Division of Labor and Productivity Advantage of Cities: Theory and Evidence from Brazil

ABFER Economic Transformation of Asia

Lin Tian
INSEAD

May 2019

Research question

- Why are firms more productive in larger cities?

Research question

- Why are firms more productive in larger cities?
  - Central question in urban economics
    - Enormous policy implications

Research question

- Why are firms more productive in larger cities?
  - Central question in urban economics
    - Enormous policy implications
  - Quantitatively important

Research question

- Why are firms more productive in larger cities?
  - Central question in urban economics
    - Enormous policy implications
  - Quantitatively important
- Hypothesis inspired by Adam Smith (1776):

Research question

- Why are firms more productive in larger cities?
  - Central question in urban economics
    - Enormous policy implications
  - Quantitatively important
- Hypothesis inspired by Adam Smith (1776):

Larger cities facilitate greater **division of labor within firms**, making firms there more productive

Research question

- Why are firms more productive in larger cities?
  - Central question in urban economics
    - Enormous policy implications
  - Quantitatively important
- Hypothesis inspired by Adam Smith (1776):

  Larger cities facilitate greater **division of labor within firms**, making firms there more productive

- Division of labor: extent of worker specialization within firms What is division of labor?

Research question

- Why are firms more productive in larger cities?
  - Central question in urban economics
    - Enormous policy implications
  - Quantitatively important
- Hypothesis inspired by Adam Smith (1776):

  Larger cities facilitate greater **division of labor within firms**, making firms there more productive

- Division of labor: extent of worker specialization within firms What is division of labor?
- Research question:

  Is **division of labor within firms** *an important mechanism driving* **productivity advantage** *in larger cities?*

Preview of Key Results

I. New stylized fact using Brazilian firm-level data
   - Great division of labor within firms in larger cities

II. Theoretical framework
   - Embed division of labor into a spatial equilibrium model
   - Incorporate different channels that generate the observed correlation between *division of labor* and *city size*

III. Empirical support for the proposed theory
   - Quasi-experiment

IV. Structural analysis: reduced-form evidence from III
   - Quantify importance of *division of labor*

$\Longrightarrow$ **Main result:** Division of labor accounts for 15% of productivity advantage in larger cities.

Related Literature

Preview of Key Results

I. New stylized fact using Brazilian firm-level data
   - Great division of labor within firms in larger cities

II. Theoretical framework
   - Embed division of labor into a spatial equilibrium model
   - Incorporate different channels that generate the observed correlation between *division of labor* and *city size*

III. Empirical support for the proposed theory
   - Quasi-experiment

IV. Structural analysis: reduced-form evidence from III
   - Quantify importance of *division of labor*

$\Longrightarrow$ **Main result:** Division of labor accounts for 15% of productivity advantage in larger cities.

Related Literature

Preview of Key Results

I. New stylized fact using Brazilian firm-level data
   - Great division of labor within firms in larger cities

II. Theoretical framework
   - Embed division of labor into a spatial equilibrium model
   - Incorporate different channels that generate the observed correlation between *division of labor* and *city size*

III. Empirical support for the proposed theory
   - Quasi-experiment

IV. Structural analysis: reduced-form evidence from III
   - Quantify importance of *division of labor*

$\Longrightarrow$ **Main result:** Division of labor accounts for 15% of productivity advantage in larger cities.

Related Literature

Preview of Key Results

I. New stylized fact using Brazilian firm-level data
   - Great division of labor within firms in larger cities

II. Theoretical framework
   - Embed division of labor into a spatial equilibrium model
   - Incorporate different channels that generate the observed correlation between *division of labor* and *city size*

III. Empirical support for the proposed theory
   - Quasi-experiment

IV. Structural analysis: reduced-form evidence from III
   - Quantify importance of *division of labor*

$\Longrightarrow$ **Main result:** Division of labor accounts for 15% of productivity advantage in larger cities.

Related Literature

Preview of Key Results

I. New stylized fact using Brazilian firm-level data
   - Great division of labor within firms in larger cities

II. Theoretical framework
   - Embed division of labor into a spatial equilibrium model
   - Incorporate different channels that generate the observed correlation between *division of labor* and *city size*

III. Empirical support for the proposed theory
   - Quasi-experiment

IV. Structural analysis: reduced-form evidence from III
   - Quantify importance of *division of labor*

$\Longrightarrow$ **Main result:** Division of labor accounts for 15% of productivity advantage in larger cities.

Related Literature

Model to guide measurement

Model to guide measurement

- Production function

$$Q_z^s = \left(\int_0^{c_s} q_z^s(t)^{\frac{\epsilon - 1}{\epsilon}} dt \right)^{\frac{\epsilon}{\epsilon - 1}}$$

  - Tasks are complementary: $0 < \epsilon < 1$
  - c_s: the total length of complementary tasks

Model to guide measurement

- Production function

$$Q_z^s = \left(\int_0^{c_s} q_z^s(t)^{\frac{\epsilon-1}{\epsilon}} dt \right)^{\frac{\epsilon}{\epsilon-1}}$$

  - Tasks are complementary: $0 < \epsilon < 1$
  - c_s: the total length of complementary tasks

- Increasing returns at worker level

$$l(i,t,z) = \int_0^{\infty} \mathbf{1}_z^s(i,t,u)du + z$$

  - Worker time split between production and training
  - Training cost z constant across all $t \in [0, c_s]$

Properties

Occupations and Division of Labor

- Production organization:
  1. A partition $\mathcal{J} = \{J_k\}_{k=1}^{N}$ of the sets of tasks $[0, c_s]$.
  2. A mapping $\mathcal{C}(N) : [0, l] \to \{\mathcal{J}\}$

Occupations and Division of Labor

- Production organization:
  1. A partition $\mathcal{J} = \{J_k\}_{k=1}^{N}$ of the sets of tasks $[0, c_s]$.
  2. A mapping $\mathcal{C}(N) : [0, l] \to \{\mathcal{J}\}$
- Optimal contract:
  - Each worker specializes in one occupation
  - Each occupation has the same number of tasks, $\frac{c_s}{N}$

$\Longrightarrow$ Intuition: All tasks are symmetric and all workers are identical Proof

Occupations and Division of Labor

- Production organization:
  1. A partition $\mathcal{J} = \{J_k\}_{k=1}^{N}$ of the sets of tasks $[0, c_s]$.
  2. A mapping $\mathcal{C}(N) : [0, l] \rightarrow \{\mathcal{J}\}$
- Optimal contract:
  - Each worker specializes in one occupation
  - Each occupation has the same number of tasks, $\frac{c_s}{N}$

$\Longrightarrow$ Intuition: All tasks are symmetric and all workers are identical Proof

- Greater division of labor $\Longrightarrow$ Fewer number of tasks $\Longrightarrow$ More occupations

Occupations and Division of Labor

- Production organization:
  1. A partition $\mathcal{J} = \{J_k\}_{k=1}^{N}$ of the sets of tasks $[0, c_s]$.
  2. A mapping $\mathcal{C}(N) : [0, l] \rightarrow \{\mathcal{J}\}$
- Optimal contract:
  - Each worker specializes in one occupation
  - Each occupation has the same number of tasks, $\frac{c_s}{N}$

$\Longrightarrow$ Intuition: All tasks are symmetric and all workers are identical Proof

- Greater division of labor $\Longrightarrow$ Fewer number of tasks $\Longrightarrow$ More occupations

Stylized Fact

- Data: Brazilian RAIS 2010
  - matched employer-employee records for private establishments in manufacturing sectors
- Division of labor: proxied by number of 6-digit occupation codes within an establishment
  - A range of alternative measures for robustness

Stylized Fact

- Data: Brazilian RAIS 2010
  - matched employer-employee records for private establishments in manufacturing sectors
- Division of labor: proxied by number of 6-digit occupation codes within an establishment
  - A range of alternative measures for robustness

$$\log N_{jms} = \alpha_0 + \alpha_1 \log L_m + \mathbf{X}_{jms} + \varepsilon_{jms}$$

- where $\mathbf{X}_{jms}$:
  - Establishment size
  - Industry FE
  - Occupation categories (3-digit occ codes)
  - State FE
  - Market access
  - Size of local employment in sector s
  - Skill intensity

Greater division of labor within firms in larger cities

Dependent variable	Log no of occupations within an establishment				
	All tradable		Export intensive	Mono-estb firms	Homogeneous
	(1)	(2)	(3)	(4)	(5)
Log (city size)	.0514***	.0206***	.0204***	.0189***	.0166***
	(.0033)	(.0042)	(.0041)	(.0041)	(.0097)
Controls	No	Yes	Yes	Yes	Yes
Obs	266529	266529	97815	250380	34058
R-sq	.126	.854	.848	.862	.881

Standard errors clustered by city in parentheses. Significance levels: * 10%, ** 5%, ***1%. All regressions include state and sector FEs. Establishment-level controls are establishment size, skill intensity, and occupation categories within the firm. City-level controls are share of high-skilled workers, average wage, sector diversity, and the size of local sectoral employment. Occupations are measured by 6-digit Brazilian CBO codes. Sectors are measured by 5-digit Brazilian CNAE codes. Homogeneous sectors include corrugated and solid fiber boxes, white pan bread, carbon black, roasted coffee beans, ready-mixed concrete, oak flooring, motor gasoline, block ice, processed ice, hardwood plywood, and raw cane sugar (Foster, Haltiwanger and Syverson, 2008).

Example | Plots | Specialization index | 4-digit occupation codes | Bins of firm sizes | Population size | Variation of tasks within firms | Division of labor and complexity

Greater division of labor within firms in larger cities

Dependent variable	Log no of occupations within an establishment				
	All tradable		Export intensive	Mono-estb firms	Homogeneous
	(1)	(2)	(3)	(4)	(5)
Log (city size)	.0514*** (.0033)	.0206*** (.0042)	.0204*** (.0041)	.0189*** (.0041)	.0166*** (.0097)
Controls	No	Yes	Yes	Yes	Yes
Obs	266529	266529	97815	250380	34058
R-sq	.126	.854	.848	.862	.881

Standard errors clustered by city in parentheses. Significance levels: * 10%, ** 5%, ***1%. All regressions include state and sector FEs. Establishment-level controls are establishment size, skill intensity, and occupation categories within the firm. City-level controls are share of high-skilled workers, average wage, sector diversity, and the size of local sectoral employment. Occupations are measured by 6-digit Brazilian CBO codes. Sectors are measured by 5-digit Brazilian CNAE codes. Homogeneous sectors include corrugated and solid fiber boxes, white pan bread, carbon black, roasted coffee beans, ready-mixed concrete, oak flooring, motor gasoline, block ice, processed ice, hardwood plywood, and raw cane sugar (Foster, Haltiwanger and Syverson, 2008).

- Both division of labor and production location are endogenous

Example | Plots | Specialization index | 4-digit occupation codes | Bins of firm sizes | Population size
Variation of tasks within firms | Division of labor and complexity

Three Key Ingredients

1. Cities emerge endogenously from homogenous sites
   - Constrained in housing land supply $\leftarrow$ congestion force

Three Key Ingredients

1. Cities emerge endogenously from homogenous sites
    - Constrained in housing land supply $\leftarrow$ congestion force
2. Homogeneous workers, mobile across space
    - Spatial equilibrium $\implies$ higher wages in larger cities

Three Key Ingredients

1. Cities emerge endogenously from homogenous sites
   - Constrained in housing land supply $\leftarrow$ congestion force
2. Homogeneous workers, mobile across space
   - Spatial equilibrium $\implies$ higher wages in larger cities
3. Heterogenous firms: monopolistic competition
   - Single product: freely traded across space
   - Exogenously differ in "complexity" of production technology:

Three Key Ingredients

1. Cities emerge endogenously from homogenous sites
   - Constrained in housing land supply $\leftarrow$ congestion force
2. Homogeneous workers, mobile across space
   - Spatial equilibrium $\implies$ higher wages in larger cities
3. Heterogenous firms: monopolistic competition
   - Single product: freely traded across space
   - Exogenously differ in "complexity" of production technology:
     - ⋆ Across sectors, c_s: *Computer* vs *Shoe*

Three Key Ingredients

1. Cities emerge endogenously from homogenous sites
   - Constrained in housing land supply $\leftarrow$ congestion force
2. Homogeneous workers, mobile across space
   - Spatial equilibrium $\Longrightarrow$ higher wages in larger cities
3. Heterogenous firms: monopolistic competition
   - Single product: freely traded across space
   - Exogenously differ in "complexity" of production technology:
     - ⋆ Across sectors, c_s: *Computer* vs *Shoe*
     - ⋆ Within sector, z: *Nike* vs *local shoe factory*

Two Reduced-Form Assumptions

- A firm chooses division of labor N:
  - increases productivity, raises "costs" (e.g., coordination costs, Becker & Murphy, 1992)

Two Reduced-Form Assumptions

- A firm chooses division of labor N:
  - increases productivity, raises "costs" (e.g., coordination costs, Becker & Murphy, 1992)
- **Assumption 1:** Benefits of division of labor
  - More complex firms and sectors benefit more from greater division of labor
    Microfoundation An example

Two Reduced-Form Assumptions

- A firm chooses division of labor N:
  - increases productivity, raises "costs" (e.g., coordination costs, Becker & Murphy, 1992)
- **Assumption 1:** Benefits of division of labor
  - More complex firms and sectors benefit more from greater division of labor
    Microfoundation An example
- **Assumption 2:** Costs of division of labor
  - Larger cities lower costs of division of labor Microfoundation: infrastructure
    Microfoundation: learning

Two Reduced-Form Assumptions

- A firm chooses division of labor N:
  - increases productivity, raises "costs" (e.g., coordination costs, Becker & Murphy, 1992)
- **Assumption 1:** Benefits of division of labor
  - More complex firms and sectors benefit more from greater division of labor
  Microfoundation An example
- **Assumption 2:** Costs of division of labor
  - Larger cities lower costs of division of labor Microfoundation: infrastructure
  Microfoundation: learning

$\Longrightarrow$ Positive assortative matching between *production complexity* and *city size*

- $L_s^*(z)$ is an increasing function of z

Firm's problem First-order conditions
Spatial Eqm Definition GE quantities Existence and Uniqueness Stability

Equilibrium characteristics

Proposition

In equilibrium, within a sector, firm's division of labor, profit, revenue, and productivity all increase with city size.

Equilibrium characteristics

Proposition

In equilibrium, within a sector, firm's division of labor, profit, revenue, and productivity all increase with city size.

- In equilibrium, more complex firms sort into larger cities
- N is higher in larger cities
  - Selection: more complex firms occupy larger cities, choosing larger N
  - Treatment: larger cities reduce cost of increasing N for all firms
- Firms located in larger cities are bigger (in revenue) and more productive

Descriptive evidence Cross-sector Characteristics

Impact of ICT infrastructure improvement

- Larger cities provide better ICT infrastructure in equilibrium Microfoundation
- Hypothesis: facilitate greater division of labor, e.g., by reducing coordination or information frictions (Bolton and Dewatripont, 1994; Bloom and Garicano, 2008; Garicano and Heaton, 2010)

Proposition

In equilibrium, in response to an ICT improvement,

(i) *all firms increase their division of labor;*

(ii) *the increase is larger for firms in more complex sectors; and*

(iii) *the increase is larger for firms located in bigger cities.*

Impact of ICT infrastructure improvement

- Larger cities provide better ICT infrastructure in equilibrium Microfoundation
- Hypothesis: facilitate greater division of labor, e.g., by reducing coordination or information frictions (Bolton and Dewatripont, 1994; Bloom and Garicano, 2008; Garicano and Heaton, 2010)

Proposition

In equilibrium, in response to an ICT improvement,

(i) *all firms increase their division of labor;*

(ii) *the increase is larger for firms in more complex sectors; and*

(iii) *the increase is larger for firms located in bigger cities.*

- Improvement in ICT infrastructure benefits complex sectors and firms more

$\Longrightarrow$ Larger increase for more complex sectors Definition and firms

Impact of ICT infrastructure improvement

- Larger cities provide better ICT infrastructure in equilibrium Microfoundation
- Hypothesis: facilitate greater division of labor, e.g., by reducing coordination or information frictions (Bolton and Dewatripont, 1994; Bloom and Garicano, 2008; Garicano and Heaton, 2010)

Proposition

In equilibrium, in response to an ICT improvement,

(i) *all firms increase their division of labor;*

(ii) *the increase is larger for firms in more complex sectors; and*

(iii) *the increase is larger for firms located in bigger cities.*

- Improvement in ICT infrastructure benefits complex sectors and firms more

  $\Longrightarrow$ Larger increase for more complex sectors Definition and firms

- More complex firms locate in bigger cities

  $\Longrightarrow$ Larger increase for firms in bigger cities

Quasi-experiment

- Plausibly exogenous variation in ICT infrastructure from a quasi-experiment:
  - Expansion of broadband backbones in Brazil 2012-2014

Quasi-experiment

- Plausibly exogenous variation in ICT infrastructure from a quasi-experiment:
  - Expansion of broadband backbones in Brazil 2012-2014
- Difference-in-differences estimation New backbones

Quasi-experiment

- Plausibly exogenous variation in ICT infrastructure from a quasi-experiment:
  - Expansion of broadband backbones in Brazil 2012-2014
- Difference-in-differences estimation New backbones
- Two objectives:

