converge on peers' destinations, access suppliers

**Takeaway:** transport infrastructure can facilitate trade not only by moving *goods*, but by moving *people*—and the knowledge they carry

# Appendix

## Robustness: Concurrent Policies

Each control constructed at the *city-sector-year* level, to match the treatment's variation:

- **Hukou**: restrictiveness  $\times$  low-skill intensity  $\times$  export-demand shock
- **Minimum wage**: firm wage exposure  $\times$  local minimum-wage growth (post-2004)
- **College expansion**: province enrollment-capacity shock  $\times$  sector skill share
- **Industrial parks**: sector targeted by a park in the firm's city-year
- **VAT rebate**: province-sector export VAT rebate rate

|           | Hukou               | Min. wage           | College             | Parks               | VAT                 | Baseline            |
|-----------|---------------------|---------------------|---------------------|---------------------|---------------------|---------------------|
| $x_{cst}$ | 0.040***<br>(0.007) | 0.041***<br>(0.007) | 0.035***<br>(0.007) | 0.042***<br>(0.007) | 0.054***<br>(0.016) | 0.042***<br>(0.007) |

Dep. var.: log export revenue. Each column adds one policy control to the baseline.

Estimates stable across all controls

[Broadband](#)[Back](#)

## Robustness: Broadband Internet

**Concern:** internet expansion could independently aid exports and be correlated with HSR rollout

China's fiber backbone was largely built in the 1990s—well before our study period.  
Directly testing the correlation:

|                | Internet use     | log(Internet use) |
|----------------|------------------|-------------------|
| HSR connection | 1.176<br>(4.849) | -0.041<br>(0.036) |
| Observations   | 2,271            | 2,271             |
| $R^2$          | 0.437            | 0.993             |

City and year FE. Internet use = number of internet subscribers (China City Statistical Yearbook).

- HSR connection is *uncorrelated* with internet penetration  $\Rightarrow$  broadband cannot account for the results

## Robustness: Persistent Trends

**Placebo:** regress *pre*-treatment export changes on *future* connectivity changes—if persistent city-sector trends drove results, future connectivity would predict past growth

|                   | $\Delta_{t-1} r_s^x$ | $\Delta_{t-1} N^x$ |
|-------------------|----------------------|--------------------|
| $\Delta_t X_{CS}$ | 0.002<br>(0.009)     | 0.009<br>(0.043)   |
| Observations      | 207,972              | 207,972            |
| $R^2$             | 0.058                | 0.051              |

- Future connectivity does *not* predict past export growth (coefficients  $\approx 0$ , insignificant)
- No evidence that persistent city-sector trends drive the baseline

[Back](#)

# Robustness: Endogenous HSR Placement

**Concern:** a city's largest sector may have influenced where HSR was routed. Drop the largest sector within each city and re-estimate

|           | Drop largest sector |                     | Baseline            |                     |
|-----------|---------------------|---------------------|---------------------|---------------------|
|           | Revenue             | Destinations        | Revenue             | Destinations        |
| $x_{cst}$ | 0.048***<br>(0.008) | 0.182***<br>(0.040) | 0.042***<br>(0.007) | 0.202***<br>(0.038) |

- Excluding the sectors most likely to shape HSR routing leaves the estimates unchanged
- Robust to dropping the largest **3** and **5** sectors in each city

Back

# Robustness: Alternative Connectivity Measures

|                        | Revenue ( $r_{zt}^x$ ) | Destinations ( $N_{zt}^x$ ) |
|------------------------|------------------------|-----------------------------|
| Continuous travel time | 0.030*** (0.006)       | 0.105*** (0.037)            |
| IHS( $x_{cst}$ )       | 0.071*** (0.012)       | 0.318*** (0.064)            |
| Export-share weights   | 0.074*** (0.021)       | 0.418*** (0.097)            |
| Num.-firm weights      | 0.076*** (0.012)       | 0.267*** (0.060)            |
| Lagged treatment       | 0.044*** (0.013)       | 0.288*** (0.101)            |
| Baseline               | 0.042*** (0.007)       | 0.202*** (0.038)            |

- Sign, significance, and rough magnitude robust across all alternative measures

[Back](#)

