Follow Thy Neighbor: The Role of First Exporters

Hao Fe[1] **Yang Liang**[1] Mary E. Lovely[2]

[1]San Diego State University
[2]Peterson Institute for International Economics

SEA 94th Annual Meeting

Motivation

- Firms face considerable uncertainty when entering and exporting to new foreign markets.
  - Bartelsman et al. 2009; Eaton et al. 2008; Albornoz et al. 2012; Freund and Pierola, 2010

Motivation

- Firms face considerable uncertainty when entering and exporting to new foreign markets.
    - Bartelsman et al. 2009; Eaton et al. 2008; Albornoz et al. 2012; Freund and Pierola, 2010
- To reduce the uncertainty, firms can learn from self-exploration in an uncertain environment (within firm, e.g., Rauch and Watson, 2003)

Motivation

- Firms face considerable uncertainty when entering and exporting to new foreign markets.
  - Bartelsman et al. 2009; Eaton et al. 2008; Albornoz et al. 2012; Freund and Pierola, 2010
- To reduce the uncertainty, firms can learn from self-exploration in an uncertain environment (within firm, e.g., Rauch and Watson, 2003)
- However, self-discovery can be costly in foreign markets and learning from others can be a potentially more advantageous strategy (Hausmann and Rodrik, 2003)
  - Firm Structure and Production Network (Bernard et al., 2018..., Lovely et al., 2005, Monarch, 2016...)

Motivation

- Firms face considerable uncertainty when entering and exporting to new foreign markets.
  - Bartelsman et al. 2009; Eaton et al. 2008; Albornoz et al. 2012; Freund and Pierola, 2010
- To reduce the uncertainty, firms can learn from self-exploration in an uncertain environment (within firm, e.g., Rauch and Watson, 2003)
- However, self-discovery can be costly in foreign markets and learning from others can be a potentially more advantageous strategy (Hausmann and Rodrik, 2003)
  - Firm Structure and Production Network (Bernard et al., 2018..., Lovely et al., 2005, Monarch, 2016...)
- Can firms learn from neighboring firms before undertaking risky investments?

Motivation

- Recent studies have started to explore how learning from neighbors affects firms' export decisions.
  - e.g. Koenig 2009; Koenig et al. 2010; Iacovone and Javorcik 2010; Fernandes and Tang 2014; Kamal and Sundaram 2016; He and Lugovskyy 2018...

Motivation

- Recent studies have started to explore how learning from neighbors affects firms' export decisions.
  - e.g. Koenig 2009; Koenig et al. 2010; Iacovone and Javorcik 2010; Fernandes and Tang 2014; Kamal and Sundaram 2016; He and Lugovskyy 2018...
- **Big Challenge: Identification**
  - Studies face one crucial question: how to define "neighboring firms"
  - Neighborhood is too large $\rightarrow$ **overestimation**
    - self-selection
    - common-market demand shock

Motivation

- Recent studies have started to explore how learning from neighbors affects firms' export decisions.
  - e.g. Koenig 2009; Koenig et al. 2010; Iacovone and Javorcik 2010; Fernandes and Tang 2014; Kamal and Sundaram 2016; He and Lugovskyy 2018...
- **Big Challenge: Identification**
  - Studies face one crucial question: how to define "neighboring firms"
  - Neighborhood is too large $\rightarrow$ **overestimation**
    - self-selection
    - common-market demand shock
- Due to the lack of detailed data on the location of exporters, most recent empirical studies treated city, province, or country as the geographic level where the learning happens.

Motivation

- Recent studies have started to explore how learning from neighbors affects firms' export decisions.
  - e.g. Koenig 2009; Koenig et al. 2010; Iacovone and Javorcik 2010; Fernandes and Tang 2014; Kamal and Sundaram 2016; He and Lugovskyy 2018...
- **Big Challenge: Identification**
  - Studies face one crucial question: how to define "neighboring firms"
  - Neighborhood is too large → **overestimation**
    - self-selection
    - common-market demand shock
- Due to the lack of detailed data on the location of exporters, most recent empirical studies treated city, province, or country as the geographic level where the learning happens.
- This paper: uses detailed GIS data for Chinese firms, and provides new empirical evidence with a new identification strategy at a much finer geographical level
  - We believe learning occurs in a geographic level smaller than city
  - Focusing on small regions helps reduce endogeneity problems

What we do?