Quasi-experiment

- Plausibly exogenous variation in ICT infrastructure from a quasi-experiment:
  - Expansion of broadband backbones in Brazil 2012-2014
- Difference-in-differences estimation New backbones
- Two objectives:
  1. Establish existence of one possible explanation (ICT infrastructure)

$$\log N_{jt} = \alpha + \beta Backbone_{jt} + \delta_j + \delta_t + \varepsilon_{jt}$$

Quasi-experiment

- Plausibly exogenous variation in ICT infrastructure from a quasi-experiment:
  - Expansion of broadband backbones in Brazil 2012-2014
- Difference-in-differences estimation New backbones
- Two objectives:
  1. Establish existence of one possible explanation (ICT infrastructure)

$$\log N_{jt} = \alpha + \beta Backbone_{jt} + \delta_j + \delta_t + \varepsilon_{jt}$$

  2. Test complementarity assumptions

$$\log N_{jt} = \alpha + \beta Backbone_{jt} + \gamma Backbone_{jt} \times \log L_{c(j),t_0} + \delta_j + \delta_t + \varepsilon_{jt}$$

$$\log N_{jt} = \alpha + \beta Backbone_{jt} + \omega Backbone_{jt} \times \log c_{s(j),t_0} + \delta_j + \delta_t + \varepsilon_{jt}$$

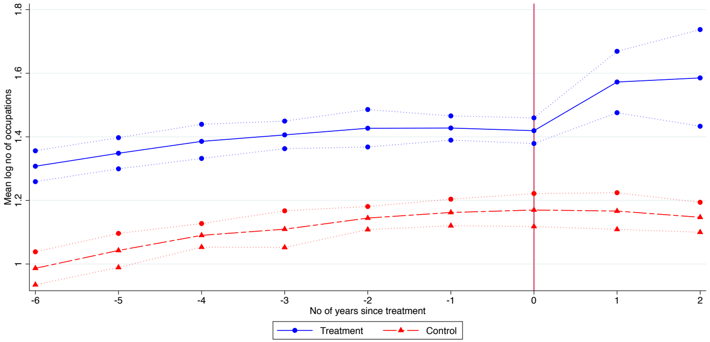
Details of PNBL

Identifying assumptions I

- Identifying assumption: common trend
  - **Parallel trends** in log N_{jt} before the program
  - **No systematically different shocks** after the program

Identifying assumptions I

- Identifying assumption: common trend
  - **Parallel trends** in log N_{jt} before the program
  - **No systematically different shocks** after the program
- No significant difference in pre-trends:



Specialization index

Identifying assumptions II

- Alignment of new backbones pre-determined
- The order in which locations are served approximately geographically determined

Broadband availability increases the division of labor

- Number of occupations increases by 1.3% for firms in treated areas, relative to others
- Increases significantly higher for
  - firms in larger cities
  - firms producing more complex products

Dependent variable	Log (No of occs)			
	(1)	(2)	(3) Interm. inputs	(4) G3 exp share
$Backbone_{jt}$	.0127***	.0015	.0015	.0074**
	(.0028)	(.003)	(.0038)	(.0032)
$Backbone_{jt} \times \log L_{ct_0}$		.0077***		
		(.0008)		
$Backbone_{jt} \times \log c_{st_0}$			.0139***	.004***
			(.0031)	(.0012)
Mean of outcome	1.45	1.45	1.45	1.45
Obs	777096	777096	777096	777096
R-sq	.853	.853	.853	.854

Robust standard errors clustered by municipality in parentheses. Significance levels: * 10%, ** 5%, ***1%. All regressions include a constant term, establishment and year FEs.

Examples | Alternative theories | Other outcome variables | Back to policy evaluation

Robustness I

- Varying distance around the backbone network used to define if an area is served
  - Served if distance $<$ 100km, 200km, 300km, 400km Results
- Adding *lead* variables: $t-1$ and $t-2$
  - insignificant coefficients $\implies$ supporting parallel trends assumption Results
- Firms may reorganize and reallocate resources across establishments in response to the new ICT infrastructure
  - Excluding multiple-establishment firms Results
- Origin and destination locations for the new backbones tend to be larger cities
  - Excluding terminal locations Results
- Locations near submarine cable landing points are typically in or near mega-cities
  - Excluding all establishments located within 100km of the landing points Results
- Areas connected to broadband networks before PNBL may be on a different growth path
  - Excluding firms connected to broadband network before PNBL Results
- Areas that were never treated may be on a different growth path
  - Restricting sample to establishments that are eventually treated Results
- Removing new workers hired after the program Results

Robustness II

- There may be city-specific time trends
  - Adding city-specific linear trends Results
- Results are driven by locations very near or far from the new backbone cables
  - Excluding municipalities that are either very near ($< 10th$ percentile) or very far ($> 90th$ percentile) from the backbone network Results
- Firms may have anticipated the change in ICT infrastructure
  - Excluding data from 2010 and 2011 Results
- Rural areas or mega cities may be on a different growth path compared to urban areas
  - Drop rural areas (density $<$ 400 persons/km^2) Results
  - Drop mega cities (density $>$ 90th percentile) Results
- Export-intensive firms may benefit more as the ICT infrastructure enhances international communication
  - Separate firms into two groups based on sector-level share of exports Results
- Possible spatial correlation biasing standard errors
  - Use Conley SE (Conley, 1999; Conley, 2008) Results
- Possible serial correlation biasing standard errors
  - Non-parametric permutation tests Results
- Combining two interaction terms in a single regression Results

Structural estimations

- Model extensions:
  - Standard RF agglomeration
  - Spatial sorting of firms
  - Imperfect sorting of firms

Structural estimations

- Model extensions:
  - Standard RF agglomeration
  - Spatial sorting of firms
  - Imperfect sorting of firms
- Method of simulated moments
  - Moments: quasi-experiment + cross-sectional data

Structural estimations

- Model extensions:
  - Standard RF agglomeration
  - Spatial sorting of firms
  - Imperfect sorting of firms
- Method of simulated moments
  - Moments: quasi-experiment + cross-sectional data
- Identification: two complementarity parameters (N, z), (N, L)

Structural estimations

- Model extensions:
  - Standard RF agglomeration
  - Spatial sorting of firms
  - Imperfect sorting of firms
- Method of simulated moments
  - Moments: quasi-experiment + cross-sectional data
- Identification: two complementarity parameters (N, z), (N, L)
  - Average city-level increase in division of labor
    - ⋆ Firms in larger cities increase more in response to the ICT shock
    - ⋆ Greater increase if either complementarity is higher

Structural estimations

- Model extensions:
  - Standard RF agglomeration
  - Spatial sorting of firms
  - Imperfect sorting of firms
- Method of simulated moments
  - Moments: quasi-experiment + cross-sectional data
- Identification: two complementarity parameters (N, z), (N, L)
  - Average city-level increase in division of labor
    - Firms in larger cities increase more in response to the ICT shock
    - Greater increase if either complementarity is higher
  - Within-city variation in firm's division of labor
    - Larger variation if complementarity (N, z) is higher

Model fit

- Estimated model fits targeted moments well Results
- Moments not targeted:
  - Sector product complexity lines up well with empirical proxies Some examples Correlation
  - City-size distribution well-approximated by Zipf's Law Results
  - City-level changes in division of labor across all sectors Details

Division of labor and size advantage of cities

- Productivity advantages of larger cities:

$$\log \psi_{js} = \beta_0 + \beta_1 \log L_j + \delta_s + \iota_j \tag{1}$$

- $\hat{\beta}_1 = 0.083$: consistent with 0.02-0.10 estimates in the literature (Rosenthal and Strange, 2004; Melo et al., 2009)

Division of labor and size advantage of cities

- Productivity advantages of larger cities:

$$\log \psi_{js} = \beta_0 + \beta_1 \log L_j + \delta_s + \iota_j \quad (1)$$

- $\hat{\beta}_1 = 0.083$: consistent with 0.02-0.10 estimates in the literature (Rosenthal and Strange, 2004; Melo et al., 2009)

- Shutting down division of labor N Details
  - $\triangle \hat{\beta}_1 = 0.013$
  - Division of labor: 15% of the productivity gains in larger cities

Division of labor and size advantage of cities

- Productivity advantages of larger cities:

$$\log \psi_{js} = \beta_0 + \beta_1 \log L_j + \delta_s + \iota_j \tag{1}$$

- $\hat{\beta}_1 = 0.083$: consistent with 0.02-0.10 estimates in the literature (Rosenthal and Strange, 2004; Melo et al., 2009)

- Shutting down division of labor N Details
  - $\triangle\hat{\beta}_1 = 0.013$
  - Division of labor: 15% of the productivity gains in larger cities
  - 7%–20% due to natural advantages (Ellison and Glaeser, 1999; Roos, 2005) 10% due to knowledge spillover (Serafinelli, 2015)

Division of labor and size advantage of cities

- Productivity advantages of larger cities:

$$\log \psi_{js} = \beta_0 + \beta_1 \log L_j + \delta_s + \iota_j \tag{1}$$

- $\hat{\beta}_1 = 0.083$: consistent with 0.02-0.10 estimates in the literature (Rosenthal and Strange, 2004; Melo et al., 2009)

- Shutting down division of labor N Details
  - $\triangle\hat{\beta}_1 = 0.013$
  - Division of labor: 15% of the productivity gains in larger cities
  - 7%–20% due to natural advantages (Ellison and Glaeser, 1999; Roos, 2005) 10% due to knowledge spillover (Serafinelli, 2015)

- Shutting down systematic choice of L Details
  - $\triangle\hat{\beta}_1 = 0.0064$
  - Sorting of firms: about 50% of the productivity differences through **division of labor**

Alternative approach

Conclusions

- I study how division of labor within firms relates to spatial productivity differences
- New fact: positive correlation between firm's division of labor and city size
- A parsimonious model generating the stylized fact in equilibrium
  - Sorting of firms + direct effect of city size $\Longrightarrow$ spatial distributions of the division of labor and productivity
- Quasi-experiment: strong empirical support for the proposed theory
- Structural analysis: the division of labor accounts for 15% of productivity advantages in larger cities
  - Half due to sorting of firms; half due to direct effect of city size

"**The greatest improvement in the productive powers of labour**, *and the greater part of the skill, dexterity, and judgment with which it is anywhere directed, or applied,* **seem to have been the effects of the division of labour**."

– Adam Smith, the Wealth of Nations (1776)

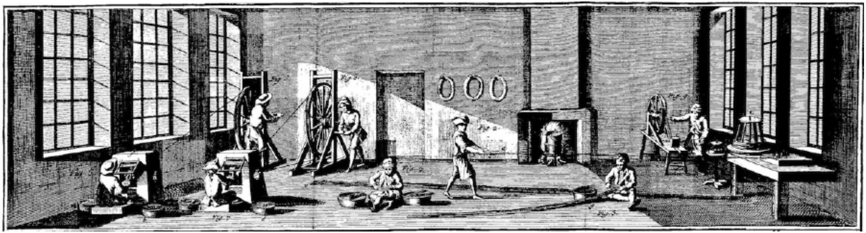


Illustration of the pin factory, Denis Diderot *Encyclopépie* (1772)

APPENDICES

Related Literature

- **Agglomeration economies**: Black and Henderson (1999), Duranton and Puga (2003), Rosenthal and Strange (2004), Melo et al. (2009), Eeckhout and Kircher (2011), Davis and Dingel (2015), Davis and Dingel (2016), De la Roca and Puga (2016), Behrens et al. (2015), Gaubert (2016)
  - I investigate an under-explored mechanism that explains productivity advantage in larger cities.
- **Theories of division of labor**: Becker and Murphy (1992), Costinot (2008), Chaney and Ossa (2013)
  - I develop the first theory of division of labor in a spatial equilibrium setting.
- **Empirical work on division of labor**: Baumgardner (1988), Garicano and Hubbard (2009), Duranton and Jayet (2011)
  - I provide the first economy-wide empirical evidence on the relationship between firm's division of labor and city size.
- **Impact of ICT infrastructure**: Sinai and Waldfogel (2004), Clarke and Wallsten (2006), Commander et al. (2011), Hjort and Poulsen (2016), Fort (2017), Almaida et al. (2017)
  - I study the role of ICT infrastructure in facilitating greater worker specialization within firms.

Back

Preview of results: Theory

- Spatial equilibrium model + internal organizations of firms
- Firms, exogenously heterogeneous in *complexity* of production technology
  - More complex firms **benefit** more from greater division of labor An example

Preview of results: Theory

- Spatial equilibrium model + internal organizations of firms
- Firms, exogenously heterogeneous in *complexity* of production technology
  - More complex firms **benefit** more from greater division of labor An example
  - **Cost** of greater division of labor lower in cities with better infrastructure, e.g. information and communications technology (ICT)

Preview of results: Theory

- Spatial equilibrium model + internal organizations of firms
- Firms, exogenously heterogeneous in *complexity* of production technology
  - More complex firms **benefit** more from greater division of labor An example
  - **Cost** of greater division of labor lower in cities with better infrastructure, e.g. information and communications technology (ICT)
- Larger cities offer better infrastructure

Preview of results: Theory

- Spatial equilibrium model + internal organizations of firms
- Firms, exogenously heterogeneous in *complexity* of production technology
  - More complex firms **benefit** more from greater division of labor An example
  - **Cost** of greater division of labor lower in cities with better infrastructure, e.g. information and communications technology (ICT)
- Larger cities offer better infrastructure

$\Longrightarrow$ **More complex** firms sort into **larger cities** in equilibrium

Preview of results: Theory

- Spatial equilibrium model + internal organizations of firms
- Firms, exogenously heterogeneous in *complexity* of production technology
  - More complex firms **benefit** more from greater division of labor An example
  - **Cost** of greater division of labor lower in cities with better infrastructure, e.g. information and communications technology (ICT)
- Larger cities offer better infrastructure

$\Longrightarrow$ **More complex** firms sort into **larger cities** in equilibrium

- Positive correlation between the division of labor and city size
  (A) More complex firms occupy larger cities, choosing greater division of labor
  (B) Larger cities increase the division of labor for all firms

Preview of results: Theory

- Spatial equilibrium model + internal organizations of firms
- Firms, exogenously heterogeneous in *complexity* of production technology
  - More complex firms **benefit** more from greater division of labor An example
  - **Cost** of greater division of labor lower in cities with better infrastructure, e.g. information and communications technology (ICT)
- Larger cities offer better infrastructure

$\Longrightarrow$ **More complex** firms sort into **larger cities** in equilibrium

- Positive correlation between the division of labor and city size
  (A) More complex firms occupy larger cities, choosing greater division of labor
  (B) Larger cities increase the division of labor for all firms
- In response to an exogenous improvement in infrastructure in a city
  (i) All firms affected increase their division of labor, and
  (ii) The increases are higher for the firms located in larger cities and
  (iii) ... for the firms producing more complex products

Preview of results: Theory

- Spatial equilibrium model + internal organizations of firms
- Firms, exogenously heterogeneous in *complexity* of production technology
  - More complex firms **benefit** more from greater division of labor An example
  - **Cost** of greater division of labor lower in cities with better infrastructure, e.g. information and communications technology (ICT)
- Larger cities offer better infrastructure

$\Longrightarrow$ **More complex** firms sort into **larger cities** in equilibrium

- Positive correlation between the division of labor and city size
  (A) More complex firms occupy larger cities, choosing greater division of labor
  (B) Larger cities increase the division of labor for all firms
- In response to an exogenous improvement in infrastructure in a city
  (i) All firms affected increase their division of labor, and
  (ii) The increases are higher for the firms located in larger cities and
  (iii) ... for the firms producing more complex products
- Firms from multiple sectors can co-exist in the same city
  - Geographic distribution of firms in more complex sectors first-order stochastically dominates that of less complex sectors

Back to Outline Jump to conclusion

Preview of results: Empirics

Examine theoretical predictions using causal evidence + correlation analysis:

(1) Combine a quasi-experiment with the panel of establishment-level data

- Improvement in ICT infrastructure: expansion of broadband backbones PNBL
- Find reduced-form evidence supporting model predictions:
  - (i) Establishments benefiting from faster internet increase their division of labor
  - (ii) The effects are higher for establishments in larger cities
  - (iii) The effects are also higher for establishments producing in more complex sectors

Preview of results: Empirics

Examine theoretical predictions using causal evidence + correlation analysis:

(1) Combine a quasi-experiment with the panel of establishment-level data
- Improvement in ICT infrastructure: expansion of broadband backbones
  PNBL
- Find reduced-form evidence supporting model predictions:
  (i) Establishments benefiting from faster internet increase their division of labor
  (ii) The effects are higher for establishments in larger cities
  (iii) The effects are also higher for establishments producing in more complex sectors

(2) Document additional descriptive evidence consistent with other model predictions
- Within a sector, firms are larger in bigger cities
- Across sectors, firms in more complex sectors are more likely to locate in bigger cities

Back to Outline Jump to conclusion

Preview of results: Structural analysis

(1) The division of labor accounts for **16%** of the overall firm productivity differences across cities

- Comparable to 20% by natural advantages (Ellison and Glaeser, 1999); 10% by knowledge spillover (Serafinelli, 2015)
- About **half** due to firm sorting and the other **half** due to direct city effects