# Robustness: Staggered Treatment

Recent literature flags bias in two-way FE designs with staggered rollout. We discretize the treatment and re-estimate with robust estimators

|                        | Revenue ( $r_{zt}^x$ ) | Destinations ( $N_{zt}^x$ ) |
|------------------------|------------------------|-----------------------------|
| Standard staggered DDD | 0.054*** (0.020)       | 0.231*** (0.026)            |
| Borusyak et al. (2022) | 0.050** (0.021)        | 0.239*** (0.075)            |
| Sun & Abraham (2021)   | 0.982*** (0.128)       | 0.276*** (0.091)            |

Discretized (binary) treatment; coefficients not directly comparable to the continuous baseline.

- Consistent and significant on both margins across all estimators

Back

# Robustness: Sample Composition

|                                     | Revenue ( $r_{zt}^x$ ) | Destinations ( $N_{zt}^x$ ) |
|-------------------------------------|------------------------|-----------------------------|
| Balanced panel (export 2007 & 2013) | 0.047*** (0.009)       | 0.229*** (0.047)            |
| Drop never-HSR cities               | 0.046*** (0.008)       | 0.210*** (0.040)            |
| Drop state-owned enterprises        | 0.040*** (0.007)       | 0.206*** (0.034)            |
| Baseline                            | 0.042*** (0.007)       | 0.202*** (0.038)            |

- Estimates stable across all sample definitions

[Back](#)

# Robustness: Alternative Clustering

Baseline clusters standard errors at the city level. We re-cluster at the *city-sector-year* level

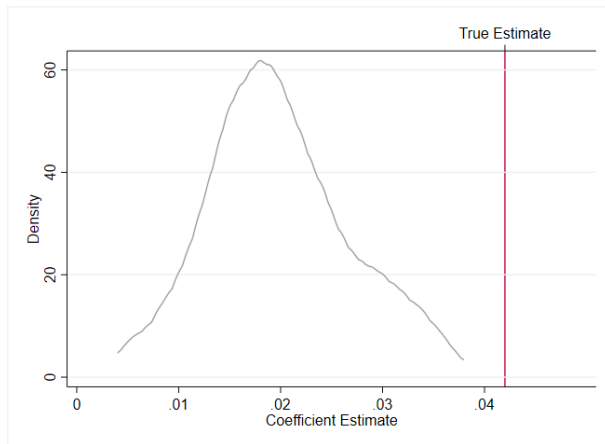
|           | City-sector-year    |                     | Baseline (city)     |                     |
|-----------|---------------------|---------------------|---------------------|---------------------|
|           | Revenue             | Destinations        | Revenue             | Destinations        |
| $x_{cst}$ | 0.042***<br>(0.007) | 0.202***<br>(0.034) | 0.042***<br>(0.007) | 0.202***<br>(0.038) |

- Significance essentially unchanged under alternative clustering

Back

## Robustness: Permutation Test

**Test:** randomly reassign HSR connections across city-pairs, preserving the number of treated pairs each year; reconstruct connectivity and re-estimate, over 200 simulations



- None of the 200 placebo estimates exceeds the actual coefficient  $\Rightarrow p = 0$

# Heterogeneity: Firm Size

- Smaller exporting firms respond more to the HSR driven improvement in access to other exporting firms.

|                           | (1)                 | (2)                 |
|---------------------------|---------------------|---------------------|
| Dependent Variable        | log(export revenue) |                     |
| Independent Variable      | Measure 1           | Measure 2           |
| Access to Exporters       | 0.016*<br>(0.009)   | 0.011*<br>(0.050)   |
| Access to Exporters×Small | 0.060**<br>(0.024)  | 0.090***<br>(0.027) |
| Observations              | 2,032,156           | 2,032,156           |
| R-squared                 | 0.740               | 0.740               |

