Production Networks, Geography and Firm Performance

Andrew B. Bernard  Andreas Moxnes  Yukiko U. Saito
Tuck @ Dartmouth  University of Oslo  RIETI
CEPR and NBER  CEPR and NBER

National Bank of Belgium October 23, 2014
Motivation and Questions

What determines buyer-supplier (firm-to-firm) connections?
What are the consequences for firm performance?

Mechanism:
- Firms have comparative advantage (CA) in producing a given task.
- But searching for suppliers (observing price/quality) is costly.
- Trade-off between benefits from exploiting CA and cost of search.
Motivation and Questions

What determines buyer-supplier (firm-to-firm) connections?
What are the consequences for firm performance?

Mechanism:
- Firms have comparative advantage (CA) in producing a given task.
- But searching for suppliers (observing price/quality) is costly.
- Trade-off between benefits from exploiting CA and cost of search.
Implications

1) Variation in firm output and productivity across space (Sveikauskas 1975, Glaeser and Mare 2001, Combes et al 2012).
   - Central locations have lower costs of finding high quality suppliers & outsource more.

2) Substantial heterogeneity in firm sales (w/in localities and industries).
   - High productivity firms have an incentive to search harder for good suppliers.

3) Effect of infrastructure on firm performance.
   - Lowers the cost of searching for suppliers.
Implications

1) Variation in firm output and productivity across space (Sveikauskas 1975, Glaeser and Mare 2001, Combes et al 2012).
   - Central locations have lower costs of finding high quality suppliers & outsource more.

2) Substantial heterogeneity in firm sales (w/in localities and industries).
   - High productivity firms have an incentive to search harder for good suppliers.

3) Effect of infrastructure on firm performance.
   - Lowers the cost of searching for suppliers.
Implications

1) Variation in firm output and productivity across space (Sveikauskas 1975, Glaeser and Mare 2001, Combes et al 2012).
   - Central locations have lower costs of finding high quality suppliers & outsource more.

2) Substantial heterogeneity in firm sales (within localities and industries).
   - High productivity firms have an incentive to search harder for good suppliers.

3) Effect of infrastructure on firm performance.
   - Lowers the cost of searching for suppliers.
Three Components of the Paper

- Facts about (Japanese) production network.
  - Comprehensive data on (nearly) complete production network.

- Model of producers and domestic sourcing.
  - Building on Antras, Fort and Tintelnot (2014).

- ‘Natural’ experiment testing predictions of model (and effects of infrastructure).
  - Kyushu Shinkansen expansion in 2004.
  - Up to 75% fall in travel time for persons, 0% for goods.

Disclaimer: This paper is not about relocation of inputs or firms - within-firm identification only.
Literature


Data Sources

TSR (Tokyo Shoko Research):
- Credit reporting agency (1 of 2 in Japan)
- Buyer-supplier linkages in 2005 & 2010 + firm sales & geolocation.
- 950,000+ firms in private sector.
  - Close to complete coverage of firms with 5+ employees.
    - Not limited to a particular sector.
    - More than 50% of all firms in Japan (relative to census).

Kikatsu:
- All firms with 50+ employees & capital of more than 30 million yen (USD 300,000).
Each firm provides a rank ordered list of suppliers & customers (max 24).

We use union of own-reported and other-reported information.

- A supplies B if both firms are in the TSR data and
  - A reports B as customer or
  - B reports A as supplier.
(TSR) In-degree = 2 (1 own-reported + 1 other-reported)

(TSR) Out-degree = 2 (1 own-reported + 1 other-reported)
Geography of the Network

5% sample of the location-pairs with 5+ connections in 2005:

Sources in green, recipients in blue.
The Production Network: Facts

Key relationships that inform the model:

1. Larger firms have more suppliers.
2. The majority of connections is formed locally.
3. Larger firms have suppliers in more locations and their distance to suppliers is higher.
4. Within firms, distance is positively correlated with supplier/customer quality.
5. Negative degree assortativity among sellers and buyers.
Fact 1: Larger firms have more suppliers.
Fact II: The majority of connections is formed locally.