Preview of results: Structural analysis

(1) The division of labor accounts for **16%** of the overall firm productivity differences across cities

- Comparable to 20% by natural advantages (Ellison and Glaeser, 1999); 10% by knowledge spillover (Serafinelli, 2015)
- About **half** due to firm sorting and the other **half** due to direct city effects

(2) Evaluate impacts of the spatially targeted ICT infrastructure program

- Short-term (no reallocation of firms): 3.7 log-point increase in productivity in affected areas
- Long-term (full GE impacts):
  - Local: 9.5 log-point increase in productivity
  - Reduces spatial inequality by 0.7-1.4%
  - Aggregate: 0.39 log-point increase in productivity; 0.38 log-point increase in welfare

Back to Outline Jump to conclusion

Cities and Individuals

Continuum of homogeneous sites: cities emerge endogenously, with population L

- Constrained in housing land supply $\leftarrow$ congestion force

Continuum of homogeneous individuals

- Agents consume both housing, h, and a bundle of freely traded goods, X:

$$U = \left(\frac{X}{\eta}\right)^{\eta} \left(\frac{h}{1-\eta}\right)^{1-\eta}, \text{ where } X = \prod_{s=1}^{S} X_s^{\xi_s}$$

- Within s, agents choose varieties according to a CES aggregator:

$$X_s = \left[\int x_s(z)^{\frac{\sigma_s - 1}{\sigma_s}} dz\right]^{\frac{\sigma_s}{\sigma_s - 1}} \text{ where } \sigma_s > 1, \tag{2}$$

- Given spatial mobility, utility must be equalized across space Derivation

$$w(L) = \bar{w}\left((1-\eta)L\right)^{\frac{1-\eta}{\eta}}, \quad \text{where } \bar{w} = \bar{U}^{1/\eta} P \tag{3}$$

Back to Production Set-up | Back to Production Process | Back to Profit Function | Back to Estimation

Local infrastructure provision

- Following Henderson (1974), assume a class of local land developers
  - fully tax local land owners $\Longrightarrow$ revenue equals profit from housing

$$(1-\eta)Lw(L)$$

  - invest in local infrastructure, $\mathcal{I}$, to attract firms

- Assume free entry and perfect competition, the equilibrium investment in local infrastructure is

$$\mathcal{I}(L) = (1-\eta)Lw(L) = ((1-\eta)L)^{\frac{1}{\eta}}\,\bar{w}. \tag{4}$$

- Better infrastructure, e.g. ICT infrastructure, lowers coordination cost of division of labor
  - Complementarity between $\mathcal{I}$ and N $\Longrightarrow$ complementarity between L and N

Back to production process | Back to model predictions | Back to empirics

Henry George Theorem

- Henry George Theorem (Arnott and Stiglitz, 1979)
  - When the population size of a city is optimal, the provision of public infrastructure is efficient if the expenditure equals aggregate land rent
- Alternative microfoundation (Fujita and Thisse, 2002)
  - Local residents vote to determine the local provision of infrastructure
  - Optimal provision is unanimously selected if the local government implements a housing tax equivalent to housing rent

Back

An example

- Consider an island with two factories: **shoe** and **computer**
- Many identical workers: 300 working days for each worker
  - 1 day to learn any task; 1 day to perform the task (once)

	Shoe	Computer
Number of tasks (Complexity)	10	100
Organization type 1: 1 partition		
Training time per worker	10	100
Production time per worker	290	200
Organization type 2: 2 partitions		
Training time per worker	5	50
Production time per worker	295	250
Gain from specialization	**5 (2%)**	**50 (25%)**

Back to Production Technology

Worker's problem

A worker chooses consumption of tradable good X, housing h and location l to maximize utility:

$$\max_{X,h,L} \left(\frac{X}{\eta}\right)^{\eta} \left(\frac{h}{1-\eta}\right)^{1-\eta}$$

- subject to:

$$PX + p^{h}(L)h = w(L)$$

Equilibrium expenditure on housing:

$$p^{h}(L)h = (1-\eta)w(L)$$

Aggregate expenditure on housing in city L:

$$p^{h}(L)H = (1-\eta)w(L)L$$

Back

Local wages

- Equilibrium housing rents:

$$p^h(L) = \frac{(1-\eta)w(L)L}{H} = (1-\eta)w(L)L$$

- Given spatial mobility, utility must be equalized across space

$$\bar{U} = \left[\frac{w(L)}{P}\right]^{\eta} \left[\frac{L^{-1}}{1-\eta}\right]^{1-\eta}$$

- Re-arrange to get equilibrium local wage equation

$$w(L) = \bar{w}\left[(1-\eta)L\right]^{\frac{1-\eta}{\eta}}$$

where $\bar{w} = \bar{U}^{1/\eta}P$ is a general equilibrium variable.

Back

Production technology I: Complementarity

Leontief production technology (Costinot, 2008):

$$Q_z^s = \int_0^\infty \min_{t \in [0, c_s]} [\mathbf{1}_z^s(t, u)]\, du \quad (5)$$

where

- t: complementary tasks
- c_s: the total length of complementary tasks
  - higher $c_s \implies$ more complex product
- $\mathbf{1}_z^s(t, u)$:

$$\begin{cases} \mathbf{1}_z^s(t, u) = 1 & \text{if } t \in [0, c_s] \text{ is performed on the } u\text{-th unit} \\ \mathbf{1}_z^s(t, u) = 0 & \text{otherwise} \end{cases}$$

Properties Back

Production technology II: Increasing returns

Amount of labor required by worker i performing task t for firm z:

$$l(i,t,z) = \int_0^{\infty} \mathbf{1}_z^s(i,t,u)du + z$$

where

- $\int_0^{\infty} \mathbf{1}_z^s(i,t,u)du$: total number of units in which task t is performed by i
  - $\mathbf{1}_z(n,t,u) = 1$ if n performs task t on unit u;
  - $\mathbf{1}_z(n,t,u) = 0$ otherwise.
- z: training time needed to pursue competency in performing t
  - assume cost constant across all $t \in [0, c_s]$
  - higher-$z \implies$ higher learning costs
- Worker-level training cost to produce at least one unit of output:

$$\int_0^{c_s} z \; dt = c_s z.$$

Back

Optimal contracts and jobs

- All tasks are symmetric and all workers are identical, Proof
  - All workers specialize in one job (i.e. partition of tasks)
  - All jobs include the same number of tasks, $\frac{c_s}{N}$
- For a firm with z, all workers hired have

$$1 - \frac{c_s z}{N}$$

units of time available for production.

- Worker productivity: Detail

$$A(N, z, c_s) = \frac{1}{c_s} - \frac{z}{N}$$

- $A(N, z, c_s)$ is:
  - increasing in N
  - log-supermodular in $(z,\ N)$ and $(c_s,\ N)$.

Back

Production technology I: Complementarity

Leontief production technology (Costinot, 2008):

$$Q_z^s = \int_0^\infty \min_{t \in [0, c_s]} \left[\mathbf{1}_z^s(t, u) \right] du \tag{6}$$

- Every task is essential: if $t \in [0, c_s]$ is not performed on unit u, then u is not produced.
- Tasks are unit-specific: A task performed on unit u cannot be used on $u' \neq u$.

Back

Worker productivity

- In my set-up, all tasks are necessary to produce one unit of output.
- A profit-maximizing firm must allocate the same amount of labor,

$$\frac{l}{N}\left(1 - \frac{zc_s}{N}\right),$$

to each job

- Since each job includes $\frac{c_s}{N}$ tasks, we get:

$$\begin{aligned} A(N, z, c_s)l &= \frac{l}{N}\left(1 - \frac{c_s}{N}\right)\frac{N}{c_s} \\ &= \left(\frac{1}{c_s} - \frac{z}{N}\right)l \end{aligned}$$

$\implies \quad A(N, z, c_s) = \frac{1}{c_s} - \frac{z}{N}$

Back to Optimal Contracts

Alternative interpretation: Coordination costs

- Each worker may make a mistake in her production process with probability:

$$e^{-\frac{1}{\theta}}$$

- Since each task is essential, once a mistake is made, the entire unit is not produced.
- Since all tasks are unit-specific, a unit is only produced if all workers produce without any mistake.
- Probability that a given unit is produced is

$$e^{-\frac{N}{\theta}}$$

- By WLLN, total output is:

$$Q_z^s(N) = e^{-\frac{N}{\theta}} \int_0^\infty \min_{t \in [0, c_s]} \left[\mathbf{1}_z^s(t, u) \right] du$$

Back

The firm's problem

Within s, firm z's problem is to choose its size l, price p, **division of labor** N, and **production location** L to maximize profit:

$$\max_{l,p,N,L} \underbrace{pQ}_{\text{revenue}} - \underbrace{w(L)l}_{\text{costs}} \tag{7}$$

The firm's problem

Within s, firm z's problem is to choose its size l, price p, **division of labor** N, and **production location** L to maximize profit:

$$\max_{l,p,N,L} \underbrace{pQ}_{\text{revenue}} - \underbrace{w(L)l}_{\text{costs}} \tag{7}$$

subject to:

- Production technology:

$$Q_s(z) = A(N, z, c_s)H(N, L)l$$

The firm's problem

Within s, firm z's problem is to choose its size l, price p, **division of labor** N, and **production location** L to maximize profit:

$$\max_{l,p,N,L} \underbrace{pQ}_{\text{revenue}} - \underbrace{w(L)l}_{\text{costs}} \tag{7}$$

subject to:

- Production technology:

$$Q_s(z) = A(N, z, c_s)H(N, L)l$$

- Demand:

$$p_s(z) = Q_s(z)^{-\frac{1}{\sigma_s}} R_s^{\frac{1}{\sigma_s}} P_s^{\frac{\sigma_s - 1}{\sigma_s}}$$

Back

Optimal division of labor

Solving for p and l, profit function for firm z:

$$\max_{N,L} \pi_s(N, L; z) = \max_{N,L} \ \kappa_s \left(\frac{A(N, z, c_s)H(N, L)}{w(L)} \right)^{\sigma_s - 1} \tag{8}$$

Preferences

- FOC wrt to N:

$$\underbrace{\frac{A_N}{A}}_{\text{Marginal benefit}} = - \underbrace{\frac{H_N}{H}}_{\text{Marginal cost}}$$

$$\implies N_s(z, L) \equiv \arg\max_N \pi_s(N, L; z)$$

- FOC wrt L is:

$$\underbrace{\frac{H_L}{H}}_{\text{higher productivity}} = \underbrace{\frac{1-\eta}{\eta}\frac{1}{L}}_{\text{more expensive labor costs}}$$

Back

Definition: Spatial Equilibrium

An equilibrium of a set of cities $\mathcal{L}$ characterized by a city-size distribution $f_L(\cdot)$, a wage schedule $w(L_c)$, a housing price schedule $p_H(L_c)$ and for each sector $s = 1 \ldots S$, a location function $L_s(z)$, and employment function $l_s(z)$, a production function $Q_s(z)$, an ideal price index P_s and a mass of firms M_s such that,

1. workers maximize utility, given $w(L_c), p_H(L_c)$ and P_s;
2. utility is equalized across all inhabited cities;
3. firms maximize profits, given $w(L_c)$ and P_s;
4. landowners maximize profits given $w(L_c)$ and $p_H(L_c)$;
5. factors, goods and housing market clear. In particular, the labor market clears in each city; and
6. firms earn zero profits.

Back

General equilibrium quantities

Remaining $(2S+2)$ unknowns:

- $\bar{w}$: set as the numeraire; R: aggregate revenue
- M_s and P_s, for all $s = 1 \dots S$

S free entry condition for each sector

$$f_{E_s} P = \kappa_s \xi_s R P_s^{\sigma_s - 1} \int_z \left[\frac{A(N, z, c_s) H(N, L)}{w(L)} \right]^{\sigma_s - 1} dF_s(z)$$

where P is the aggregate price index for all tradable sectors. Given Cobb-Douglas preference,

$$P = \prod_{s=1}^{S} \left(\frac{P_s}{\xi_s} \right)^{-\xi_s}.$$

S conditions for aggregate sectoral production

$$1 = \kappa_s \sigma_s M_s P_s^{\sigma_s - 1} \int_z \left[\frac{A(N, z, c_s) H(N, L)}{w(L)} \right]^{\sigma_s - 1} dF_s(z)$$

National labor market clearing condition

$$\bar{L} = \sum_{s=1}^{S} \kappa_s (\sigma_s - 1) M_s \xi_s R P_s^{\sigma_s - 1} \int_z \frac{\left[\frac{A(N,z,c_s) H(N,L)}{w(L)} \right]^{\sigma_s - 1}}{w(L)^{\sigma_s}} dF_s(z)$$

Back

City-size distribution

The city-size distribution is pinned down by the labor market clearing conditions. For all $L > L_0$:

$$\int_{L_0}^{L} r f_L(n) dn = \sum_{s=1}^{S} M_s \int_{z_s^*(L_0)}^{z_s^*(L)} l(z, L_s^*(z)) dF_s(z), \tag{9}$$

where $L_0 \equiv \inf(\mathcal{L})$, the smallest city size in equilibrium.

Differentiating this with respect to L and dividing by L on both sides gives the city size density,

$$f_L(L_c) = \frac{\sum_{s=1}^{S} M_s \mathbf{1}_s(L) l(z_s^*(L)) f_s(z_s^*(L)) \frac{dz_s^*(L)}{dL}}{L} \tag{10}$$

where $\mathbf{1}_s(L) = 1$ if sector s has firms in city L, and 0 otherwise.

Back

Stability Proof I

Fix the set of equilibrium cities as well as the set of firms located in each cities.
Consider a city. In equilibrium, its population is L and it has m firms of draw z
Labor demand for each firm is

$$l = \frac{(\sigma_s - 1)^{\sigma_s}}{\sigma_s^{\sigma_s}} \frac{(A(N_s(z), z, c_s)H(N_s(z), L))^{\sigma_s - 1}}{w(L)^{\sigma_s}} R_s P_s^{\sigma_s - 1}.$$

Applying the local labor market condition,

$$m \frac{(\sigma_s - 1)^{\sigma_s}}{\sigma_s^{\sigma_s}} \frac{(A(N_s(z), z, c_s)H(N_s(z), L))^{\sigma_s - 1}}{w(L)^{\sigma_s}} R_s P_s^{\sigma_s - 1} = L,$$

This pins down the wage $w(L)$ as a function of L.
Recall that worker indirect utility is given by:

$$U(L) \propto w(L)^{\eta} L^{-(1-\eta)}$$

The equilibrium is stable if worker utility increases if a small mass of individuals move away from the city.

Back

Stability Proof II

I prove by contradiction, i.e. suppose $\frac{\partial \log U(L)}{\partial \log L} > 0$ instead.

$$\frac{\partial \log U(L)}{\partial \log L} = \eta \frac{w'(L)L}{w(L)} - (1-\eta) > 0$$

Differentiating local labor market clearing condition with respect with L, we get

$$m \frac{(\sigma_s - 1)^{\sigma_s}}{\sigma_s^{\sigma_s}} \frac{(A(N_s(z), z, c_s))^{\sigma_s - 1}}{w(L)^{\sigma_s}} R_s P_s^{\sigma_s - 1} \left[(\sigma_s - 1) \frac{\partial H}{\partial L} - \sigma_s \frac{w'(L)}{w(L)} \right] = 1.$$

From FOC wrt L in the firm's problem,

$$Lf'(L) \left[\frac{2c_s z}{c_s z f(L) + 1} - \frac{1}{f(L)} \right] = \frac{b(1-\eta)}{\eta}$$

and that by assumption,

$$\frac{w'(L)}{w(L)} L > b \frac{1-\eta}{\eta}$$

Combining, we get

$$Lf'(L) \left[\frac{2c_s z}{c_s z f(L) + 1} - \frac{1}{f(L)} \right] < -b \frac{1-\eta}{\eta} < 0$$

Contradiction.