# Heterogeneity: Product Complexity

- Startz (2019): some products, such as apparel and electronics, have style evolve more quickly than others.
- Firms exporting products with more quickly-evolving styles respond more to the HSR driven improvement in access to other exporting firms.
- Results robust to using Rauch (1999) index

|                            | (1)                 | (2)                  |
|----------------------------|---------------------|----------------------|
| Dependent Variable         | log(export revenue) |                      |
| Independent Variable       | Measure 1           | Measure 2            |
| Access to Exporters        | 0.006<br>(0.011)    | 0.015*<br>(0.009)    |
| Access to Exporters×Simple | -0.026*<br>(0.014)  | -0.025***<br>(0.010) |
| Observations               | 2,023,156           | 2,023,156            |
| R-squared                  | 0.738               | 0.738                |

# Heterogeneity: Firms in Core vs. Peripheral Cities

- Export firms located in peripheral cities respond more to the HSR driven improvement in access to other exporting firms.

|                               | (1)                 | (2)                  |
|-------------------------------|---------------------|----------------------|
| Dependent Variable            | log(export revenue) |                      |
| Independent Variable          | Measure 1           | Measure 2            |
| Access to Exporters           | 0.053***<br>(0.010) | 0.041***<br>(0.008)  |
| Access to Exporters×Core City | -0.026*<br>(0.014)  | -0.025***<br>(0.010) |
| Observations                  | 2,032,156           | 2,032,156            |
| R-squared                     | 0.738               | 0.738                |

# Heterogeneity: Openness of Export Destinations

- Less acute information frictions when exporting to more open and easily accessed markets (Lovely et al., 2005)
  - ▶ Classify export destinations by openness (trade per dollar GDP, controlling for country size and proximity to economic activity outside of the country)
- Firms exporting to less open destinations respond more to the HSR driven improvement in access to other exporting firms.

| Dependent Variable                       | (1)                  | (2)                  |
|------------------------------------------|----------------------|----------------------|
| Independent Variable                     | log(export revenue)  |                      |
|                                          | Measure 1            | Measure 2            |
| Access to Exporters                      | 0.082***<br>(0.012)  | 0.064***<br>(0.009)  |
| Access to Exporters×Destination Openness | -0.068***<br>(0.013) | -0.055***<br>(0.012) |
| Observations                             | 1,974,606            | 1,974,606            |
| R-squared                                | 0.735                | 0.735                |

# Heterogeneity: Processing Trade

- Firms in ordinary trade respond more to the HSR driven improvement in access to other exporting firms than those in processing trade

| Dependent Variable          | (1)<br>log(export revenue) | (2)                  |
|-----------------------------|----------------------------|----------------------|
| Independent Variable        | Measure 1                  | Measure 2            |
| Access to Exporters         | 0.047***<br>(0.007)        | 0.033***<br>(0.006)  |
| Access to Exporters×Process | -0.399***<br>(0.116)       | -0.288***<br>(0.100) |
| Observations                | 2,032,156                  | 2,032,156            |
| R-squared                   | 0.739                      | 0.739                |

# Heterogeneity: Ownership Status

- No significant difference between SOEs and privately owned firms

|                             | (1)                 | (2)                 |
|-----------------------------|---------------------|---------------------|
| Dependent Variable          | log(export revenue) |                     |
| Independent Variable        | Measure 1           | Measure 2           |
| Access to Exporters         | 0.052***<br>(0.008) | 0.028***<br>(0.006) |
| Access to Exporters×Private | -0.003<br>(0.012)   | -0.025<br>(0.015)   |
| Observations                | 2,032,156           | 2,032,156           |
| R-squared                   | 0.740               | 0.740               |

# Destination Level Analysis

- We consider impact of  $x_{cst}$  on firm export revenue to a specific destination  $n$

$$y_{zsnt} = \alpha_{zn} + \alpha_{snt} + \alpha_{ct} + \beta x_{cst} + \iota_{znt},$$

|              | (1)                 | (2)                 |
|--------------|---------------------|---------------------|
|              | log(export revenue) |                     |
|              | Measure 1           | Measure 2           |
| $x_{cst}$    | 0.012*<br>(0.006)   | 0.004***<br>(0.000) |
| Observations | 15,717,657          | 15,717,657          |
| R-squared    | 0.726               | 0.726               |