- Put together a novel dataset that combines both the China Custom Records (firm-destination-product-year) and detailed geographic locations (neighborhood(neighborhood/sub-district) $<$ district $<<$ city).

What we do?

- Put together a novel dataset that combines both the China Custom Records (firm-destination-product-year) and detailed geographic locations (neighborhood(neighborhood/sub-district) $<$ district $<<$ city).

- Allow us to redefine much detailed clusters and relax the ”networks” from only using city- or province-level political boundaries.

What we do?

- Put together a novel dataset that combines both the China Custom Records (firm-destination-product-year) and detailed geographic locations (neighborhood(neighborhood/sub-district) $<$ district $<<$ city).
- Allow us to redefine much detailed clusters and relax the "networks" from only using city- or province-level political boundaries.
- Examine the effect of a newly formed seller-market route on other firms exporting behaviors.

What we do?

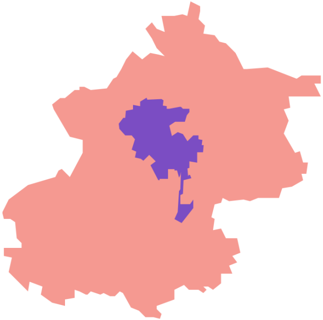
- Put together a novel dataset that combines both the China Custom Records (firm-destination-product-year) and detailed geographic locations (neighborhood(neighborhood/sub-district) $<$ district $<<$ city).
- Allow us to redefine much detailed clusters and relax the "networks" from only using city- or province-level political boundaries.
- Examine the effect of a newly formed seller-market route on other firms exporting behaviors.
- How does this likelihood vary with firm characteristics (i.e. size, ownership type, or product-relatedness)?

Data Sources

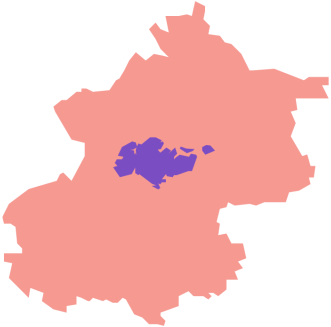
1. Chinese customs transaction records database between 2000 and 2006, for each transaction, the dataset reports
   - Firm name, exact address, type (state-owned, private, foreign, etc.)
   - Product category (6,986 HS 8-digit categories; 97 HS 2-digit categories), exporting quantity and value
   - Destination country (232 destinations), transportation mode (road, sea, rail or air)

2. We then geo-code firm addresses:
   - ArcGIS World geocoding feature + google API
   - Nine largest exporting municipalities/cities in China ($> 35\%$ exporting firms, $> 75\%$ total exports, 54,107 firms, 98% matching rate, , 3790 miles^2) [How Large?]
   - District/county Level (12 districts per city):
   - Subdistrict/neighborhood Level (124 subdistricts per city): 1 miles^2, 129 firms on average.

How large are these urban areas? - Map Battle



(a) Beijing is 13 × as big as Los Angeles

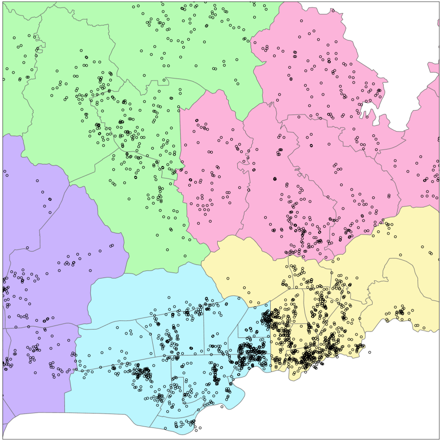


(b) Beijing is 23 × as big as Singapore

Comparative sizes of Beijing with Los Angeles and Singapore.