Back

Microfoundation: complementarity between N and L

- Human capital:
  - intensive b: depth
  - extensive $K = \frac{1}{N}$: breadth
- Learning costs: $\gamma(b, K)$ is convex in (b, K) (Kim, 1989)
- Cost of knowledge acquisition depends on the local aggregate level of b and K (Davis & Dingel, 2014; De la Roca & Puga, 2017)

$$\mathbf{b}(L) = \int_{i \in L} b(i)\, di, \quad \mathbf{K}(L) = \sup\{K(i)\}_{i \in L}, \quad \text{where } i \text{ denotes a worker in city } L.$$

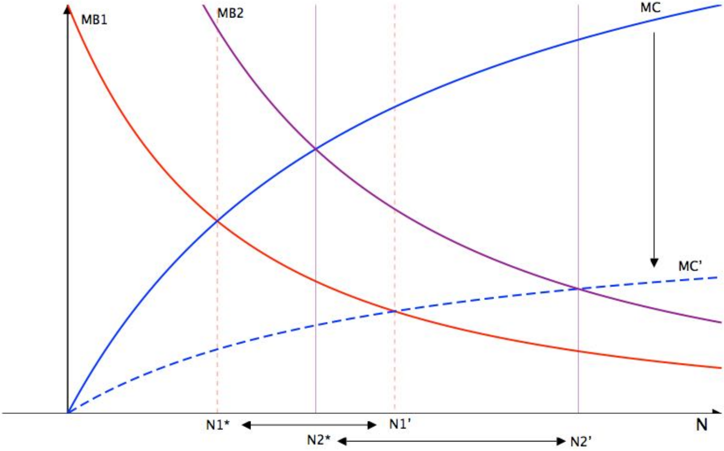
  - $\mathbf{K}(L)$ same everywhere: all tasks are needed to produce a unit of output
  - $\mathbf{b}(L)$ increasing in city size
- Larger cities have a comparative advantage in pursuing intensive human capital

$$\gamma_{bL} < \gamma_{KL} = 0$$

- $\gamma_{bK} > 0 \implies \gamma_{bN} < 0$: high-$N$ firms have a comparative advantage in pursing intensive human capital
- High-N firms benefit more from being in larger cities: cheaper to hire specialists (small K and high b)

Back

Impacts of a shock to ICT infrastructure



Back

Market access

- Market access: potential demand for goods and services produced in city m (see e.g., Fujita, Krugman and Venables, 2001)

$$MarketAccess_m = \sum_{k \neq m} \frac{Y_k}{d_{mk}},$$

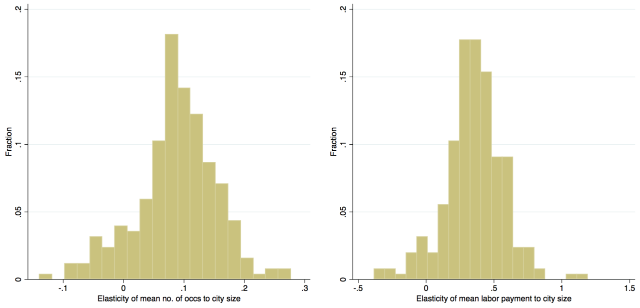
where

  - Y_m: income in city m
  - d_{mk}: straight-line distance between two cities m and k

Back

Within-sector characteristics

- Model: within sector, firm's division of labor and revenue increase with city size
- Division of labor: 91% positive
  - Significantly negative: *growth of grains* and *sawmill*
- Labor payment (proportional to revenue): 94% positive
  - None significantly negative



Back

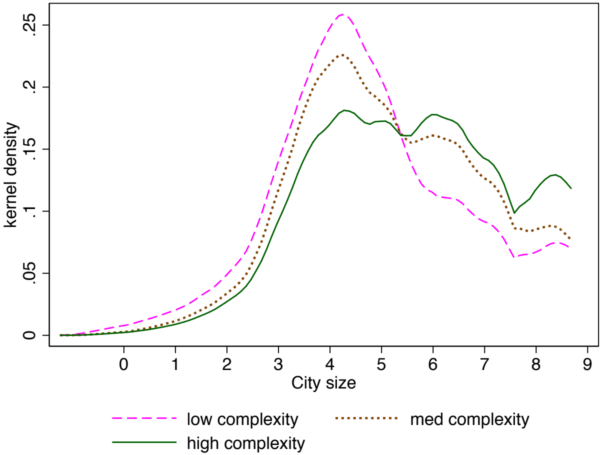
Sectoral-product complexity

- Sector-level complexity c_s: Examples
  - Measure 1: number of intermediate inputs (Dietzenbacher et al., 2005; Levchenko, 2007)
  - Measure 2: export share of goods by the G3 (US, EU and Japan) economies (Hausmann et al., 2006; Wang and Wei, 2010)
    - ⋆ Goods exported by the advanced economies are more complex

Back

Descriptive evidence consistent with model prediction

- Model: geographic distribution of high-c_s sectors FOSD that of low-c_s sectors



Export share Back

Geographic distribution of sectors

- FOSD: larger share of firms in high-c_s sectors in larger cities

$$share_s = \alpha_0 + \alpha_1 \log c_s + \mathbf{X}_s + \varepsilon_s$$

where

- $share_s$: share of establishments in sector s in *larger* cities
- *Larger* cities: the largest cities that host half of the population

Dep var:	Share of establishments in large cities					
	Intermediate inputs			G3 exp share		
	(1)	(2)	(3)	(4)	(5)	(6)
$\log c_s$	.151***	.149***	.127***	.043***	.042***	.029***
	(.0258)	(.0258)	(.0262)	(.009)	(.009)	(.011)
No of Firms	No	Yes	Yes	No	Yes	Yes
Skill Intensity	No	No	Yes	No	No	Yes
Obs	239	239	239	239	239	239
R-sq	.15	.155	.215	.091	.102	.143

Robust standard errors in parentheses. Significance levels: * 10%, ** 5%, ***1%. Sectors are defined at 4-digit Brazilian CNAE level.

Back

Division of labor

- Adam Smith, the Wealth of Nations (1776)

> *"The greatest improvement in the productive powers of labour ... seem(s) to have been the effects of the division of labour... (T)he extent of this division must always be limited by the extent of the market."*

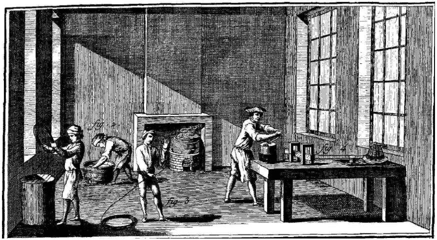


Illustration of the pin factory, Denis Diderot *Encyclopépie* (1772)

Back

Division of labor

- Adam Smith, the Wealth of Nations (1776)

> *"The greatest improvement in the* **productive powers of labour** *...
> seem(s) to have been the effects of* **the division of labour**...
> *(T)he extent of this division must always be limited by
> the extent of the market."*

(1) Greater division of labor within the factory increases productivity

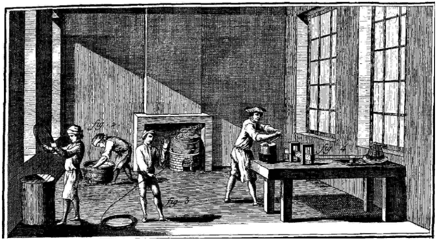


Illustration of the pin factory, Denis Diderot *Encyclopépie* (1772)

Back

Division of labor

- Adam Smith, the Wealth of Nations (1776)

> *"The greatest improvement in the* **productive powers of labour** *... seem(s) to have been the effects of* **the division of labour***...* **(T)he extent of this division** *must always be limited by* **the extent of the market***."*

(1) Greater division of labor within the factory increases productivity

(2) The extent of the local market, i.e. city size, limits the factory's division of labor

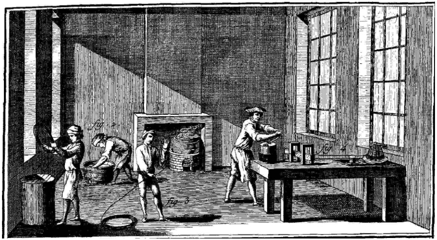
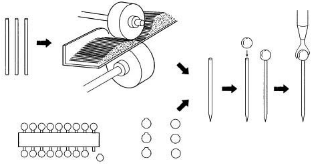


Illustration of the pin factory, Denis Diderot *Encyclopépie* (1772)

Back

What is division of labor?

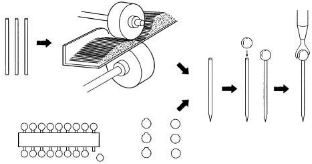
- Division of labor: the extent of worker specialization within a firm.
- Production of any good involves completing a number of different tasks



Tasks involved in making a pin (http://www.madehow.com/Volume-7/Straight-Pin.html)

What is division of labor?

- Division of labor: the extent of worker specialization within a firm.
- Production of any good involves completing a number of different tasks



Tasks involved in making a pin (http://www.madehow.com/Volume-7/Straight-Pin.html)

- A firm organizes its production by partitioning the tasks into smaller groups

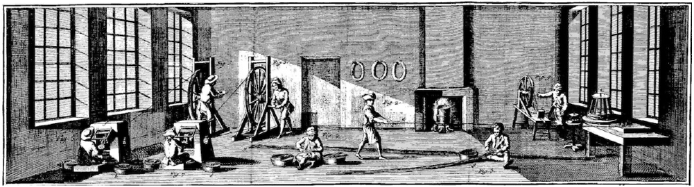
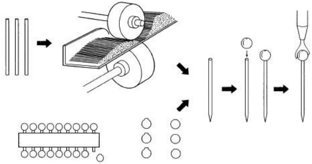


Illustration of the pin factory, Denis Diderot *Encyclopépie* (1772)

What is division of labor?

- Division of labor: the extent of worker specialization within a firm.
- Production of any good involves completing a number of different tasks



Tasks involved in making a pin (http://www.madehow.com/Volume-7/Straight-Pin.html)

- A firm organizes its production by partitioning the tasks into smaller groups

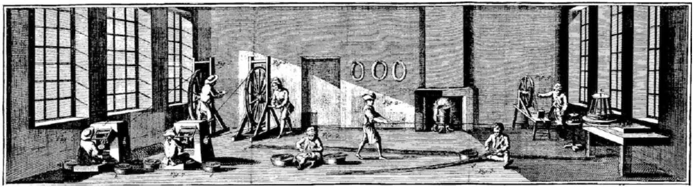


Illustration of the pin factory, Denis Diderot *Encyclopépie* (1772)

- Within a firm: the more partitions $\Longrightarrow$ the greater the division of labor

Back to Intro

Definition: Division of Labor

Measured by the **heterogeneity in occupation codes** within a firm:

1. Identify relevant occupations within a firm
   - Relevant: occupations directly involved in the production process
   - Remove occupations that involve managerial and supervisory tasks primarily

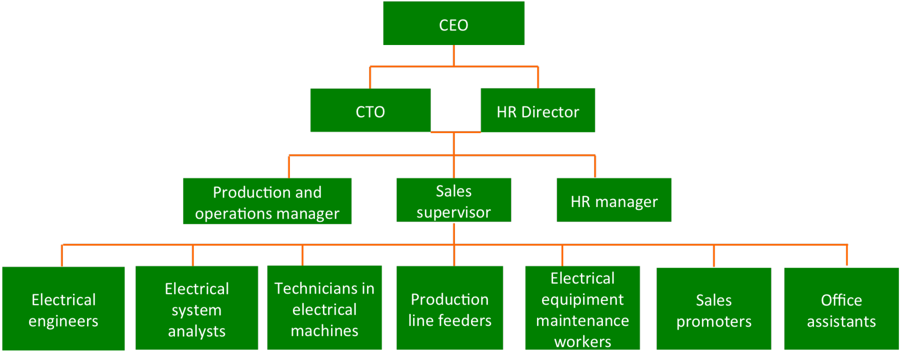
   Algorithm Organization structure

2. Definitions:
   a. Number of occupations within a firm
   b. Specialization index: one minus the Herfindahl index across occupations within a firm (Michaels, 2007; Duranton & Jayet, 2011)

$$N_j = 1 - \sum_{o=1}^{\mathcal{O}} \left(\frac{l_j(o)}{l_j} \right)^2, \text{ where } o \text{ is an occupation.}$$

Back to Definition Back to Empirics

Organization structure



Back to Definition

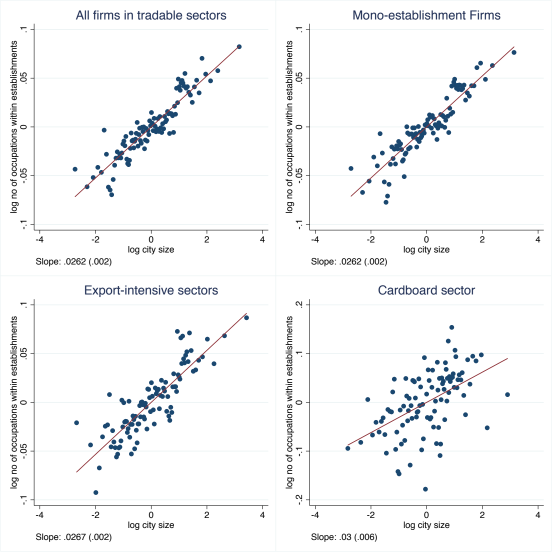
An example: division of labor and city size

No of employees = 15

Firm A (small city)		**Firm B (large city)**	
Machine operator	4	Machine operator	1
		Machine preparer	1
		Machine operation Technician	1
		Machine maintenance Technician	1
Office assistant	5	Office assistant	1
		Administrative assistant	1
		Personal assistant	1
		Office courier	1
		Telephone operator	1
Warehouse clerk	2	Warehouse clerk	1
		Warehouse weighing agent	1
Sales assistant	4	Sales assistant	1
		Specialized sales promoter	1
		Sales technician	1
		Sales and service technician	1
No. of occs (division of labor)	**4**	**No. of occs (division of labor)**	**15**

Back to results

Greater division of labor within firms in larger cities



Back to Introduction | Back to baseline

Greater division of labor within firms in larger cities

Dependent var	Specialization index					
	All tradable			Export intensive	Mono-estb firms	Cardboard
	(1)	(2)	(3)	(4)	(5)	(6)
Log (estb size)	.1721***	.1716***	.1701***	.157***	.1773***	.1505***
	(.001)	(.0011)	(.0011)	(.001)	(.001)	(.0033)
Log (city size)		.0141***	.0147***	.0159***	.015***	.0158***
		(.0009)	(.0011)	(.001)	(.0011)	(.0033)
Log (mkt access)			.0024	-.005	.0012	-.0592***
			(.0046)	(.007)	(.0046)	(.0199)
Log (ind size)			.0018	.0018	.0017	-.0084
			(.0016)	(.0081)	(.0059)	(.003)
Obs	304503	304503	304503	115449	284592	4123
R-sq	.577	.581	.581	.569	.587	.553

Standard errors clustered by city-sector in parentheses. Significance levels: * 10%, ** 5%, ***1%. All regressions include a state FE, a sector FE and the skill intensity within the establishments. Occupations are measured by 6-digit Brazilian CBO codes.

Back to baseline

Greater division of labor within firms in larger cities

Dependent variable	Log no of occupations within an establishment					
	All tradable			Export intensive	Mono-estb firms	Cardboard
	(1)	(2)	(3)	(4)	(5)	(6)
Log (estb size)	.6266***	.6256***	.6256***	.6257***	.6283***	.6312***
	(.0015)	(.0016)	(.0016)	(.0016)	(.0016)	(.0026)
Log (city size)		.0156***	.0156***	.0158***	.0162***	.0169***
		(.0014)	(.0015)	(.0018)	(.0017)	(.0017)
Log (mkt access)			.0002	.0004	.0009	.0313***
			(.0088)	(.0088)	(.009)	(.012)
Log (ind size)			.002*	.0003	.0003	.0024
			(.0011)	(.001)	(.001)	(.002)
Obs	304504	304504	304504	115449	284592	4123
R-sq	.853	.854	.854	.854	.849	.868

Standard errors clustered by city-sector in parentheses. Significance levels: * 10%, ** 5%, ***1%. All regressions include a state FE, a sector FE and the skill intensity within the establishments. Occupations are measured by 4-digit Brazilian CBO codes.

Back to baseline

Greater division of labor within firms in larger cities

Dependent variable	Log no of occupations within an establishment					
	All tradable			Export intensive	Mono-estb firms	Cardboard
	(1)	(2)	(3)	(4)	(5)	(6)
Log (estb size)	.6556***	.6548***	.6551***	.6585***	.6575***	.6691***
	(.0017)	(.0017)	(.0017)	(.0026)	(.0017)	(.0044)
Log (city size)		.0293***	.032***	.033***	.0319***	.0327***
		(.0061)	(.0024)	(.0022)	(.0023)	(.0068)
Log (mkt access)			.0511	.0848***	.0509***	.0121
			(.0089)	(.0124)	(.0092)	(.0398)
Log (ind size)			-.001	-.0003**	.0001	-.0081
			(0.011)	(.0015)	(.0012)	(.0053)
Obs	304504	304504	304504	115449	284592	4123
R-sq	.862	.864	.864	.854	.849	.889

Standard errors clustered by city-sector in parentheses. Significance levels: * 10%, ** 5%, ***1%. All regressions include a state FE, a sector FE and the skill intensity within the establishments. Occupations are measured by 4-digit Brazilian CBO codes.

Back to baseline

Greater division of labor within firms in larger cities

- Separately estimate correlation for each decile

Dependent variable:	Log no of occupations within an establishment		
1st decile	.0005***	6th decile	.0324***
	(.0001)		(.0026)
2nd decile	.0045***	7th decile	.0366***
	(.001)		(.0033)
3rd decile	.0145***	8th decile	.0472***
	(.0014)		(.0039)
4th decile	.0186***	9th decile	.0502***
	(.0018)		(.0046)
5th decile	.0253***	10th decile	.045***
	(.0022)		(.004)

Standard errors clustered by city in parentheses. Significance levels: * 10%, ** 5%, ***1%. All regressions include state and sector FEs, and city-level controls including share of high-skilled workers, average wage, sector diversity, and scale of the sector within cities. Occupations are measured by 6-digit Brazilian CBO codes.