Robust standard errors clustered by prefectures in parentheses

\*\*\*  $p < 0.01$ , \*\*  $p < 0.05$ , \*  $p < 0.1$

# Lagged Treatment

| VARIABLES                      | (1)<br>log(export revenue) | (2)<br>log(export revenue) | (3)<br>No. exp destinations | (4)<br>No. exp destinations |
|--------------------------------|----------------------------|----------------------------|-----------------------------|-----------------------------|
| Measure 1                      | 0.044***<br>(0.013)        |                            | 0.288***<br>(0.101)         |                             |
| Measure 2                      |                            | 0.027***<br>(0.008)        |                             | 0.180**<br>(0.074)          |
| Observations                   | 2,032,156                  | 2,032,156                  | 2,032,156                   | 2,032,156                   |
| R-squared                      | 0.738                      | 0.738                      | 0.805                       | 0.805                       |
| *** p<0.01, ** p<0.05, * p<0.1 |                            |                            |                             |                             |

Back

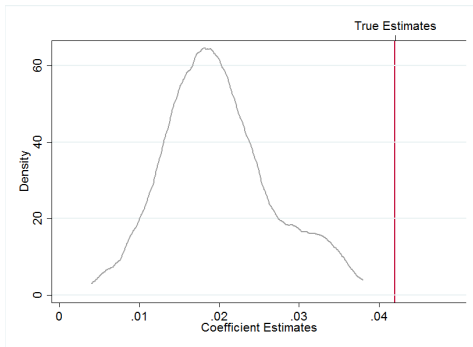
## Exclude the Largest Sector

| VARIABLES                      | (1)<br>log(export revenue) | (2)                 | (3)<br>No. exp destinations | (4)                 |
|--------------------------------|----------------------------|---------------------|-----------------------------|---------------------|
| Measure 1                      | 0.048***<br>(0.008)        |                     | 0.182***<br>(0.040)         |                     |
| Measure 2                      |                            | 0.030***<br>(0.007) |                             | 0.145***<br>(0.036) |
| Observations                   | 1,740,255                  | 1,740,255           | 1,740,255                   | 1,740,255           |
| R-squared                      | 0.734                      | 0.734               | 0.802                       | 0.802               |
| *** p<0.01, ** p<0.05, * p<0.1 |                            |                     |                             |                     |

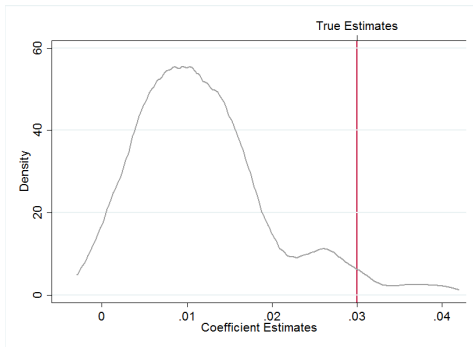
Back

# Permutation Test

- The p-values derived from the permutation tests are quantitatively similar to the baseline results.



(a) Measure1



(b) Measure2

*Notes:* For the first measure, we randomly select a set of city pairs for each year during 2008 and 2013. The number of connected city pairs is specified to be the same as in reality in the corresponding year. For the second measure, we reshuffle the reduction of commuting time between city pairs caused by the HSR network expansion and re-compute the second measure. The distribution of coefficient estimates are based on 200 simulations. The p-values of the estimates are 0 for Measure 1 and 0.02 for Measure 2, both of

# No Control of City Time-varying Effects

| VARIABLES    | (1)<br>log(export revenue) | (2)                | (3)<br>No. exp destinations | (4)              |
|--------------|----------------------------|--------------------|-----------------------------|------------------|
| Measure 1    | 0.026*<br>(0.013)          |                    | 0.052<br>(0.063)            |                  |
| Measure 2    |                            | 0.033**<br>(0.014) |                             | 0.102<br>(0.065) |
| Observations | 2,032,190                  | 2,032,190          | 2,032,190                   | 2,032,190        |
| R-squared    | 0.735                      | 0.735              | 0.802                       | 0.802            |