Median (mean) distance to connections: 30 (172) km.
Fact III: Larger firms have suppliers in more locations

Slope = 0.27

kernel = epanechnikov, degree = 0, bandwidth = .21, pwidth = .32
Fact III: Larger firms have higher distance to suppliers

Median distance to suppliers

Sales, mill yen

Kernel = epanechnikov, degree = 0, bandwidth = .41, pwidth = .61

Slope = 0.32
Fact IV: Distance and supplier/buyer quality

<table>
<thead>
<tr>
<th>Suppliers</th>
<th>Within-buyer regressions</th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>In-degree</td>
<td>Out-degree</td>
<td>Employment</td>
<td>Sales/worker</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Distance</td>
<td>0.12&lt;sup&gt;a&lt;/sup&gt;</td>
<td>0.17&lt;sup&gt;a&lt;/sup&gt;</td>
<td>0.15&lt;sup&gt;a&lt;/sup&gt;</td>
<td>0.04&lt;sup&gt;a&lt;/sup&gt;</td>
<td></td>
</tr>
</tbody>
</table>

includes industry, island, supplier-prefecture and buyer fixed effects

<table>
<thead>
<tr>
<th>Customers</th>
<th>Within-supplier regressions</th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>In-degree</td>
<td>Out-degree</td>
<td>Employment</td>
<td>Sales/worker</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Distance</td>
<td>0.07&lt;sup&gt;a&lt;/sup&gt;</td>
<td>0.06&lt;sup&gt;a&lt;/sup&gt;</td>
<td>0.06&lt;sup&gt;a&lt;/sup&gt;</td>
<td>0.01&lt;sup&gt;a&lt;/sup&gt;</td>
<td></td>
</tr>
</tbody>
</table>

includes industry, island, buyer-prefecture and supplier fixed effects

- Within buyers, supplier quality is increasing in distance.
- Within suppliers, buyer quality is increasing in distance.
Fact V: Negative degree assortativity

A firm with more suppliers - those suppliers have fewer customers.
A firm with fewer suppliers - those suppliers have more customers.
Fact V: Negative assortativity revisited

Note: 2005 data. The horizontal and vertical axes show deciles \((d_x, d_y)\) of the sales distribution for a seller and buyer respectively. The colors represent the number of matches between sellers in the \(d_x\)'th decile and buyers in the \(d_y\)'th decile, with darker colors representing more matches.
The Model

Key elements:

- Firms with productivity $z$ require a unit continuum of tasks to operate (materials processing, accounting, ..).
- They can produce tasks themselves or outsource.
- Task prices/quality vary across locations.
- Observing them requires a cost $f$.

$\rightarrow$ High $z$ firms outsource more tasks and search more locations $\rightarrow$ MC ↓.

$\rightarrow$ Lower search costs $f$ facilitate more/better buyer-seller matches $\rightarrow$ MC ↓→ firm-level growth.
The Model

We build on Antras et al (2014) and introduce:

- In-house production vs outsourcing margin.
- Continuum of locations, because of 1000s of locations in our data.

Model for domestic economy → Abstract from wage and average productivity differences across locations.
The Model: Details

- Downstream firms use labor and a unit continuum of tasks $\omega$. PF $y = zl^\alpha \nu^{1-\alpha}$ where $\nu$ is CES task composite, $z$ is efficiency.
- No trade costs on final good. Fixed measure of firms in each location $m(i)$.
- Firm-specific tasks can be produced in-house or outsourced.
  - Firm has stochastic in-house task productivities from Frechet $(T_0, \theta)$.
  - Firm observes average price of location $i$ but not individual prices $p(\omega, i)$.
  - Knowing every $p(\omega, i)$ in $i$ costs $f(j)$, paid in terms of labor.
- Upstream firms:
  - Use only labor. Task productivity in location $i$ from Frechet $(T, \theta)$.
  - Iceberg trade costs $\tau(i, j)$.
- Perfect competition in upstream stage, monopolistic competition downstream.
Outsourcing

Solve by backwards induction:

- Conditional on locations searched, firm chooses share of tasks to outsource $o(z,j)$ and allocation across locations.
- Firm chooses highest trade cost location $\bar{\tau}(z,j)$ it is willing to search.
  - $\bar{\tau}(z,j)$ chosen to balance the benefit of lower MC against the cost of search.
More productive firms outsource more tasks and therefore have more suppliers:

$$\frac{\partial \ln o(z,j)}{\partial \ln z} > 0,$$

More productive firms search more and costlier locations:

$$\frac{\partial \ln \bar{\tau}}{\partial \ln z} > 0$$

Negative degree assortivity: Higher $z$ (high indegree) → firm reaches costlier locations → suppliers there are on average not very competitive in $z$’s home market (low avg. outdegree).
A Distributional Assumption

- Every location faces a density of trade costs $g(\tau,j)$.
- Assume $g()$ inverse Pareto with shape $\gamma > \theta$ and support $[1, \tau_H]$.
  - A location has few nearby markets and many remote ones.
- Density fits empirical distance distribution well.
Two Propositions