Back to baseline

Division of labor and complexity

- Sector-level complexity: Examples
  - Measure 1: number of intermediate inputs (Dietzenbacher et al., 2005; Levchenko, 2007)
  - Measure 2: export share of goods by the G3 (US, EU and Japan) economies (Hausmann et al., 2006; Wang and Wei, 2010)
    - ⋆ Goods exported by advanced economies are more complex

Division of labor and complexity

- Sector-level complexity: Examples
  - Measure 1: number of intermediate inputs (Dietzenbacher et al., 2005; Levchenko, 2007)
  - Measure 2: export share of goods by the G3 (US, EU and Japan) economies (Hausmann et al., 2006; Wang and Wei, 2010)
    - Goods exported by advanced economies are more complex

$$\log N_{jms} = \alpha_0 + \alpha_1 \log c_s + \mathbf{X}_{jms} + \varepsilon_{jms}$$

- where $\mathbf{X_{jms}}$:
  - Establishment size
  - Size of local employment in sector s
  - Skill intensity
  - Occupation category
  - City FE
  - 2-digit Industry FE

Back

Fact 2: Greater division of labor within firms in more complex sectors

Dependent variable	Log no. of occupations					
	No. of intermediate inputs			G3 export share		
	All tradable		Mono-estb firms	All tradable		Mono-estb firms
	(1)	(2)	(3)	(4)	(5)	(6)
Log (complexity)	.0423***	.0363***	.0372***	5.481***	.5388***	.632***
	(.0145)	(.0043)	(.0043)	(.5432)	(.1756)	(.1376)
Controls	No	Yes	Yes	No	Yes	Yes
Obs	304503	304503	284592	304503	304503	284592
R-sq	.035	.787	.79	.039	.787	.79

Standard errors clustered by sector in parentheses. Significance levels: * 10%, ** 5%, ***1%. All regressions include a city FE and a 2-digit industry FE. Occupations are measured by 6-digit Brazilian CBO codes. Sectors are defined at 4-digit Brazilian CNAE codes.

Results using specialization index
Results using 4-digit occupation codes

Fact 2: Greater division of labor within firms in more complex sectors

Dependent variable	Log no. of occupations					
	No. of intermediate inputs			G3 export share		
	All tradable		Mono-estb firms	All tradable		Mono-estb firms
	(1)	(2)	(3)	(4)	(5)	(6)
Log (complexity)	.0423***	.0363***	.0372***	5.481***	.5388***	.632***
	(.0145)	(.0043)	(.0043)	(.5432)	(.1756)	(.1376)
Controls	No	Yes	Yes	No	Yes	Yes
Obs	304503	304503	284592	304503	304503	284592
R-sq	.035	.787	.79	.039	.787	.79

Standard errors clustered by sector in parentheses. Significance levels: * 10%, ** 5%, ***1%. All regressions include a city FE and a 2-digit industry FE. Occupations are measured by 6-digit Brazilian CBO codes. Sectors are defined at 4-digit Brazilian CNAE codes.

Results using specialization index Results using 4-digit occupation codes

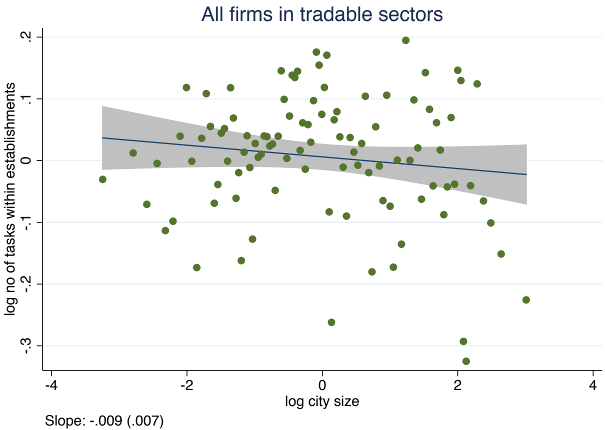
Two stylized facts:

1. Positive correlation between *division of labor* and *city size*
2. Positive correlation between *division of labor* and *complexity*

Back

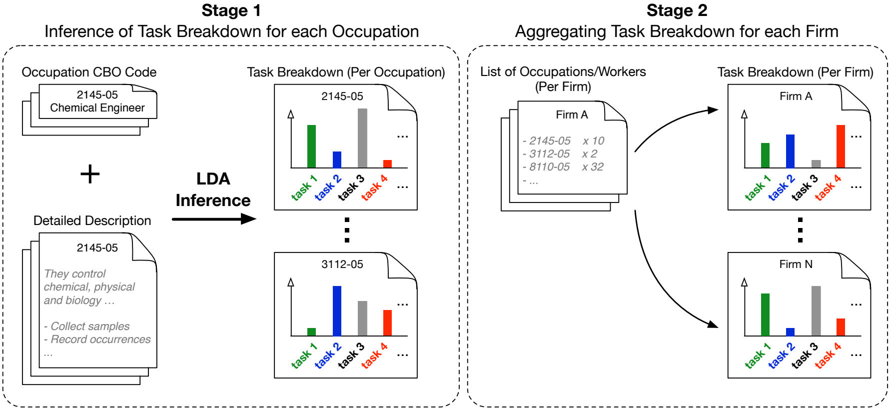
Spatial variation in tasks within firms

- No significant spatial pattern in set of tasks performed within firms
  - Controlling for firm size, sector FE, etc.



Construction of tasks within firms | Back to baseline

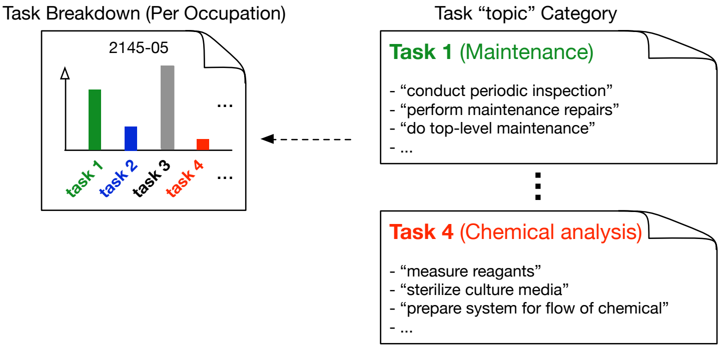
Construction of tasks within firms: overview



Back to plot
Back to main slides

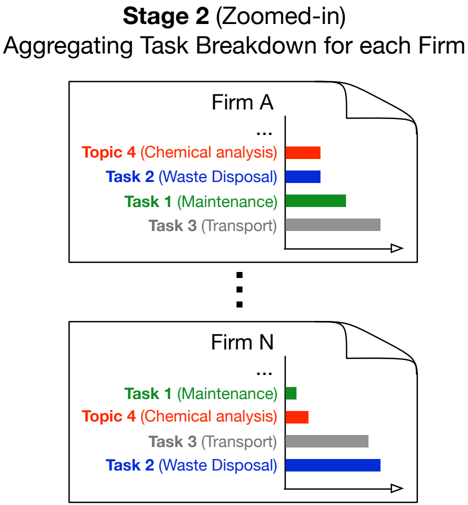
Construction of tasks within firms: Stage 1

Stage 1 (Zoomed-in)
Inference of Task Breakdown for each Occupation



Back to plot | Back to main slides

Construction of tasks within firms: Stage 2



Back to plot Back to main slides

Greater division of labor within firms in more complex sectors

Dependent variable	Log no. of occupations					
	No. of intermediate inputs			G3 export share		
	All tradable		Mono-estb firms	All tradable		Mono-estb firms
	(1)	(2)	(3)	(4)	(5)	(6)
Log (complexity)	.0388***	.0386***	.0382***	2.162***	.311***	.2207***
	(.0046)	(.0022)	(.0022)	(.1996)	(.0678)	(.0634)
Controls	No	Yes	Yes	No	Yes	Yes
Obs	304503	304503	284592	304503	304503	284592
R-sq	.044	.555	.561	.046	.553	.558

Standard errors clustered by city in parentheses. Significance levels: * 10%, ** 5%, ***1%.
All regressions include a city FE. Occupations are measured by 6-digit Brazilian CBO codes.
Sectors are defined at 4-digit Brazilian CNAE codes.

Back to baseline

Greater division of labor within firms in larger cities

Dependent variable	Log no. of occupations					
	No. of intermediate inputs			G3 export share		
	All tradable	Mono-estb firms	Export intensive	All tradable	Mono-estb firms	Export intensive
	(1)	(2)	(3)	(4)	(5)	(6)
Log (complexity)	.1109***	.1072***	.1107***	2.456***	2.613***	3.057***
	(.0041)	(.0045)	(.0043)	(.2518)	(.3366)	(.2246)
Obs	304503	115449	284592	304503	115449	284592
R-sq	.846	.86	.841	.844	.858	.839

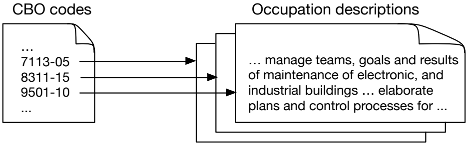
Standard errors clustered by city-sector in parentheses. Significance levels: * 10%, ** 5%, ***1%. All regressions include a state FE, a city FE and the skill intensity within the establishments. Occupations are measured by 6-digit Brazilian CBO codes.

Back to baseline

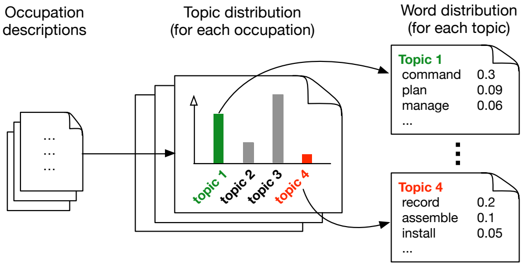
Latent Dirichlet Allocation (LDA)

- LDA: widely-used topic modeling technique in machine learning (Blei et al., 2003)

Stage 1: Preprocessing / Translation



Stage 2: Inference of Topic Distribution by LDA



Back to main slides Back to Definition

Sector-product complexity

- Number of intermediate inputs

Most complex sectors	Least complex sectors
Aircraft and parts	Knitting mills
Electronic computing equipment	Fabricated textile products
Drugs and medicines	Meat Products
Optical and health service supplies	Wood products
Communication equipment	Paper products

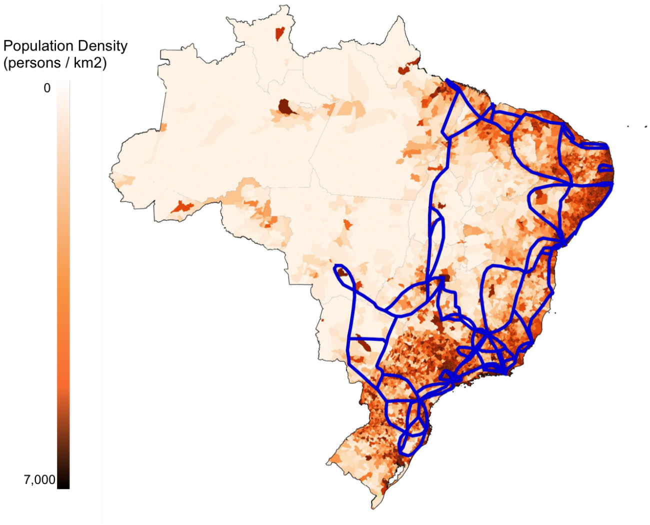
- Export share by G3 economies

Most complex sectors	Least complex sectors
Motor vehicles	Reproduction of recorded media
Refined petroleum products	Fabricated metal products
Motor vehicle engines	Basic metals
Chemicals	Textiles
Drugs and medicines	Wooden containers

Back

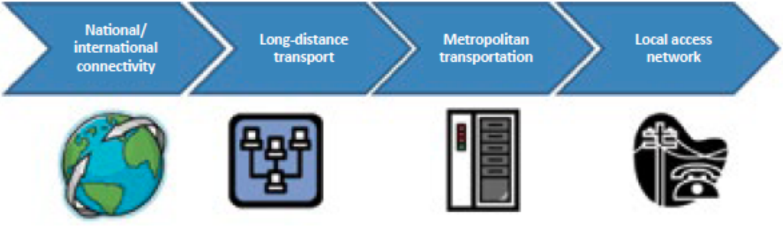
Population density and broadband access in Brazil

- Correlation between pop density and broadband penetration ratio (as of 2010): 0.79



Back

Supply chain of Broadband Internet in Brazil



Source: Anatel

Back

National Broadband Plan: Alignment of backbones

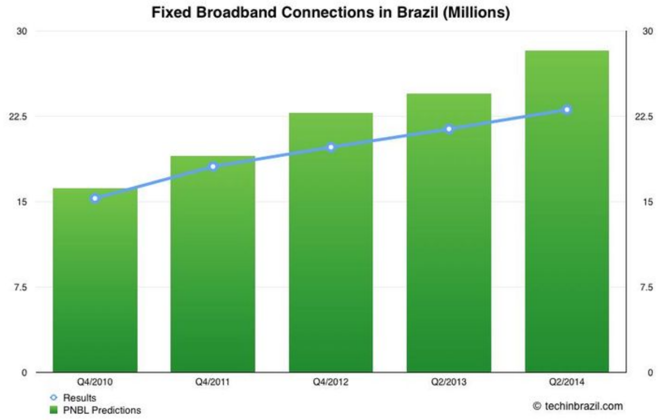
Back to Intro Back to Empirics

National Broadband Plan (PNBL)

- Up to 2010, broadband access in Brazil had been extremely uneven
  - Direct result of a lack of infrastructure for private internet providers in the remote, low-density areas (World Bank, 2011; Knight et al, 2016)
- The federal government launched the largest ICT infrastructure project in Brazil, to ensure that broadband access is available in these poorly served areas:
  - Proposed in 2010, implemented in 2011
  - Budget: $600Mil USD a year for 4 years
  - Target: triple broadband usage
- From Q1/2011 to Q2/2014,
  - Broadband connection coverage increased from 681 to 2930 municipalities
  - Increase amounts to 40% of the total population
- Variation:
  - New broadband trunk infrastructure added during the program: a major initiative ($720Mil USD)
  - Trunk infrastructure alignment follows existing unused government-owned optic fibre cable lines, and existing high-voltage power grid lines

Back

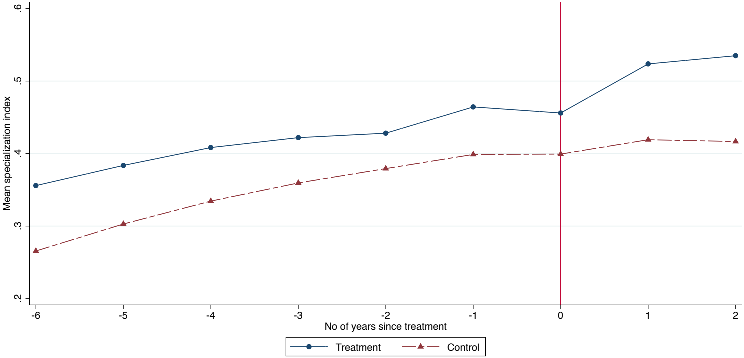
Increase in fixed broadband connections in Brazil



Back

Pre-trend: specialization index

- No significant difference in pre-trends:



Back

Examples of changes in firm's division of labor

Old occupations	Newly added occupations
Firm 1: Manufacturing of medical, dental and optic appliance	
Electrical technician	Electrical technician in assembly and installation of machinery and equipment
	Electrical technician in manufacturing
	Electrical maintenance technician
	Electrical machine maintenance technician
Sales assistant	Specialized sales promoter
	Sales technician
	Sales and service technician
Office assistant	Administrative assistant
	Office courier
	Accounting assistant
	Billing assistant
	Telephone operators
Firm 2: Manufacturing of pharmaceutical products	
Machine operator	Pharmaceutical machine operator
	Machine operation preparer
	Pharmaceutical machine technician
	Boiler operator
Warehouse clerk	Warehouse weighing agent
Office assistant	Administrative assistant
	Personal assistant
	Import and export service assistant

Back to results

Impacts of Broadband Backbone

- Impacts of the improved infrastructure
  - No significant impacts on population, migration of workers and firms
  - Significant increase in the number of firms

Dependent variable	Population (1)	Migration of workers (2)	No. of firms (3)	Relocation of firms (4)
$Backbone_{jt}$	.0258 (.0287)	.0711 (.0566)	.0148*** (.0024)	.04 (.1018)
Obs	5022	3618	5022	1062
R-sq	.987	.716	.986	.225

Robust standard errors clustered by city in parentheses. Significance levels: * 10%, ** 5%, ***1%.
All regressions include a constant term, city and year FEs.