\*\*\*  $p < 0.01$ , \*\*  $p < 0.05$ , \*  $p < 0.1$

Back

# A Balanced Panel of Firms

| VARIABLES                      | (1)<br>log(export revenue) | (2)                 | (3)<br>No. exp destinations | (4)                 |
|--------------------------------|----------------------------|---------------------|-----------------------------|---------------------|
| Measure 1                      | 0.047***<br>(0.009)        |                     | 0.229***<br>(0.047)         |                     |
| Measure 2                      |                            | 0.040***<br>(0.008) |                             | 0.170***<br>(0.057) |
| Observations                   | 975,743                    | 975,743             | 975,743                     | 975,743             |
| R-squared                      | 0.751                      | 0.751               | 0.842                       | 0.842               |
| *** p<0.01, ** p<0.05, * p<0.1 |                            |                     |                             |                     |

Back

# Dropping Cities with no HSR station by 2016

| VARIABLES    | (1)<br>lnexp        | (2)<br>lnexp        | (3)<br>N_destination | (4)<br>N_destination |
|--------------|---------------------|---------------------|----------------------|----------------------|
| $X_{cst}$    | 0.046***<br>(0.008) | 0.032***<br>(0.006) | 0.210***<br>(0.040)  | 0.112***<br>(0.039)  |
| Observations | 1,834,222           | 1,834,222           | 1,834,222            | 1,834,222            |
| R-squared    | 0.740               | 0.740               | 0.807                | 0.807                |
| Firm FEs     | Y                   | Y                   | Y                    | Y                    |
| CityXyr FEs  | Y                   | Y                   | Y                    | Y                    |
| SecXyr FEs   | 97 sectors          | 97 sectors          | 97 sectors           | 97 sectors           |

Robust standard errors in parentheses

\*\*\*  $p < 0.01$ , \*\*  $p < 0.05$ , \*  $p < 0.1$

# Internet Connectivity

- Internet not positively correlated with HSR connection

| VARIABLES    | (1)<br>Internet usage | (2)<br>ln(Internet usage) |
|--------------|-----------------------|---------------------------|
| HSR dummy    | 1.176<br>(4.849)      | -0.041<br>(0.036)         |
| Observations | 2,271                 | 2,271                     |
| R-squared    | 0.437                 | 0.993                     |

\*\*\*  $p < 0.01$ , \*\*  $p < 0.05$ , \*  $p < 0.1$

# Clustering Standard Errors at City-Sector-Year Level

| VARIABLES    | (1)<br>Inexp        | (2)<br>Inexp        | (3)<br>N_destination | (4)<br>N_destination |
|--------------|---------------------|---------------------|----------------------|----------------------|
| $X_{cst}$    | 0.042***<br>(0.007) | 0.030***<br>(0.007) | 0.202***<br>(0.034)  | 0.105**<br>(0.044)   |
| Observations | 2,032,156           | 2,032,156           | 2,032,156            | 2,032,156            |
| R-squared    | 0.738               | 0.738               | 0.805                | 0.805                |
| Firm FEs     | Y                   | Y                   | Y                    | Y                    |
| CityXyr FEs  | Y                   | Y                   | Y                    | Y                    |
| SecXyr FEs   | 97 sectors          | 97 sectors          | 97 sectors           | 97 sectors           |

Robust standard errors in parentheses

\*\*\*  $p < 0.01$ , \*\*  $p < 0.05$ , \*  $p < 0.1$

# Alternative Mechanisms: Measure Details

## 1. Changes in access to suppliers, induced by HSR connection

$$\sum_{s'} \psi_{s'}^s \sum_{c' \neq c} \left[ \sum_{z' \in \mathcal{Z}_{s' t_0}(c')} r_{s' t_0}(z') \right] f(d_{cc' t}),$$

where  $\psi_{s'}^s$  is the input share from sector  $s'$  for a firm in sector  $s$ .