Shenzhen's Districts and Neighborhood/Subdistricts

subdistricts of the same district (same color)



Exporters in 2000

Exporters in 2006

Empirical Approach

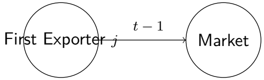
Empirical Challenges

- Firms sort into geographical locations when they choose their offices.
  - This might be driven by some local policies promoting exports. (i.e., SEZs, FTZs)
- Firms that are geographically close tend to be similar (company type, size, products, etc)
- Thus it is difficult to tell the correlated exporting behaviors are due to
  - confounders: firm sorting/demand shocks hit firms with similar characteristics
    OR
  - firms' learning from each other

Identification Strategies

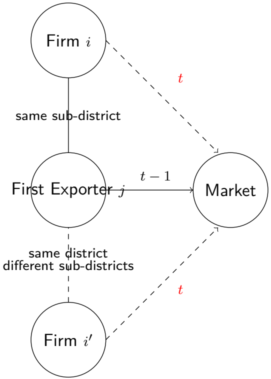
- We compare the exporting outcomes of firms who are located in the same district/county but different sub-districts .
- The key identifying assumption is (Bayer, Ross and Topa 2008; Schmutte 2015):
  - Within these small geographical levels (districts/counties), variation across subdistricts/neighborhood in network connection is uncorrelated with unobserved factors that affect the exporting outcome.
- The economic rationale: firms sort into districts but not subdistricts/neighborhood.

Defining markets; forming firm pairs (dyads)



market is defined as: product (p) $\times$ destination country (c)

Defining markets; forming firm pairs (dyads)



market is defined as: product (p) $\times$ destination country (c)

Baseline

Dyadic Model (Pair, i, j)

Firm i, district d, destination country c, product p year t:

$$\text{export}_{i,p,c,t} = \text{same neighborhood}_{i,j(p,c,t-1)} + \chi_{d,t} + \chi_{p,c,t} + \alpha_i + \epsilon_{i,p,c,t},$$

- a foreign market is then defined by: product (p) $\times$ destination country (c).
- Peer Effect: same neighborhood$_{i,j(p,c,t-1)} = 1$, if, at $t-1$, there is a first exporter j begins exporting to a new market in a district and the new exporter co-locate in the same neighborhood (0 for other neighborhoods).
- export$_{i,p,c,t} = 1$, if firm i exports to market $p \times c$ at t.
- $\chi_{d,t}$ includes: district $\times$ year FE $\rightarrow$ district specific export supply shocks
- $\chi_{p,c,t}$ includes: product $\times$ destination $\times$ year FE $\rightarrow$ destination-sector specific import demand shocks
- α_i firm FE
- Standard errors in parentheses are clustered at the district level.

Summary Statistics

Firm Descriptive Statistics

Firm-year level	First exporters		Neighboring firms	
	Mean	St.D	Mean	St.D
Number of products (HS8)	57.785	124.427	22.135	71.090
Number of destination countries	20.867	21.966	9.113	14.651
Export quantity (in millions)	16.279	147.862	7.296	92.644
Export value (2006 RMB in millions)	21.723	377.997	10.247	223.362
Proportion of ordinary trade	0.795	0.354	0.652	0.445
Foreign owned	0.434	0.496	0.618	0.486
Joint venture	0.210	0.407	0.257	0.437
Wholly foreign owned	0.225	0.417	0.362	0.481
Domestic	0.556	0.497	0.368	0.482
State owned	0.289	0.453	0.146	0.354
Private owned	0.223	0.416	0.195	0.396
Collective enterprises	0.000	0.000	0.000	0.008
Other domestic	0.044	0.205	0.026	0.159
Distance to closest highway (miles)	1.055	1.335	0.938	1.113
Distance to closest subway (miles)	24.763	58.922	17.577	49.277
Number of firms	15,059		37,601	

Note: Summary statistics for the first exporters and their neighboring firms in our sample. All statistics are on the basis of the firm-year level sample. There are two trade modes in general: processing and ordinary. Processing trade refers to trade flows by Chinese firms importing raw materials or intermediate inputs from abroad, processing them locally and exporting the value-added goods. Ordinary trade includes all other trade flows. All firm type variables are indicators and their summaries represent the fraction of the sample with the associated characteristic.

Summary Statistics

Number of Firms by City

	Number of firms	Number of first exporters
Beijing	3891	2000
Tianjin	3478	1787
Shanghai	9571	3324
Shenzhen	9314	2230
Ningbo	2470	1199
Guangzhou	3942	1691
Hangzhou	1960	1067
Nanjing	1433	854
Fuzhou	1542	907

Note: This table includes firms in the main estimation sample.