Back to RF results Back to policy evaluation

Alternative theories

- Alternative interpretations:
  - Addition of IT-related jobs
    - Mechanical increase in the number of occupations
    - Robustness test: dropping IT-related occupation codes Results
  - Expansion of firm boundary
    - Adjustment time tends to be longer than internal organizations
    - Existing evidence showing the opposite may be true (Fort, 2017)
    - Robustness tests: dropping new occupation categories Results; testing changes in set of tasks performed Results
  - Skill-biased technological changes:
    - New ICT infrastructure increases skill intensities within firms Results
    - Increase the number of occupations if high-skill occupations more specialized
    - Separately estimate impacts to high and low-skill occupations Results
- Supplementary evidence: change in share of managers / supervisors
  - Managers and supervisors play a coordinating role within an organization
  - Reduction in coordination costs $\implies$ lower share of managers within the establishment (Bloom et al, 2014; McElheran, 2014) Results

Back to results

Robustness: varying the radius I

Dependent variable	Log (No of occs)				Specialization index			
	(1)	(2)	(3) Interm. inputs	(4) G3 exp share	(5)	(6) Interm. inputs	(7) G3 exp share	(8)
				Radius: 100km				
$Backbone_{jt}$	.0072***	-.0009	-.0007	-.0029	.0562***	-.0003	.0447***	.0513***
	(.0025)	(.0027)	(.0037)	(.0029)	(.0158)	(.0107)	(.013)	(.0148)
$Backbone_{jt} \times \log L_{ct_0}$		.0095***				.014***		
		(.0009)				(.0035)		
$Backbone_{jt} \times \log c_{st_0}$			.0125***	.0046***			.0141***	.0059***
			(.0033)	(.0013)			(.0043)	(.0013)
				Radius: 200km				
$Backbone_{jt}$	.0108***	-.0011	.0002	.0053*	.0722***	.0005	.0606***	.0674***
	(.0026)	(.0029)	(.0037)	(.003)	(.0173)	(.0099)	(.0144)	(.0164)
$Backbone_{jt} \times \log L_{ct_0}$		.0084***				.0136***		
		(.0008)				(.0035)		
$Backbone_{jt} \times \log c_{st_0}$			.0131***	.0043***			.0143***	.0061***
			(.0031)	(.0012)			(.0044)	(.0013)
Mean of outcome	1.45	1.45	1.45	1.45	.43	.43	.43	.43
Obs	777096	777096	777096	777096	777096	777096	777096	777096
R-sq	.853	.853	.853	.854	.716	.717	.716	.716

Robust standard errors clustered by municipality in parentheses. Significance levels: * 10%, ** 5%, ***1%. All regressions include a constant term, establishment and year FEs.

Back

Robustness: varying the radius II

Dependent variable	Log (No of occs)				Specialization index			
	(1)	(2)	(3) Interm. inputs	(4) G3 exp share	(5)	(6) Interm. inputs	(7) G3 exp share	(8)
				Radius: 300km				
$Backbone_{jt}$	.0146***	.0033	.0041	.0096***	.0973***	.0267***	.0851***	.0925***
	(.0031)	(.0033)	(.004)	(.0034)	(.0191)	(.0084)	(.0162)	(.0181)
$Backbone_{jt} \times \log L_{ct_0}$		.0076***				.0136***		
		(.0008)				(.0035)		
$Backbone_{jt} \times \log c_{st_0}$			.0131***	.004***			.0151***	.0062***
			(.0031)	(.0012)			(.0043)	(.0013)
				Radius: 400km				
$Backbone_{jt}$	.0098**	-.0037	-.0004	.005	.0869***	.0234**	.0749***	.082***
	(.0047)	(.005)	(.0053)	(.005)	(.0193)	(.0104)	(.0164)	(.0184)
$Backbone_{jt} \times \log L_{ct_0}$		.0081***				.0127***		
		(.0008)				(.0039)		
$Backbone_{jt} \times \log c_{st_0}$			.0128***	.004***			.015***	.0063***
			(.003)	(.0012)			(.0045)	(.0014)
Mean of outcome	1.45	1.45	1.45	1.45	.43	.43	.43	.43
Obs	777096	777096	777096	777096	777096	777096	777096	777096
R-sq	.853	.853	.853	.854	.716	.718	.717	.717

Robust standard errors clustered by municipality in parentheses. Significance levels: * 10%, ** 5%, ***1%. All regressions include a constant term, establishment and year FEs.

Back

Robustness: adding *lead* variables

Dependent variable	Log (No of occs)			Specialization index		
	(1)	(2)	(3)	(4)	(5)	(6)
$Backbone_{jt}$	.0127***	.0122***	.0126***	.0855***	.0843***	.0849***
	(.0047)	(.0045)	(.0047)	(.017)	(.0169)	(.0171)
$Lead_{j,t-1}$		-.0043	-.004		.0098	.0094
		(.0029)	(.0027)		(.04)	(.039)
$Lead_{j,t-2}$			.0021			.0034
			(.0028)			(.0022)
Mean of outcome	1.45	1.45	1.45	.43	.43	.43
Obs	777096	777096	777096	777096	777096	777096
R-sq	.853	.853	.853	.717	.717	.717

Robust standard errors clustered by municipality in parentheses. Significance levels: * 10%, ** 5%, ***1%. All regressions include a constant term, establishment and year FEs.

Back

Robustness: excluding IT related jobs

Dependent variable	Log (No of occs)				Specialization index			
	(1)	(2)	(3) Interm. inputs	(4) G3 exp share	(5)	(6)	(7) Interm. inputs	(8) G3 exp share
$Backbone_{jt}$	.0124***	.0020	.0019	.007*	.086***	.011	.0734***	.0819***
	(.0031)	(.0021)	(.0019)	(.0042)	(.0073)	(.0079)	(.0241)	(.0136)
$Backbone_{jt} \times \log L_{ct_0}$		.0067***				.0126***		
		(.0018)				(.0035)		
$Backbone_{jt} \times \log c_{st_0}$			.0123***	.0027*			.016***	.0038**
			(.0023)	(.0016)			(.0047)	(.0015)
Mean of outcome	1.42	1.42	1.42	1.42	.42	.42	.42	.42
Obs	721629	721629	721629	721629	721629	721629	721629	721629
R-sq	.851	.850	.850	.850	.714	.713	.713	.715

Robust standard errors clustered by municipality in parentheses. Significance levels: * 10%, ** 5%, ***1%. All regressions include a constant term, establishment and year FEs.

Back

Robustness: dropping occ categories that did not exist before

Dependent variable	Log (No of occs)				Specialization index			
	(1)	(2)	(3) Interm. inputs	(4) G3 exp share	(5)	(6)	(7) Interm. inputs	(8) G3 exp share
$Backbone_{jt}$	.0124***	.001	.031	.008*	.076***	.012	.0743***	.0808***
	(.0037)	(.0021)	(.0029)	(.0032)	(.0063)	(.01)	(.0142)	(.013)
$Backbone_{jt} \times \log L_{ct_0}$		.0068***				.0126***		
		(.0018)				(.0035)		
$Backbone_{jt} \times \log c_{st_0}$			.0132***	.003**			.0143***	.0058***
			(.0053)	(.0015)			(.0057)	(.0015)
Mean of outcome	1.42	1.42	1.42	1.42	.42	.42	.42	.42
Obs	777096	777096	777096	777096	777096	777096	777096	777096
R-sq	.851	.850	.850	.851	.715	.715	.714	.715

Robust standard errors clustered by municipality in parentheses. Significance levels: * 10%, ** 5%, ***1%. All regressions include a constant term, establishment and year FEs.

Back

Robustness: dropping new hires after the policy

Dependent variable	Log (No of occs)				Specialization index			
	(1)	(2)	(3) Interm. inputs	(4) G3 exp share	(5)	(6)	(7) Interm. inputs	(8) G3 exp share
$Backbone_{jt}$	.0054***	.0001	.0003	.008	.0411*	.006	.0334**	.0404*
	(.002)	(.0012)	(.0093)	(.0042)	(.025)	(.01)	(.0168)	(.023)
$Backbone_{jt} \times \log L_{ct_0}$		.0038**				.0076**		
		(.0018)				(.0037)		
$Backbone_{jt} \times \log c_{st_0}$			.032***	.001			.0064	.0038***
			(.0013)	(.0015)			(.007)	(.0018)
Mean of outcome	1.44	1.44	1.44	1.44	.42	.42	.42	.42
Obs	777087	777087	777087	777087	777087	777087	777087	777087
R-sq	.749	.749	.749	.749	.615	.615	.615	.615

Robust standard errors clustered by municipality in parentheses. Significance levels: * 10%, ** 5%, ***1%. All regressions include a constant term, establishment and year FEs.

Back

Impacts of ICT on share of managers within establishments

Dependent variable	Share of managers			
	(1)	(2)	(3) Interm. inputs	(4) G3 exp share
$Backbone_{jt}$	-.0114***	-.0087***	-.0072***	-.0085***
	(.0007)	(.0007)	(.001)	(.0008)
$Backbone_{jt} \times \log L_{ct_0}$		-.001***		
		(.0001)		
$Backbone_{jt} \times \log c_{st_0}$			-.0011***	-.0001
			(.0003)	(.0003)
Mean of outcome	.104	.104	.104	.104
Obs	777096	777096	777096	777096
R-sq	.731	.731	.731	.732

Robust standard errors clustered by municipality in parentheses. Significance levels: * 10%, ** 5%, ***1%. All regressions include a constant term, establishment and year FEs.

Back

Impacts of ICT on skill intensity within establishments

Dependent variable	Skill intensity			
	(1)	(2)	(3) Interm. inputs	(4) G3 exp share
$Backbone_{jt}$	.0543***	.0667***	.0389***	.0621***
	(.0009)	(.001)	(.0009)	(.001)
$Backbone_{jt} \times \log L_{ct_0}$		.0081***		
		(.0002)		
$Backbone_{jt} \times \log c_{st_0}$			.0194***	.0061***
			(.0007)	(.0004)
Mean of outcome	.07	.07	.07	.07
Obs	777096	777096	777096	777096
R-sq	.628	.63	.629	.629

Robust standard errors clustered by establishment in parentheses. Significance levels: * 10%, ** 5%, ***1%. All regressions include a constant term, establishment and year FEs.

Back

Robustness: separating low and high-skill occupations

Dependent variable	Log (No of occs)				Specialization index			
	(1)	(2)	(3) Interm. inputs	(4) G3 exp share	(5)	(6) Interm. inputs	(7) G3 exp share	(8)
Low-skill occupations								
$Backbone_{jt}$	.0931***	.0069**	.003	.0016	.063***	.00536	.0621***	.0641***
	(.0027)	(.0029)	(.0035)	(.0031)	(.0109)	(.0077)	(.0114)	(.0106)
$Backbone_{jt} \times \log L_{ct_0}$		.0078***				.0117***		
		(.0007)				(.0023)		
$Backbone_{jt} \times \log c_{st_0}$			.0075***	.0033***			.01***	.0025***
			(.0028)	(.0012)			(.0022)	(.0009)
Mean of outcome	1.12	1.12	1.12	1.12	.56	.56	.56	.56
Obs	777096	777096	777096	777096	777096	777096	777096	777096
R-sq	.835	.835	.835	.835	.618	.618	.618	.618
High-skill occupations								
$Backbone_{jt}$	.0131***	.0012	.0027	.0077**	.0905***	.0052	.0581***	.0478***
	(.0036)	(.0038)	(.0049)	(.0039)	(.0116)	(.0095)	(.0164)	(.0125)
$Backbone_{jt} \times \log L_{ct_0}$		.0093***				.02***		
		(.0009)				(.003)		
$Backbone_{jt} \times \log c_{st_0}$			.0193***	.0042***			.0198***	.0087***
			(.0037)	(.0012)			(.0064)	(.0012)
Mean of outcome	.88	.88	.88	.88	.44	.44	.44	.44
Obs	469224	469224	469224	469224	469224	469224	469224	469224
R-sq	.818	.818	.818	.819	.68	.68	.68	.681

Robust standard errors clustered by municipality in parentheses. Significance levels: * 10%, ** 5%, ***1%. All regressions include a constant term, establishment and year FEs.

Back

Robustness: only mono-establishment firms

Dependent variable	Log (No of occs)				Specialization index			
	(1)	(2)	(3) Interm. inputs	(4) G3 exp share	(5)	(6)	(7) Interm. inputs	(8) G3 exp share
$Backbone_{jt}$	.0128***	.0022	.0021	.0097***	.088***	.0114	.0744***	.0829***
	(.0029)	(.0031)	(.0039)	(.0033)	(.0173)	(.0089)	(.0141)	(.0163)
$Backbone_{jt} \times \log L_{ct_0}$		.0072***				.0147***		
		(.0008)				(.0035)		
$Backbone_{jt} \times \log c_{st_0}$			.0132***	.0021			.017***	.0068***
			(.0032)	(.0013)			(.0047)	(.0015)
Mean of outcome	1.43	1.43	1.43	1.43	.42	.42	.42	.42
Obs	721629	721629	721629	721629	721629	721629	721629	721629
R-sq	.851	.851	.851	.851	.713	.715	.713	.714

Robust standard errors clustered by municipality in parentheses. Significance levels: * 10%, ** 5%, ***1%. All regressions include a constant term, establishment and year FEs.

Back

Robustness: excluding all terminal locations

Dependent variable	Log (No of occs)				Specialization index			
	(1)	(2)	(3) Interm. inputs	(4) G3 exp share	(5)	(6)	(7) Interm. inputs	(8) G3 exp share
$Backbone_{jt}$	.0107***	.0003	-.0003	.0051	.0846***	.0091	.0714***	.0793***
	(.0029)	(.0032)	(.0039)	(.0033)	(.0171)	(.0085)	(.014)	(.0161)
$Backbone_{jt} \times \log L_{ct_0}$		.0067***				.0147***		
		(.0008)				(.0034)		
$Backbone_{jt} \times \log c_{st_0}$			.0137***	.0041***			.0165***	.0069***
			(.0032)	(.0013)			(.0045)	(.0014)
Mean of outcome	1.45	1.45	1.45	1.45	.43	.43	.43	.43
Obs	738702	738702	738702	738702	738702	738702	738702	738702
R-sq	.853	.853	.853	.854	.715	.717	.715	.715

Robust standard errors clustered by municipality in parentheses. Significance levels: * 10%, ** 5%, ***1%. All regressions include a constant term, establishment and year FEs.

Back

Robustness: excluding locations near submarine landing points

Dependent variable	Log (No of occs)				Specialization index			
	(1)	(2)	(3) Interm. inputs	(4) G3 exp share	(5)	(6)	(7) Interm. inputs	(8) G3 exp share
$Backbone_{jt}$	.0104***	.0026	-.0058	.0068*	.062***	.0113	.0548***	.0585***
	(.0032)	(.0041)	(.0042)	(.0037)	(.0072)	(.0074)	(.0065)	(.0069)
$Backbone_{jt} \times \log L_{ct_0}$		.0038***				.011***		
		(.0013)				(.0017)		
$Backbone_{jt} \times \log c_{st_0}$			.0207***	.0026*			.0092***	.0048***
			(.0035)	(.0014)			(.0028)	(.0009)
Mean of outcome	1.41	1.41	1.41	1.41	.43	.43	.43	.43
Obs	606294	606294	606294	606294	606294	606294	606294	606294
R-sq	.85	.85	.85	.85	.719	.72	.719	.72

Robust standard errors clustered by municipality in parentheses. Significance levels: * 10%, ** 5%, ***1%. All regressions include a constant term, establishment and year FEs.

Back

Robustness: excluding locations connected to broadband before PNBL

Dependent variable	Log (No of occs)				Specialization index			
	(1)	(2)	(3) Interm. inputs	(4) G3 exp share	(5)	(6)	(7) Interm. inputs	(8) G3 exp share
$Backbone_{jt}$	.0159***	-.0022	-.0081	.0135***	.0456***	.0299***	.0429***	.0428***
	(.0044)	(.0052)	(.0056)	(.005)	(.0071)	(.008)	(.0066)	(.0068)
$Backbone_{jt} \times \log L_{ct_0}$		.0086***				.0034*		
		(.0014)				(.002)		
$Backbone_{jt} \times \log c_{st_0}$			.0314***	.002			.0036	.0037***
			(.0044)	(.0018)			(.0031)	(.001)
Mean of outcome	1.37	1.37	1.37	1.37	.41	.41	.41	.41
Obs	388539	388539	388539	388539	388539	388539	388539	388539
R-sq	.847	.847	.847	.847	.72	.72	.72	.72

Robust standard errors clustered by municipality in parentheses. Significance levels: * 10%, ** 5%, ***1%. All regressions include a constant term, establishment and year FEs.

Back

Robustness: only treated municipalities

Dependent variable	Log (No of occs)				Specialization index			
	(1)	(2)	(3) Interm. inputs	(4) G3 exp share	(5)	(6)	(7) Interm. inputs	(8) G3 exp share
$Backbone_{jt}$	.0144***	.0031	.0032	.009***	.0923***	.0184**	.0797***	.0873***
	(.0027)	(.003)	(.0037)	(.0031)	(.0181)	(.0083)	(.0152)	(.0172)
$Backbone_{jt} \times \log L_{ct_0}$		.0077***				.0141***		
		(.0008)				(.0033)		
$Backbone_{jt} \times \log c_{st_0}$			.0139***	.004***			.0156***	.0064***
			(.0031)	(.0012)			(.0043)	(.0013)
Mean of outcome	1.46	1.46	1.46	1.46	.43	.43	.43	.43
Obs	764541	764541	764541	764541	764541	764541	764541	764541
R-sq	.854	.854	.854	.854	.717	.719	.717	.717

Robust standard errors clustered by municipality in parentheses. Significance levels: * 10%, ** 5%, ***1%. All regressions include a constant term, establishment and year FEs.