Proposition

(i) Lower search costs $f(j)$ lead to growth in sales among downstream firms in $j$. (ii) Sales growth is stronger in input-intensive (low $\alpha$) relative to labor intensive (high $\alpha$) industries.

$$\frac{\partial \ln r(z,j)}{\partial \ln f(j)} < 0 \quad \text{and} \quad \frac{\partial^2 \ln r(z,j)}{\partial \ln f(j) \partial \alpha} > 0$$

Two channels:

1. Direct: low $\alpha$ firms grow more bc inputs large share of total costs.
2. Indirect: low $\alpha$ firms search more markets when $f(j) \downarrow (|\partial \bar{\tau}/\partial f| \text{ decreasing in } \alpha)$. 
Two Propositions

Proposition

Lower search costs $f(j)$ lead to more outsourcing and suppliers from new locations (higher $\bar{\tau}$) among downstream firms in $j$.

\[
\frac{\partial o(z,j)}{\partial f(j)} < 0 \quad \text{and} \quad \frac{\partial \bar{\tau}(z,j)}{\partial f(j)} < 0.
\]
Shinkansen - A Natural Experiment

Does cheaper search improve firm performance by facilitating linkages in the production network?

- Use opening of high-speed train network (Shinkansen) in southern Japan in 2004. Key advantages:

1. Dramatic reduction in travel time between stations.
   - 75% reduction for many city pairs.

2. Goods do not travel by Shinkansen, just people.
   - No contemporaneous reduction in travel time for goods along this southern route.

3. Likely exogenous.
   - Planned decades in advance. Timing of completion was subject to substantial uncertainty.
- Rail line connecting two prefectures (Kagoshima + Kumamoto) with total population of 3.5 million (~CT).
- Travel time Kagoshima → Shin-Yatsushiro: 130 → 35 min.
- Kagoshima → Hakata: 4 → 2 hours.
800 series Shinkansen

- Operating speed 260 km/h.
- 2-3 departures / hour. Capacity 392 passengers per trainset.
Empirical Methodology

- Lower travel time should benefit input-intensive firms more than labor intensive firms (Proposition 1).
  - Lower $f(j)$ has no impact on MC of firms belonging to $\alpha = 1$ industries.
- Classify industries $k$ according to their 2003 intermediate input use:
  
  $H_k = 1 - \text{labor share of industry}_k$

- Define $Treat_f = 1$ if firm $f$ is $< 30$ km from new Shinkansen station (stations between Kagoshima and Shin-Yatsushiro).
- Dependent variables: lnSales, ln(sales/employee), TFP (Olley-Pakes); relative to industry-year means.
Empirical Methodology

- Estimate for 2000-2008 period

\[
\ln y_{fkrt} = \alpha^1_f + \alpha^2_{rt} + \beta_1 \text{Treat}_f \times H_k \times \text{Post}2004_t + \gamma X_{fkrt} + \epsilon_{fkrt},
\]

where \( \alpha^1_f \) and \( \alpha^2_{rt} \) are firm and prefecture-year fixed effects.

- Triple differences:
  - Pre to post shock (1st diff)
  - Firm near station relative to firm not near station (2nd diff).
  - High \( H_k \) relative to low \( H_k \) firms (3rd diff).

- Positive \( \beta_1 \) if high \( H_k \) firms are growing faster relative to low-\( H_k \) firms near new stations \textit{relative to elsewhere}.

- More controls:
  - Time-varying geographic controls by using average performance in \( f \)'s municipality (\( \approx 1400 \) municipalities).
  - Remaining interactions (\( \text{Treat}_f \times H_k \), etc.).
Potential Concerns

1. We pick up market access (demand side) effects:
   - No, because demand ↑ should affect both input- and labor-intensive firms.

2. We pick up different trends for input- and labor-intensive firms:
   - No, industry trends are differenced out.

3. The location of the stations are endogenous:
   - Not a problem as long as locations are not determined based on differential growth for input/labor intensive industries.

4. Pre-trends; input-intensive firms near new stations always grow faster relative to labor-intensive firms:
   - No evidence of this in placebo test.
## Results

### Table: Firm Performance

<table>
<thead>
<tr>
<th></th>
<th>Sales</th>
<th>Sales/employee</th>
<th>TFP</th>
</tr>
</thead>
<tbody>
<tr>
<td>$Treat_f \times H_j \times Post2004_t$</td>
<td>0.47**</td>
<td>0.42*</td>
<td>0.29**</td>
</tr>
<tr>
<td></td>
<td>(2.12)</td>
<td>(1.76)</td>
<td>(2.44)</td>
</tr>
<tr>
<td>Firm and city controls</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
</tr>
<tr>
<td>Prefecture-year FE</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
</tr>
<tr>
<td>Firm FE</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
</tr>
<tr>
<td>N</td>
<td>148,264</td>
<td>146,466</td>
<td>145,058</td>
</tr>
<tr>
<td>R-sq</td>
<td>0.97</td>
<td>0.92</td>
<td>0.94</td>
</tr>
</tbody>
</table>

Note: Robust t-statistics in parentheses. Dependent variables in logs.