## 2. Changes in access to port, induced by HSR connection

- ▶ Define  $\mathbb{C}_z^{port}$  as the set of port locations that exporter  $z$  uses
- ▶ Define access to port:

$$\sum_{port(z) \in \mathbb{C}_z^{port}} f(d_{cc' t}),$$

# Alternative Mechanisms: Measure Details

## 3. Changes in access to production activities, induced by HSR connection

- ▶ Define access to general production activities

$$\sum_{c' \neq c} \left[ \sum_{z' \in \mathcal{Z}_{st_0}(c')} VA_{st_0}(z') \right] f(d_{cc't})$$

- ▶ Split the measure (using TFP) into access to the most productive firms and the rests

$$\sum_{c' \neq c} \left[ \sum_{z' \in \mathcal{Z}_{st_0}(c')} VA_{st_0}^{\text{top quartile}}(z') \right] f(d_{cc't}) + \sum_{c' \neq c} \left[ \sum_{z' \in \mathcal{Z}_{st_0}(c')} VA_{st_0}^{\text{rest}}(z') \right] f(d_{cc't})$$

## Extensive Margin: Export to A Destination or Not

|                      | (1)                 | (2)                 |
|----------------------|---------------------|---------------------|
| Dependent Variable   | Export dummy        | Export dummy        |
| Independent Variable | Measure 1           | Measure 2           |
| $X_{cst}$            | 0.003***<br>(0.001) | 0.025***<br>(0.004) |
| Observations         | 12,153,853          | 12,153,853          |
| R-squared            | 0.710               | 0.710               |

*Notes:* We control for firm $\times$ destination fixed effects, city time-varying effects, sector time-varying effects and destination time-varying effects. We cluster the error terms at the city level. Robust standard errors are in parentheses: \*\*\*  $p < 0.01$ , \*\*  $p < 0.05$ , and \*  $p < 0.1$ .

## No Significant Effects on FTCs

| VARIABLES                                    | (1)<br>lnexp     | (2)<br>lnexp     |
|----------------------------------------------|------------------|------------------|
| $x_{cst}$                                    | 0.017<br>(0.016) | 0.016<br>(0.016) |
| Observations                                 | 491,074          | 491,074          |
| R-squared                                    | 0.712            | 0.712            |
| Firm FEs                                     | Y                | Y                |
| CityYr FEs                                   | Y                | Y                |
| SecYr FEs                                    | 97 sectors       | 97 sectors       |
| Robust standard errors in parentheses        |                  |                  |
| *** $p < 0.01$ , ** $p < 0.05$ , * $p < 0.1$ |                  |                  |

- Evidence consistent with export insights to overcome **production**-specific barriers, rather than distribution-specific barriers

## Pairwise Analysis: Quality

- Use destination market GDP/capita to proxy product quality (Khandelwal, 2008)
- No significant impact on average; positive impact to firms in core cities

| VARIABLES         | (1)                | (2)                | (3)                   | (4)                 |
|-------------------|--------------------|--------------------|-----------------------|---------------------|
|                   | Measure 1          |                    | Measure 2             |                     |
| $X_{cst}$         | 46.764<br>(41.272) | 60.975<br>(39.317) | -21.911<br>(42.106)   | 12.191<br>(35.866)  |
| $X_{cst}X_{core}$ |                    |                    | 166.045**<br>(77.952) | 115.253<br>(73.389) |
| Firm FEs          | Y                  | Y                  | Y                     | Y                   |
| CityXyr FEs       | Y                  | Y                  | Y                     | Y                   |
| SecXyr FEs        | 97 sectors         | 97 sectors         | 97 sectors            | 97 sectors          |
| Observations      | 2,032,156          | 2,032,156          | 2,032,156             | 2,032,156           |
| R-squared         | 0.788              | 0.788              | 0.788                 | 0.788               |

Robust standard errors in parentheses

\*\*\*  $p < 0.01$ , \*\*  $p < 0.05$ , \*  $p < 0.1$

# Controlling for Access to Goods

## 1. Better access to upstream suppliers (Bernard et al, 2019)

- ▶ Control for changes in access to suppliers, induced by HSR connection

$$\sum_{s'} \psi_{s'}^s \sum_{c' \neq c} \left[ \sum_{z' \in \mathcal{Z}_{s't_0}(c')} r_{s't_0}^x(z') \right] \mathbb{I}(HSR_{cc't}),$$

where  $\psi_s^{s'}$  is the input share from sector  $s'$  for a firm in sector  $s$ .