Back

Robustness: adding linear trends

Dependent variable	Log (No of occs)			Specialization index		
	(1)	(2) Interm. inputs	(3) G3 exp share	(4)	(5) Interm. inputs	(6) G3 exp share
$Backbone_{jt}$	.0151***	.0021	.0106***	.0891***	.0789***	.0851***
	(.0026)	(.0036)	(.003)	(.0183)	(.0158)	(.0175)
$Backbone_{jt} \times \log c_{st_0}$		.0159***	.0033***		.0125***	.0051***
		(.0031)	(.0012)		(.0039)	(.0012)
Mean of outcome	1.45	1.45	1.45	.43	.43	.43
Obs	777096	777096	777096	777096	777096	777096
R-sq	.854	.854	.855	.718	.718	.719

Robust standard errors clustered by municipality in parentheses. Significance levels: * 10%, ** 5%, ***1%. All regressions include a constant term, establishment and year FEs.

Back

Robustness: excluding locations with extreme distance

Dependent variable	Log (No of occs)				Specialization index			
	(1)	(2)	(3) Interm. inputs	(4) G3 exp share	(5)	(6)	(7) Interm. inputs	(8) G3 exp share
$Backbone_{jt}$	.0164***	.0067	.0037	.0104**	.1112***	.0068	.0935***	.1043***
	(.0041)	(.0043)	(.0052)	(.0045)	(.0223)	(.0134)	(.0196)	(.0215)
$Backbone_{jt} \times \log L_{ct_0}$		.0066***				.02***		
		(.001)				(.0014)		
$Backbone_{jt} \times \log c_{st_0}$			.0154***	.0043***			.0216***	.0085***
			(.0039)	(.0016)			(.0046)	(.0013)
Mean of outcome	1.51	1.51	1.51	1.51	.42	.42	.42	.42
Obs	450792	450792	450792	450792	450792	450792	450792	450792
R-sq	.862	.862	.862	.862	.716	.719	.716	.716

Robust standard errors clustered by municipality in parentheses. Significance levels: * 10%, ** 5%, ***1%. All regressions include a constant term, establishment and year FEs.

Back

Robustness: excluding data Yr 2010 and Yr 2011

Dependent variable	Log (No of occs)				Specialization index			
	(1)	(2)	(3) Interm. inputs	(4) G3 exp share	(5)	(6)	(7) Interm. inputs	(8) G3 exp share
$Backbone_{jt}$	.0141***	.0016	.0022	.0081**	.0864***	.0092	.0734***	.0812***
	(.0031)	(.0034)	(.0043)	(.0036)	(.0176)	(.0084)	(.0145)	(.0166)
$Backbone_{jt} \times \log L_{ct_0}$		.0086***				.0147***		
		(.001)				(.0034)		
$Backbone_{jt} \times \log c_{st_0}$			.0148***	.0045***			.0161***	.0068***
			(.0036)	(.0014)			(.0046)	(.0014)
Mean of outcome	1.44	1.44	1.44	1.44	.44	.44	.44	.44
Obs	604408	604408	604408	604408	604408	604408	604408	604408
R-sq	.846	.846	.846	.846	.702	.704	.703	.703

Robust standard errors clustered by municipality in parentheses. Significance levels: * 10%, ** 5%, ***1%. All regressions include a constant term, establishment and year FEs.

Back

Robustness: separating export-intensive industries and others

Dependent variable	Log (No of occs)				Specialization index			
	(1)	(2)	(3) Interm. inputs	(4) G3 exp share	(5)	(6) Interm. inputs	(7) G3 exp share	(8)
Export-intensive industries								
$Backbone_{jt}$	.0182***	.0146***	.0064	.023***	.0964***	.0212**	.0756***	.0864***
	(.0047)	(.005)	(.0058)	(.0051)	(.0177)	(.0093)	(.013)	(.0152)
$Backbone_{jt} \times \log L_{ct_0}$		.0027**				.014***		
		(.0013)				(.0028)		
$Backbone_{jt} \times \log c_{st_0}$			.0149***	.0051**			.0263***	.0092***
			(.004)	(.0021)			(.0064)	(.0027)
Mean of outcome	1.59	1.59	1.59	1.59	.4	.4	.4	.4
Obs	307872	307872	307872	307872	307872	307872	307872	307872
R-sq	.857	.857	.857	.857	.72	.722	.721	.721
Others								
$Backbone_{jt}$	.0131***	.0012	.0027	.0077**	.0905***	.0052	.0581***	.0478***
	(.0036)	(.0038)	(.0049)	(.0039)	(.0116)	(.0095)	(.0164)	(.0125)
$Backbone_{jt} \times \log L_{ct_0}$		.0093***				.02***		
		(.0009)				(.003)		
$Backbone_{jt} \times \log c_{st_0}$			.0193***	.0042***			.0198***	.0087***
			(.0037)	(.0012)			(.0064)	(.0012)
Mean of outcome	.88	.88	.88	.88	.34	.34	.34	.34
Obs	469224	469224	469224	469224	469224	469224	469224	469224
R-sq	.818	.818	.818	.819	.68	.68	.68	.681

Robust standard errors clustered by municipality in parentheses. Significance levels: * 10%, ** 5%, ***1%. All regressions include a constant term, establishment and year FEs.

Back

Robustness: excluding rural areas

Dependent variable	Log (No of occs)				Specialization index			
	(1)	(2)	(3) Interm. inputs	(4) G3 exp share	(5)	(6)	(7) Interm. inputs	(8) G3 exp share
$Backbone_{jt}$	.0311***	.0311***	.0199***	.0278***	.1429***	-.0049	.1292***	.1394***
	(.0039)	(.0039)	(.0058)	(.0043)	(.0357)	(.0293)	(.0352)	(.0355)
$Backbone_{jt} \times \log L_{ct_0}$		.0114***				.0221***		
		(.0017)				(.0057)		
$Backbone_{jt} \times \log c_{st_0}$			.0123**	.0022			.015***	.0037***
			(.0048)	(.0016)			(.0042)	(.0011)
Mean of outcome	1.57	1.57	1.57	1.57	.46	.46	.46	.46
Obs	372726	372726	372726	372726	372726	372726	372726	372726
R-sq	.857	.857	.857	.857	.709	.711	.709	.71

Robust standard errors clustered by municipality in parentheses. Significance levels: * 10%, ** 5%, ***1%. All regressions include a constant term, establishment and year FEs.

Back

Robustness: excluding mega cities

Dependent variable	Log (No of occs)				Specialization index			
	(1)	(2)	(3) Interm. inputs	(4) G3 exp share	(5)	(6) Interm. inputs	(7) G3 exp share	(8)
$Backbone_{jt}$	.0124***	-.004	-.0019	.0086***	.0648***	.0226***	.0575***	.0614***
	(.0028)	(.0033)	(.0039)	(.0033)	(.0013)	(.0024)	(.0016)	(.0014)
$Backbone_{jt} \times \log L_{ct_0}$		.009***				.0086***		
		(.0011)				(.0004)		
$Backbone_{jt} \times \log c_{st_0}$			.0182***	.0028**			.0093***	.0046***
			(.0033)	(.0013)			(.0013)	(.0005)
Mean of outcome	1.41	1.41	1.41	1.41	.44	.44	.44	.44
Obs	705861	705861	705861	705861	705861	705861	705861	705861
R-sq	.85	.85	.85	.85	.72	.721	.72	.721

Robust standard errors clustered by municipality in parentheses. Significance levels: * 10%, ** 5%, ***1%. All regressions include a constant term, establishment and year FEs.

Back

Robustness: accounting for spatial correlation

Dependent variable	Log (No of occs)				Specialization index			
	(1)	(2)	(3) Interm. inputs	(4) G3 exp share	(5)	(6) Interm. inputs	(7) G3 exp share	(8)
Backbone$_{jt}$	.0127***	.0015	.0015	.0074	.0855***	.0116	.0728***	.0805**
	(.0049)	(.006)	(.006)	(.005)	(.036)	(.0102)	(.034)	(.039)
Backbone$_{jt}$ $\times \log L_{ct_0}$		.0077***				.0141***		
		(.0016)				(.0065)		
Backbone$_{jt}$ $\times \log c_{st_0}$			.0139***	.004*			.0156*	.0064***
			(.0047)	(.0021)			(.0084)	(.0022)
Mean of outcome	1.45	1.45	1.45	1.45	.43	.43	.43	.43
Obs	777096	777096	777096	777096	777096	777096	777096	777096
R-sq	.853	.853	.853	.854	.717	.718	.717	.717

Conley standard errors in parentheses. Significance levels: * 10%, ** 5%, ***1%. All regressions include a constant term, establishment and year FEs.

Back

Robustness: combining interaction terms in a single equation

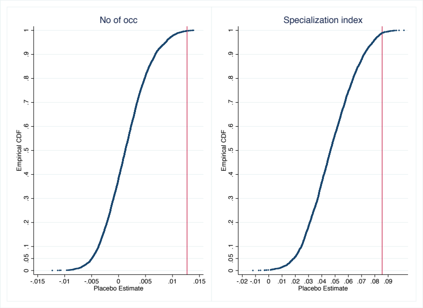
Dependent variable	Log (No of occs)		Specialization index	
	(1) Interm. inputs	(2) G3 exp share	(3) Interm. inputs	(4) G3 exp share
$Backbone_{jt}$	-.0017	-.001	.0092	.0112
	(.0041)	(.0033)	(.0081)	(.0084)
$Backbone_{jt} \times \log L_{ct_0}$	.0089***	.0075***	.0138***	.0138***
	(.0008)	(.0008)	(.0034)	(.0034)
$Backbone_{jt} \times \log c_{st_0}$	.021***	.002*	.005**	.003***
	(.0032)	(.001)	(.002)	(.001)
Mean of outcome	1.45	1.45	.43	.43
Obs	777096	777096	777096	777096
R-sq	.854	.854	.718	.719

Robust standard errors clustered by municipality in parentheses. Significance levels: * 10%, ** 5%, ***1%. All regressions include a constant term, establishment and year FEs.

Back

Non-parametric permutation test

- Serial correlation in D-in-D analysis can bias standard errors, leading to over-rejection of null hypothesis of no effect (Bertrand et al, 2002)
- Non-parametric permutation test for $\beta = 0$ (e.g. Chetty et al, 2009)
  - Maintain the alignment of new backbones
  - Sample from the set of true new backbone implementation years
  - Assign a randomly chose "fake" treatment time to each location



- P-values from 4000 permutation tests, 0.0022 (log no of occs) and 0.011 (specialization index), are very similar to the estimates from t-tests

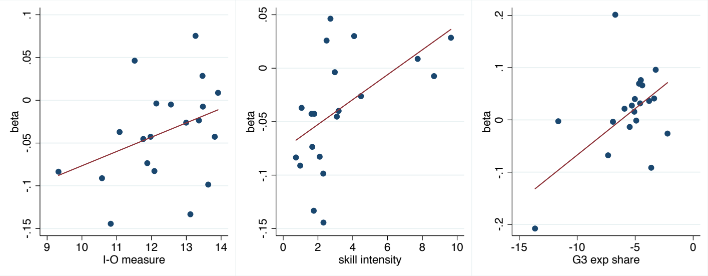
Back

Sectoral variation in the impacts of ICT improvement

- Model prediction: the impacts of new ICT infrastructure are higher for high-c sectors

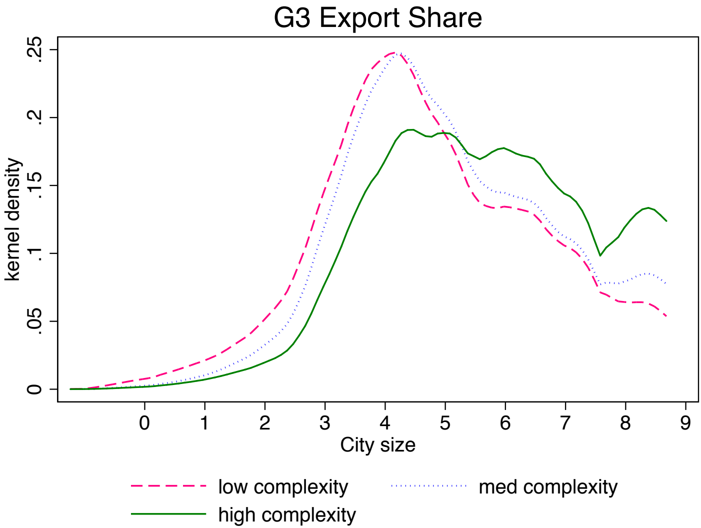
$$\log N_{jt} = \alpha + \beta_{s(j,t)} NewBackbone_{m(j)t} + \gamma_{s(j,t)} NewBackbone_{m(j)t} \times Connected_{m(j),t-1} + \zeta_{s(j,t)} Connected_{m(j),t-1} + \delta_j + \delta_t + \varepsilon_{jt}$$

Sector-complexity measure	Rank correlation with $\hat{\beta}_s$
No of intermediate inputs	.375
Skill intensity	.638
G3 export share	.64



Back

Geographic distribution of sectors: G3 export share



Back

Sector summary statistics

Sector	Wage bill*	Employment*	No. of occs*	N
Agriculture, and mining	10.94	2.05	1.42	6792
Food products, beverages and tobacco products	11	2.37	1.54	29281
Textiles	11.15	2.4	1.55	7162
Wearing apparel	10.75	2.18	1.33	33888
Leather goods and footwear, leather tanning	11.15	2.52	1.54	8255
Wood, except furniture	10.81	2.13	1.37	10324
Pulp, paper and paper products	11.45	2.58	1.78	3326
Publishing, printing and reproduction of recorded media	10.65	1.77	1.34	8311
Chemicals and chemical products	11.61	2.49	1.83	7364
Pharmaceutical products	12.18	2.88	2.08	670
Rubber and plastic products	11.5	2.58	1.73	11475
Glass, ceramic, brick and cement products	10.95	2.28	1.42	19854
Basic metals	11.6	2.5	1.81	3315
Fabricated metal products, except machinery	10.99	2.04	1.4	25693
Computer and electronic products	11.67	2.47	1.82	2757
Electrical machines	11.68	2.58	1.86	3586
Other equipments and machines	11.58	2.34	1.81	11238
Automotive vehicles	11.6	2.54	1.81	4563
Other transport equipment	11.66	2.57	1.86	859
Furniture	10.83	2.06	1.32	13444
Miscellaneous products, other mfg activities	10.89	2.1	1.39	7205

*: Average values in natural logs.

Back

Robustness: sorting

Productivity of firm j in sector s with complexity z: Moments

$$\log \psi_{js} \equiv (\log z)\,(1 + \log N)^{c_s} - \log N\,(1 + \log L)^{-\theta_s} + \alpha_s \log L + \log z(1 + \log L)^{\iota_s} + \varepsilon_{jL}$$

- Reduced-form sorting of firms: $\log z(1 + \log L)^{\iota_s}$
  - More complex firms may sort into larger cities for other reasons, beyond division of labor
  - ι_s: strength of reduced-form sorting effects
  - $\iota_s = 0$: more complex firms sort into larger cities only for division of labor

Back

First stage estimation

Direct calibration $2S+1$ parameters:

- ξ_s: Cobb-Douglas share of each s
  - Sectoral share of value-added
- σ_s: elasticity of substitution within s
  - Sectoral revenue to cost margin
- η: exponent in worker's utility function
  - Elasticity of wage to city size

Model specifications and assumptions | Moments and results
Back to Estimation

Structural results: moments targeted

- 21 moments targeted:
  - Average no. of occupations for each quartile of city size Results
  - Variance of the no. of occupations within each quartile of city size Results
  - Average firm size (in labor payment) for each quartile of city size Results
  - Firm-size distribution in labor payment: 25, 50, 75 and 90th percentiles Results
  - Share of employment in a given sector across four city bins (defined by cities with $<$ 25%, 50%, 75% of the overall workforce) Results

Back

Simulation procedure

Simulate the economy and estimate the parameters for each sector separately

(1) Start with a set of initial values for χ^s, draw 100,000 complexities and 100,000 $\times$ 400 firm-city size specific shocks.

(2) Compute optimal choice of N^*, given θ_s and c_s.