District Sorting

Sorting within Districts, R^2 Method

	Raw	District FEs
Number of products (HS8)	0.024	0.012
Number of destinations	0.032	0.006
Export value	0.000	0.000
Export quantity	0.000	0.000
Domestic firm	0.104	0.029
Ordinary trade	0.197	0.024
Processing trade	0.223	0.021
Transportation: water	0.178	0.012
Transportation: land	0.476	0.006
Transportation: air	0.096	0.011

Note: This table reports sorting within districts on observables. The input dataset contains one randomly selected firm-level observation per subdistrict and the fraction of firms (excluding the firm) in the subdistrict who share the listed characteristic or its average. Each entry is the R^2 from a regression of the firm's characteristic on the subdistrict-level average. Columns 2 controls for district specific fixed effects and reports within R^2. The sample is restricted to subdistricts with more than six firms.

Baseline Results

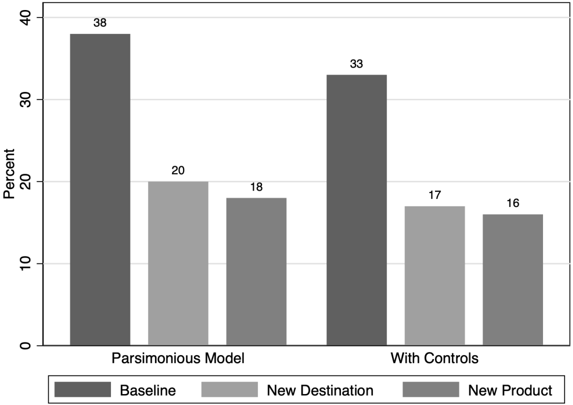
Effects of First Exporters

	(1)	(2)	(3)	(4)
Neighbor	0.014***	0.012***	0.017***	0.014***
	(0.003)	(0.003)	(0.003)	(0.003)
Percentage of baseline probability	38%	33%	46%	43%
R-squared	0.010	0.011	0.022	0.024
Observations	55,175,656	55,175,656	55,175,649	55,175,649
District by Year FEs	✓	✓	×	×
Product by Destination by Year FEs	✓	✓	×	×
Firm FEs	✓	✓	×	×
Firm by Year FEs	×	×	✓	✓
District by Product by Destination FEs	×	×	✓	✓
Additional controls	×	✓	×	✓

Note: The table reports our main estimation results. Coefficients and standard errors are multiplied by 100 for ease of readability. An observation is a pair of firms who are located in the same district. The first firm in the pair is a local (district) first exporter that exports HS-2 product p to destination c in year $t-1$. The second firm in the pair (neighboring firm) has not exported p to c by $t-1$. The dependent variable is whether the neighboring firm exports p to destination c in year t. The independent variable is whether the first exporter and the neighboring firm are at the same subdistrict. Column (1) estimates the baseline equation. Columns (2) and (4) add seven control variables to columns (1) and (3) respectively. Standard errors in parentheses are clustered at the district level.

* $p < .10$, ** $p < .05$, *** $p < .01$

Destination or Product?



Effects of First Exporters (Exported the Same Product Before)

	(1)	(2)	(3)	(4)
Neighbor	0.061***	0.054***	0.067***	0.061***
	(0.014)	(0.014)	(0.013)	(0.013)
Percentage of baseline probability	20%	17%	22%	20%
R-squared	0.044	0.044	0.115	0.115
Observations	6,723,287	6,723,287	6,723,287	6,723,287
District by Year FEs	✓	✓	×	×
Product by Destination by Year FEs	✓	✓	×	×
Firm FEs	✓	✓	×	×
Firm by Year FEs	×	×	✓	✓
District by Product by Destination FEs	×	×	✓	✓
Additional controls	×	✓	×	✓

Note: The table restricts the sample to neighboring firms with prior product-specific experience. An observation is a pair of firms who are located in the same district. The first firm in the pair is a local (district) first exporter that exports HS-2 product p to destination c in year $t-1$. The second firm in the pair (neighboring firm) has not exported p to c by $t-1$ but has exported p to other destination(s). The dependent variable is whether the neighboring firm exports p to destination c in year t. The independent variable is whether the first exporter and the neighboring firm are at the same subdistrict.