## 2. Changes in domestic shipping cost

- ▶ Control for changes in access to port, induced by HSR connection

$$\sum_{c' \in \mathbb{C}_z^{port}} \mathbb{I}(HSR_{cc't}),$$

# Controlling for Access to People

## 1. Better access to workers in other prefectures

- ▶ Control for changes in access to workers across occupations, induced by HSR connection

$$\sum_{c' \neq c} \text{CorrL}_{ic'} \mathbb{I}(\text{HSR}_{cc't}),$$

where  $\text{CorrL}_{ic'}$  is the labor correlation between firm  $i$  and city  $c'$  (Ellison et al., 2010).

## 2. Reallocation of high-skilled workers across cities

- ▶ After an opening of a HSR station, core cities have an in-flow of high-skilled workers while others have an out-flow (Yu et al., 2019).
- ▶ Core-city exporters in skill-intensive industries may benefit more than the rest.

$$\text{Skill Intensity} \times \text{Core City} \times 1_{\{\text{HSR}\}}$$

# Quality Measure Construction I

- Modify utility function to:

$$U = \left( \int_{\zeta \in \Omega_n} (\lambda_n(\zeta) q_n(\zeta))^{(\sigma-1)/\sigma} d\zeta \right)^{\sigma/(\sigma-1)},$$

where  $\lambda$  denotes quality.

- Taking logs, we get

$$\ln q_{znkt} + \sigma \ln p_{znkt} = \alpha_{nkt} + \varepsilon_{znkt},$$

- Estimated quality can be calculated using

$$\ln \hat{\lambda}_{znkt} = \frac{\hat{\varepsilon}_{znkt}}{\sigma - 1},$$

where  $\sigma$  is obtained (at HS3 level) from Broda, Greenfield, and Weinstein (2006).

## Quality Measure Construction II

- Calculate a weighted-average quality across different destinations for each firm-product pair

$$\bar{\lambda}_{zkt} = \sum_{n \in \mathcal{N}_t} \theta_{znkt} \ln(\lambda_{znkt})$$

where  $\theta_{znkt}$  denotes the quantity share of firm  $z$  for product  $k$

- Compute long difference in average quality between 2007 and 2013

$$\begin{aligned} \Delta \bar{\lambda}_{zk} &= \bar{\lambda}_{zkt} - \bar{\lambda}_{zkt-1} \\ &= \underbrace{\sum_{n \in \mathcal{N}_t \cap \mathcal{N}_{t-1}} \bar{\theta}_{znk} (\ln \lambda_{znkt} - \ln \lambda_{znkt-1})}_{\text{Intensive-Within}} + \underbrace{\sum_{n \in \mathcal{N}_t \cap \mathcal{N}_{t-1}} (\theta_{znkt} - \theta_{znkt-1}) (\bar{\lambda}_{znk} - \bar{\lambda}_{zk})}_{\text{Intensive-Across}} \\ &\quad + \underbrace{\sum_{n \in \mathcal{N}_t \setminus \mathcal{N}_{t-1}} \theta_{znkt} (\ln \lambda_{znkt} - \bar{\lambda}_{zk})}_{\text{Extensive-Entry}} - \underbrace{\sum_{n \in \mathcal{N}_{t-1} \setminus \mathcal{N}_t} \theta_{znkt-1} (\ln \lambda_{znkt-1} - \bar{\lambda}_{zk})}_{\text{Extensive-Exit}} \end{aligned}$$

- We estimate:

$$\Delta \bar{\lambda}_{zk} = \beta \Delta x_{cs} + \alpha_c + \alpha_s + \Delta \iota_{zk}$$

# Quality Upgrading: Full Decomposition

|                 | Overall             | Intensive           |                     |                  | Extensive         |                   |                    |
|-----------------|---------------------|---------------------|---------------------|------------------|-------------------|-------------------|--------------------|
|                 |                     | Total               | Within              | Across           | Total             | Entry             | Exit               |
| $\Delta x_{CS}$ | 0.058***<br>(0.016) | 0.042***<br>(0.010) | 0.037***<br>(0.009) | 0.005<br>(0.004) | 0.016*<br>(0.008) | 0.009*<br>(0.005) | -0.007*<br>(0.004) |

- Intensive margin dominated by **within-destination** upgrading (0.037), not reallocation across (0.005)
- Extensive: firms **enter** higher-quality markets (0.009) and **exit** lower-quality ones (-0.007)

Back