(3) Compute optimal choice of L^*, using:

$$\log L^* = \underset{L \in \mathcal{L}}{\arg\max} \ \log z \left(1 + \log N^*\right)^{c_s} - \frac{\log N^*}{\left(1 + \log \tilde{L}\right)^{\theta_s}} + \left(\alpha_s - \frac{1-\eta}{\eta}\right) \log \tilde{L} + \varepsilon_{j,L}$$

(4) Compute the 21 moments

(5) Find the set of χ_s that minimizes the distance between the simulated moments and the targeted moments using the particle swarm optimization algorithm.

$$\hat{\chi}_s = \arg\min \left(m_{s,data} - m_{s,sim}(\chi_s)\right)' J \left(m_{s,data} - m_{s,sim}(\chi_s)\right)$$

where J_s is the diagonal of the variance-covariance matrix of $m_{s,data}$ (Altonji and Segal, 1996)

Back

Geographic distribution of firms

- Identifies ν_z
- Density of firms located in different cities
- Conditional on sorting, distribution of z determines the density
- 4 moments:
  - Share of employment in a given sector across 4 city bins
  - City bins: cities with $< 25\%$, 50%, 75% of the overall workforce

Back

Increase in average firm's division of labor across city sizes

- Identifies $\frac{\theta}{1-c}$
  - As L increases, N increases
  - N increases more if c higher and/or θ higher
- 4 moments:
  - Average number of occupations (proxy for division of labor) for each quartile of city size

Back

Increase in average firm size across city sizes

- Identifies RF agglomeration externalities, α from $c\theta$
- When N optimally chosen, $c\theta$ complementarity between z and L
- As L increases, r increases
  - α: direct effect
  - $c\theta$: direct effect + interaction between z and L due to sorting
- 4 moments:
  - Average firm labor payment (proportional to revenue) for each quartile of city size

Back

Firm-size distribution

- Identifies distribution of the error structure ν_L and ν_z
  - Without errors, distribution of z + sorting $\implies$ firm-size distribution
  - Errors dampen sorting $\implies$ change firm-size distribution
- Intuitively,
  - ν_z: (direct) distribution of z + (indirect) matching function
  - ν_L: (indirect) matching function
- 5 moments:
  - Share of firms in 5 bins of normalized labor payment
  - 5 bins: 25, 50, 75 and 90th percentiles (Eaton et al., 2011)

Back

Change in division of labor across city sizes

- Two-step process:
  1. Estimate $\triangle H(N, L)$ to match average treatment effect: 1.27% increase
  2. Calculate $\triangle N^*$ based on simulated distribution of firms
- N increases more for higher z
  - High-z firms in larger cities
  - Sorting stronger if θ or c higher
- 4 moments:
  - Average change in division of labor for each quartile of city size

Back

Within-city variation in division of labor

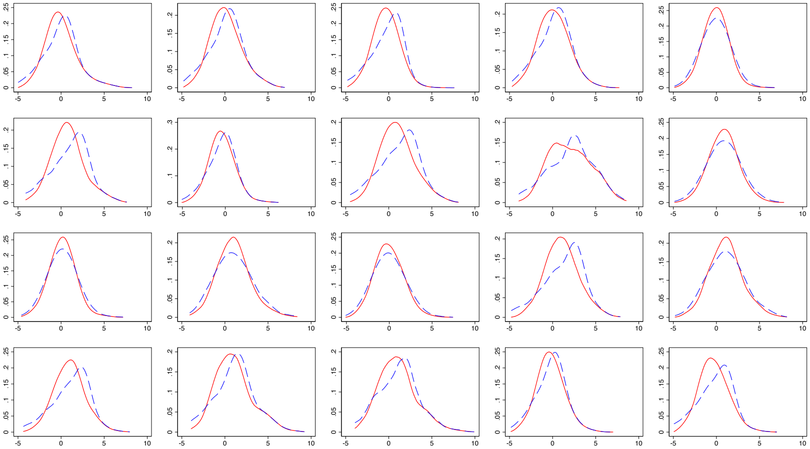
- Variation in firm's division of labor, N, given a city size, L, identifies c from θ
- Within same city, same effect on N through L
  - Differences in N driven by variation in z in that city
- Intuitively, all else equal, variation in N within a city larger if c higher
- 4 moments:
  - Variance of the no. of occupations within each quartile of city size

Back

Structural results

Distribution of firm labor payment

- Actual *vs.* simulated

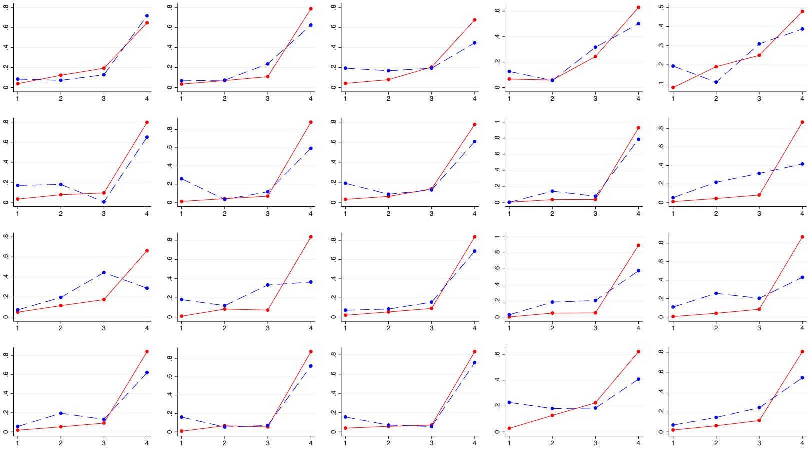


Back

Structural results

Employment share by city bin

- Actual *vs.* simulated

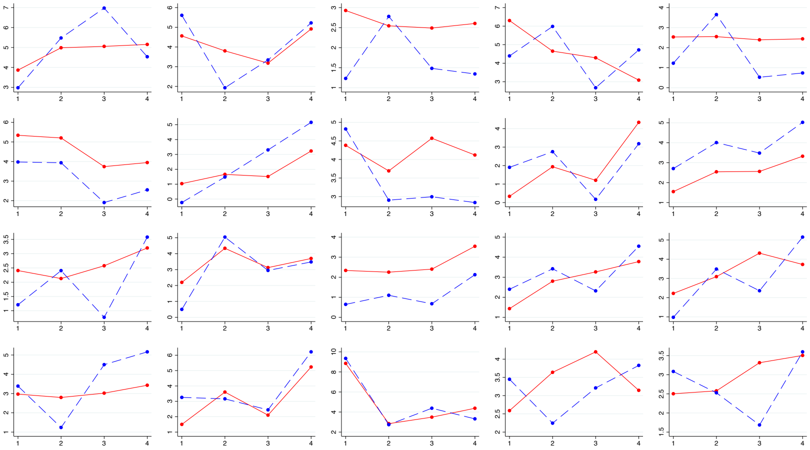


Back

Structural results

Average labor payment by city bin

- Actual *vs.* simulated

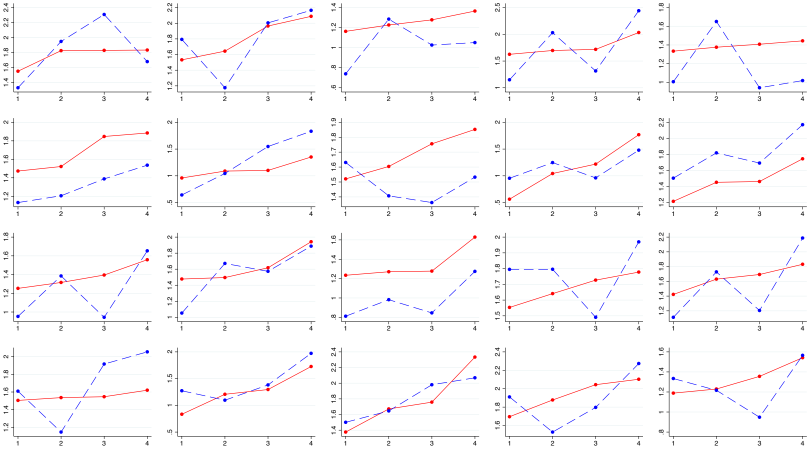


Back

Structural results

Average no of occupations by city bin

- Actual *vs.* simulated

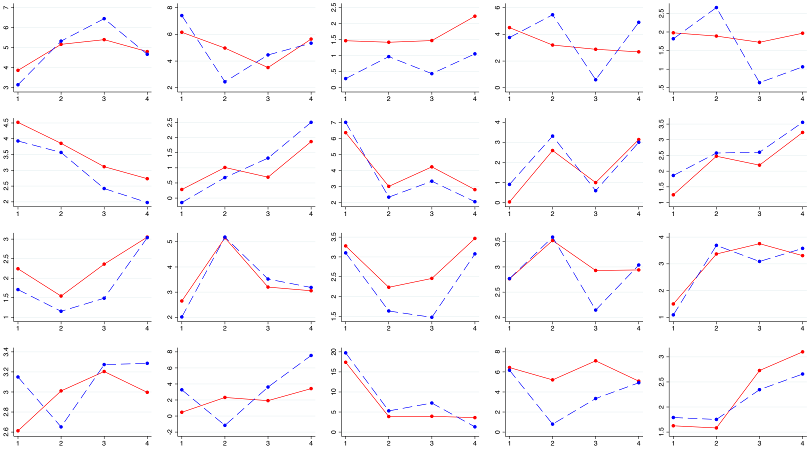


Back

Structural results

Variance of the no of occupations within city bins

- Actual *vs.* simulated



Back

Complexity estimates across sectors: examples

Sectors with the highest c_s	**Sectors with the lowest c_s**
Mfg of computer and electronic products	Mfg of pulp, paper and paper products
Mfg of automotive vehicles	Mfg of wood products, except furniture
Mfg of other transport equipment	Mfg of leather goods and footwear
Mfg of pharmaceutical products	Mfg of basic metals
Mfg of misc. products, other mfg activities	Mfg of glass, ceramic, brick and cement
Mfg of electrical machines	Mfg of wearing apparel
Mfg of other equipment and machines	Mfg of F&B, and tobacco
Mfg of furniture	Mfg of metal products
Mfg of rubber and plastic products	Mfg of textiles
Mfg of chemicals and chemical products	Agriculture & mining

Back to Structural Back to Empirics

Complexity estimates across sectors

- Sector-level complexity c_s is estimated separately for each sector
  - Relative values of c_s across sectors should be consistent with the measured sector complexities
- Rank correlations with the two data proxies Plot:
  - No of intermediate inputs: 0.68
  - Export share of goods by the G3 economies: 0.63

Sectors with the highest c_s	**Sectors with the lowest c_s**
Mfg of computer and electronic products	Mfg of pulp, paper and paper products
Mfg of automotive vehicles	Mfg of wood products, except furniture
Mfg of other transport equipment	Mfg of leather goods and footwear
Mfg of pharmaceutical products	Mfg of basic metals
Mfg of misc. products, other mfg activities	Mfg of glass, ceramic, brick and cement

Back to main slides

Moments not targeted

City-size distribution

- Zipf's law: well-established empirical fact:

$$\log \mathit{Rank} = \mathit{const} - \log \mathit{Size} + \epsilon$$

- No restriction on city-size distribution in the estimation

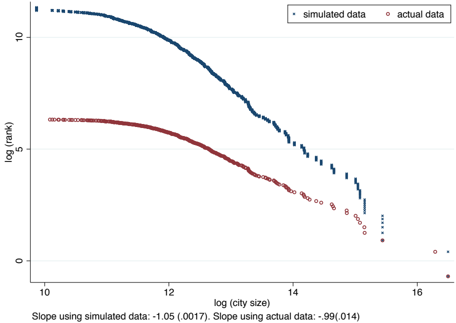
Moments not targeted

City-size distribution

- Zipf's law: well-established empirical fact:

$$\log \mathit{Rank} = \mathit{const} - \log \mathit{Size} + \epsilon$$

- No restriction on city-size distribution in the estimation



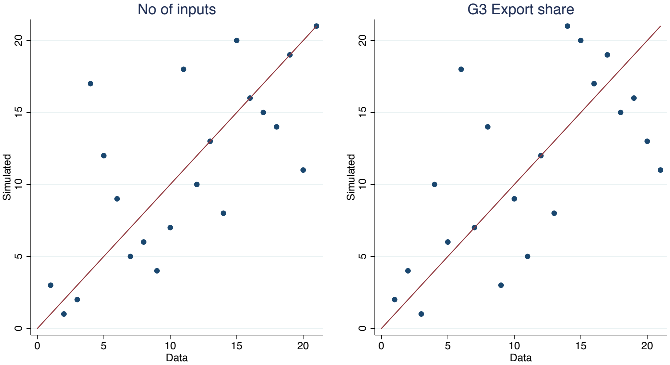
- R-squared: .89 (actual); .81 (simulated)
- Slope: -.99 (actual); -1.05 (simulated)

Back to main slides

Moments not targeted

sector-level complexity

- c_s estimated for each sector separately
- Rank correlations between estimated c_s and the two data proxies
  - No of intermediate inputs: 0.68
  - Export share of goods by the G3 economies: 0.63



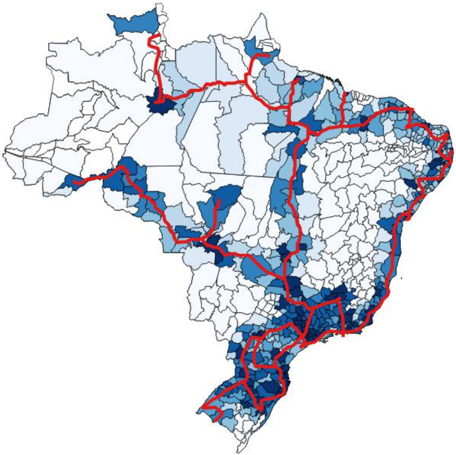
Back

Assessing structural estimates

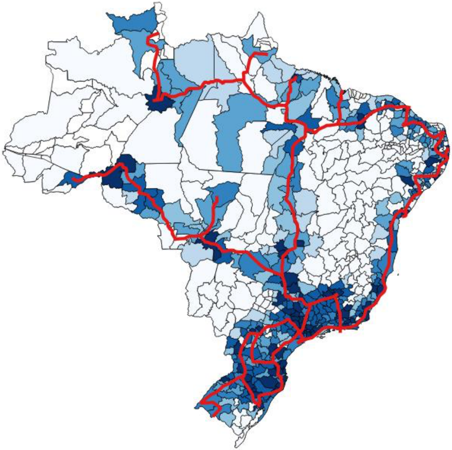
- Estimated model predicts average change in N in city m across all sectors $(\triangle \bar{N}_m)$

Assessing structural estimates

- Estimated model predicts average change in N in city m across all sectors ($\triangle\bar{N}_m$)
  - Correlation ($\triangle\bar{N}_m^{actual}$, $\triangle\bar{N}_m^{simulated}$) $= 0.73$



Actual change



Simulated change

Details Back

Details: productivity impacts of division of labor

- Determine the contribution of the division of labor on observed gains to density by shutting down changes in the extent of division of labor across firms
- Implementation steps:
  - Start by estimating

$$\log \psi_{js} = \beta_0 + \beta_1 \log L_j + \delta_s + \iota_j$$

  - Recalibrate the model by
    - fixing distribution of firms and idiosyncratic shocks
    - making firms choose their locations based on productivity shocks
    - fixing N based on average z in their sector
  - Rerun the estimation using the new simulation results
  - Compute the difference in the β_1 estimates

Back

Details: effects of sorting

- Determine the contribution of sorting by shutting down firm's systematic location choice
- Implementation steps:
  - Start by estimating

$$\log \psi_{js} = \beta_0 + \beta_1 \log L_j + \delta_s + \iota_j$$

  - Recalibrate the model by
    - fixing distribution of firms and idiosyncratic shocks
    - making firms choose their locations based on productivity shocks
    - firms **choose** N based on z and city size
  - Rerun the estimation using the new simulation results
  - Compute the difference in the β_1 estimates

Back

Division of labor and productivity

- Recall

$$\log \psi_{js} = \alpha \log L + \log z(1+\log L)^{\upsilon_s} + \log z\,(1+\log N)^{c_s} - \log N\,(1+\log L)^{-\theta_s} + \varepsilon_{jL}$$

- Remove RF agglomeration externalities, estimate

$$\log \psi_{js} = \beta_0 + \beta_1 \log L_j + \delta_s + \iota_j$$

  - $\hat{\beta}_1 = 1.5\%$
  - Division of labor: 18% of the productivity gains in larger cities

- Shutting down systematic choice of L
  - $\triangle\hat{\beta}_1 = 0.76\%$
  - Sorting of firms: about 50% of the productivity differences through **the division of labor**

Back

PNBL long-term GE impacts

- Long-term Local impacts:

$$\triangle_t \log y_m = \alpha + \beta Backbone_m + \varepsilon_m$$

where

- $\triangle_t \log y_m$ log-change in outcome variable in city m between t_0 and t_1
- $Backbone_m = 1$ if city m receives the ICT infrastructure, and 0 otherwise

Dependent var	Log change in no. of estb	Log change in city size	Log change in estb pdty
	(1)	(2)	(3)
Backbone	.0743*** (.0011)	.0751*** (.0033)	.0951*** (.002)
Obs	558	558	558
R-sq	.923	.571	.432

Significance levels: * 10%, ** 5%, ***1%. All regressions include a constant term.

Back

Assessing structural estimates

1. Calibrate change in $H(N, L)$ to match average treatment effect $\triangle N$
   - 1.3% change in $N \implies$ 5.6% reduction in costs
2. Compute $\triangle \bar{N}_m$ within treated cities based on simulated distribution of firms
   - Average $\triangle N$ higher for cities with more complex (high-z, high-c_s) firms
3. Compared predicted heterogeneity in treatment *vs.* actual changes
   - Benchmark correlation: 0.36 (uniform distribution of firms)

Back

Cross-sector characteristics

- Model prediction: in equilibrium, firms from multiple sectors can coexist in the same city

Proposition

In a spatial equilibrium, the geographic distribution of firms in a high c_s sector first-order stochastically dominates that of a lower c_s' sector.

- The matching function is increasing in z
- The slope and absolute level depend on c_s
- High-c_s firms benefit more from being in a larger city $\implies$ pushing the matching function up
- In equilibrium, more firms locate in larger cities.

Back | Skip to counterfactual results | Descriptive evidence