* $p < .10$, ** $p < .05$, *** $p < .01$

Effects of First Exporters (Exported the Same Destination Before)

	(1)	(2)	(3)	(4)
Neighbor	0.103***	0.081***	0.124***	0.102***
	(0.024)	(0.021)	(0.023)	(0.020)
Percentage of baseline probability	18%	16%	21%	19%
R-squared	0.047	0.053	0.106	0.111
Observations	5,819,092	5,819,092	5,819,092	5,819,092
District by Year FEs	✓	✓	×	×
Product by Destination by Year FEs	✓	✓	×	×
Firm FEs	✓	✓	×	×
Firm by Year FEs	×	×	✓	✓
District by Product by Destination FEs	×	×	✓	✓
Additional controls	×	✓	×	✓

Note: The table restricts the sample to neighboring firms with prior product-specific experience. An observation is a pair of firms who are located in the same district. The first firm in the pair is a local (district) first exporter that exports HS-2 product p to destination c in year $t-1$. The second firm in the pair (neighboring firm) has not exported p to c by $t-1$ but has exported p to other destination(s). The dependent variable is whether the neighboring firm exports p to destination c in year t. The independent variable is whether the first exporter and the neighboring firm are at the same subdistrict.
* $p < .10$, ** $p < .05$, *** $p < .01$

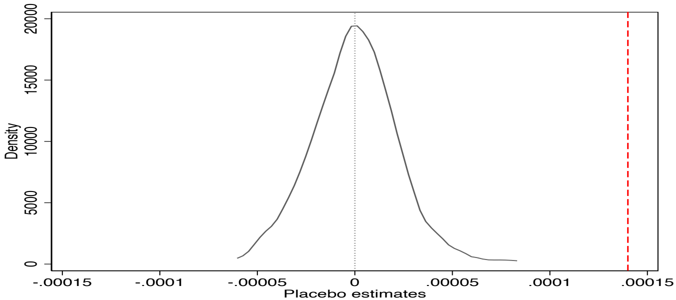
Alternative Samples

	(1)	(2)	(3)	(4)
Neighbor	0.048***	0.015***	0.013***	0.013***
	(0.015)	(0.003)	(0.003)	(0.003)
Percentage of baseline probability	40%	33%	35%	35%
R-squared	0.023	0.011	0.011	0.010
Observations	5,335,075	39,751,467	43,993,871	54,155,255

Note: The table reports the estimation results using different subsamples. Column (1) restricts to the sample of export extensive firms (export value at top 5% of all firms in CCR). Column (2) restricts to the sample of significant products (product annual export value above median of all products in CCR). Column (3) drops the clothing, textile and footwear sectors that benefited from dramatic trade liberalization over the study period. Column (4) excludes three destinations in the Greater China area, i.e. Hong Kong, Macao and Taiwan. Coefficients and standard errors are multiplied by 100 for ease of readability. All specifications use our preferred specification, i.e. equation (**??**). Standard errors in parentheses are clustered at the district level.

* $p < .10$, ** $p < .05$, *** $p < .01$

Density of Placebo Estimates



Note: The figure shows a density plot of the estimated export spillover effects using 100 placebo samples where the first exporters are randomly selected from the firms in our sample that have never exported product p to destination c. Epanechnikov kernel is used. The bandwidth is 0.00001. The vertical red dashed line represents the coefficient estimated in our preferred specification, i.e. column (1) of Table 4. The same identification strategy with district-year fixed effects, product-destination-year fixed effects and peer fixed effects is used to estimate the placebo coefficients.

Placebo Test: Relocated Firms

	(1)	(2)
Peer Effect	1.1008*	0.0006
	(0.5998)	(0.0208)
District by Year FEs	×	✓
Product by Destination by Year FEs	×	✓
Firm FEs	×	✓
Percentage of baseline probability	150%	0.08%
R-squared	0.002	0.089
Observations	34,097,865	34,097,865

Note: The table reports our placebo results using the relocated firms. We artificially create export spillovers between the relocated firms and their neighbors in the subdistricts of their new locations for the time periods before the moves, and then estimate export spillover using the same identification strategy with relocated firms taking the role of first exporters. Column (1) presents the raw estimate without any fixed effects. Column (2) use the same set of fixed effects as our preferred specification. Standard errors in parentheses are clustered at the district level.
* $p < .10$, ** $p < .05$, *** $p < .01$