- A Shinkansen station increases sales by 0.47 log points more for a firm with $H_k = 1$ relative to a firm with $H_k = 0$.
- A firm in the 9th decile of the $H_k$ distribution (industrial plastic products) increased sales by 0.10 log points more than a firm in the 1st decile of the $H_k$ distribution (general goods rental and leasing).
Robustness: Placebo


<table>
<thead>
<tr>
<th></th>
<th>Sales</th>
<th>Sales/employee</th>
<th>TFP</th>
</tr>
</thead>
<tbody>
<tr>
<td>$Treat_f \times H_j \times Post2000_t$</td>
<td>-0.30</td>
<td>-0.05</td>
<td>0.02</td>
</tr>
<tr>
<td></td>
<td>(1.05)</td>
<td>(0.22)</td>
<td>(0.17)</td>
</tr>
<tr>
<td>Firm and city controls</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
</tr>
<tr>
<td>Prefecture-year FE</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
</tr>
<tr>
<td>Firm FE</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
</tr>
<tr>
<td>N</td>
<td>66,756</td>
<td>66,756</td>
<td>66,487</td>
</tr>
<tr>
<td>R-sq</td>
<td>0.99</td>
<td>0.94</td>
<td>0.95</td>
</tr>
</tbody>
</table>

Note: Robust t-statistics in parentheses. Dependent variables in logs.
More robustness

1. Labor supply - Recruiting now easier for knowledge intensive industries (which may happen to be input intensive).
   - Calculate R&D intensity of industries, add additional interactions → no change in results.

2. The 'straw effect' - less economic activity in nearby locations.
   - Add interactions for firms 30-60km from new station → small negative effect for these firms & no change in main results.

3. Demand side again - Input intensive industries may have more remote customers.
   - Should not see TFP effects.
   - $corr(\text{avg distance to customers}, H_j) = -0.02$.

4. Drop construction industry.

5. Change 30 km threshold.
Shinkansen - New Connections

- Mechanism: Should see more supplier linkages in treated regions.
- Divide Japan into a grid consisting of $500 \times 500$ locations ($5.6^2 \text{ km}^2$).
- Number of connections from $i$ to $j$ at time $t$ is $C_{ijt}$, $t = \{2005, 2010\}$.
- Regress

$$\Delta \ln C_{ij} = \xi^1_i + \xi^2_j + \beta_1 \text{Both}_{ij} + \beta_2 \text{One}_{ij} + \gamma X_{ij} + \varepsilon_{ij},$$

where $\xi^1_i$ and $\xi^2_j$ are source and destination FE, $\text{Both}_{ij} = 1$ if both locations $i$ and $j$ get a new station, $\text{One}_{ij} = 1$ if one of them gets a new station.

- Notes:
  - Using 2005-2010 changes at the firm level is (currently) problematic because of cross-year matching issues.
  - Lower bound bc data starts in 2005.
### Table: Connections: Descriptive statistics.

<table>
<thead>
<tr>
<th></th>
<th>Mean</th>
<th>Median</th>
<th>Std. dev.</th>
<th>min</th>
<th>max</th>
</tr>
</thead>
<tbody>
<tr>
<td>$C_{ij}^{2005}$</td>
<td>7.05</td>
<td>2</td>
<td>62.35</td>
<td>1</td>
<td>16507</td>
</tr>
<tr>
<td>$C_{ij}^{2010}$</td>
<td>7.51</td>
<td>2</td>
<td>59.93</td>
<td>1</td>
<td>14808</td>
</tr>
<tr>
<td>$\Delta \ln C_{ij}$</td>
<td>0.08</td>
<td>0</td>
<td>0.56</td>
<td>-3.47</td>
<td>3.78</td>
</tr>
<tr>
<td>$Both_{ij}$</td>
<td>0.01</td>
<td>0</td>
<td>0.08</td>
<td>0</td>
<td>1</td>
</tr>
<tr>
<td>$One_{ij}$</td>
<td>0.02</td>
<td>0</td>
<td>0.14</td>
<td>0</td>
<td>1</td>
</tr>
<tr>
<td>Firms per cell</td>
<td>104.21</td>
<td>19</td>
<td>463.30</td>
<td>1</td>
<td>21,207</td>
</tr>
<tr>
<td># sources</td>
<td>7,613</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td># destinations</td>
<td>8,054</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td># obs</td>
<td>386,294</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
Shinkansen – New Connections