Heterogenous Results – Firm Ownership Type

	(1)	(2)	(3)	(4)	(5)	(6)
Neighbor	0.008**	0.008**	0.012***	0.012***	0.010**	0.008**
	(0.003)	(0.003)	(0.003)	(0.003)	(0.004)	(0.004)
Domestic firm × Neighbor	0.015**					
	(0.007)					
State owned × Neighbor		0.019**				
		(0.009)				
Private × Neighbor		0.014				
		(0.010)				
Foreign firm$_{1st}$ × Neighbor			0.006*			
			(0.003)			
Joint venture$_{1st}$ × Neighbor				0.009**		
				(0.004)		
Wholly foreign owned$_{1st}$ × Neighbor				0.003		
				(0.004)		
Same ownership type × Neighbor					0.011***	
					(0.004)	
Both domestic × Neighbor						0.013*
						(0.007)
Both foreign × Neighbor						0.009***
						(0.003)
R-squared	0.010	0.010	0.010	0.010	0.010	0.010
Observations			54,279,112			

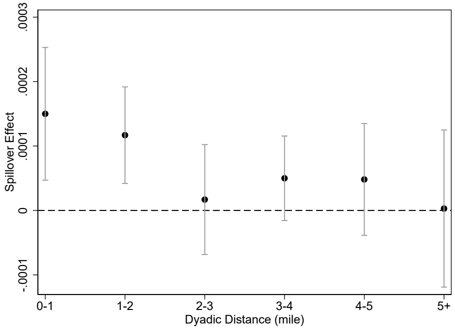
Note: The table reports our heterogeneous results when the variable of interest is interacted with firm ownership type. Coefficients and standard errors are multiplied by 100 for ease of readability. District-year fixed effects, product-destination-year fixed effects and peer fixed effects are used in all columns. Only the coefficients of neighbor and the interaction terms are presented here. Standard errors in parentheses are clustered at the district level.
* $p < .10$, ** $p < .05$, *** $p < .01$

Heterogenous Results - Firm Characteristics and Export Performance

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Neighbor	0.020***	0.023***	0.007*	0.013***	0.013***	0.016***	0.009***	0.005
	(0.004)	(0.005)	(0.004)	(0.004)	(0.004)	(0.005)	(0.004)	(0.003)
Firm characteristics								
Number of products (HS-8) (< median) × Neighbor	-0.011***							
	(0.003)							
Number of destinations (< median) × Neighbor		-0.018***						
		(0.004)						
Percentage of ordinary goods × Neighbor			0.011*					
			(0.006)					
(Number of products$_{1st}$ < median) × Neighbor				0.001				
				(0.003)				
(Number of destinations$_{1st}$ < median) × Neighbor					0.001			
					(0.004)			
Percentage of ordinary trade$_{1st}$ × Neighbor						-0.003		
						(0.005)		
Export performance								
Export value$_{1st}$ (> median) × Neighbor							0.009**	
							(0.003)	
Export quantity$_{1st}$ (> median) × Neighbor								0.009**
								(0.002)
R-squared	0.010	0.010	0.010	0.010	0.010	0.010	0.010	0.010
Observations	55,175,656	55,175,656	55,175,656	55,175,656	55,175,656	55,175,656	55,175,656	55,043,274

Note: The table reports our heterogeneous results when the variable of interest, Neighbor, is interacted with various firm and export characteristics. Coefficients and standard errors are multiplied by 100 for ease of readability. District-year fixed effects, product-destination-year fixed effects and peer fixed effects are used in all specifications. Only the coefficients and standard errors of the variable of interest and the interaction terms are reported. Standard errors in parentheses are clustered at the district level.
* $p < .10$, ** $p < .05$, *** $p < .01$

Spatial Decay



The figure shows the estimated export spillover effects by the distance between the first exporter and its matched firm in the dyad (they may or may not be in the same subdistrict). 95% confidence intervals of these estimates are also presented. The horizontal axis labels indicate the distance range. For example, 0-1 means the dyadic distance is between 0 and 1 mile. Note that the effects in the vertical axis are not multiplied by 100 as in the tables shown above.

Thank you!

yang.liang@sdsu.edu