<table>
<thead>
<tr>
<th></th>
<th>(1)</th>
<th>(2)</th>
<th>(3)</th>
<th>(4)</th>
</tr>
</thead>
<tbody>
<tr>
<td>( \text{Both}_{ij} )</td>
<td>0.07**</td>
<td>0.12**</td>
<td>0.39**</td>
<td>0.42**</td>
</tr>
<tr>
<td></td>
<td>(5.91)</td>
<td>(7.91)</td>
<td>(20.12)</td>
<td>(7.93)</td>
</tr>
<tr>
<td>( \text{One}_{ij} )</td>
<td>-0.02**</td>
<td>-0.01</td>
<td>0.19**</td>
<td>0.15**</td>
</tr>
<tr>
<td></td>
<td>(3.56)</td>
<td>(0.74)</td>
<td>(19.87)</td>
<td>(6.42)</td>
</tr>
<tr>
<td>( \ln \text{Dist}_{ij} )</td>
<td></td>
<td></td>
<td>-0.06**</td>
<td>-0.06**</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>(71.32)</td>
<td>(81.98)</td>
</tr>
<tr>
<td>( \text{Both}<em>{ij} \times \ln \text{Dist}</em>{ij} )</td>
<td></td>
<td></td>
<td>-0.01</td>
<td>(0.86)</td>
</tr>
<tr>
<td>( \text{One}<em>{ij} \times \ln \text{Dist}</em>{ij} )</td>
<td></td>
<td></td>
<td>0.01*</td>
<td>(1.87)</td>
</tr>
<tr>
<td>\text{Destination FE}</td>
<td>No</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
</tr>
<tr>
<td>\text{Source FE}</td>
<td>No</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
</tr>
<tr>
<td># obs</td>
<td>386,294</td>
<td>386,294</td>
<td>386,294</td>
<td>386,294</td>
</tr>
<tr>
<td># sources</td>
<td>7,613</td>
<td>7,613</td>
<td>7,613</td>
<td>7,613</td>
</tr>
<tr>
<td># destinations</td>
<td>8,054</td>
<td>8,054</td>
<td>8,054</td>
<td>8,054</td>
</tr>
<tr>
<td>R-sq</td>
<td>0.00</td>
<td>0.17</td>
<td>0.18</td>
<td>0.18</td>
</tr>
</tbody>
</table>

Note: Bootstrapped t-statistics in parentheses with 200 replications. Dependent variable is \( \Delta \ln C_{ij} = \ln C_{ij}^{2010} - \ln C_{ij}^{2005} \).

*** significant at the 0.01 level, ** significant at the 0.05 level, * significant at the 0.1 level.
Conclusions

- The supply network matters for firm performance:
  - Infrastructure shock generates significant performance gains.
  - Evidence that gains are related to new buyer-seller linkages, as suggested by model.

- Directions for future work:
  - The entry margin - how heterogeneous firms sort into different markets.
  - Identifying additional firm performance channels.
Appendix: Degree Distributions

- 3,783,711 supplier-customer connections.
- Among firms with positive degree:
  - Mean (median) # customers is 5.6 (1).
  - Mean (median) # suppliers is 4.9 (2).
- 1/slope is -1.32 (in-degree) and -1.50 (out-degree).
3.4 mill passengers in 2004, traveling 120km on average.
24% increase in passenger count by 2008.
> 100% increase relative to 2003 levels.
50% fare increase pre to post Shinkansen (currently USD 75 Hakata-Kagoshima).
The modal shares of railways and airlines changed from 41% to 71% and 42% to 12% respectively between Fukuoka and Kagoshima prefectures (2000 to 2005). (Tokyo Institute of Technology, 2008).
Shinkansen Factsheet

- Total length 2,388 km, connects majority of JP population.
- Share of train passenger traffic larger than in any other country.
  - Rail has 28% of total passenger km in JP; 1% in US and 8% in France
  - Car has 50% in JP, 85% in US and France (Clever et al 2008).
Shinkansen Factsheet

Shinkansen dominates medium distance travel:

![Graph showing the share of the Shinkansen in various long-distance transport modes.](image)

Share of the Shinkansen in various long-distance transport modes.