International Knowledge Diffusion: Role of Immigration and FDI

Aakash Kalyani
Boston University

November 22, 2022

Motivation: Measuring diffusion key for study of economic growth

- Innovation and R&D is highly concentrated across countries/geographies/firms.
- Subsequent diffusion important for overall economic growth.
- Limitation - A measure of diffusion.
- Traditional measure - citations - imperfect, esp. across countries.

This paper:

- Measure diffusion: using variation in similarity of a patent's text to past patents from various countries.

Correlates with trade, FDI, immigration; Large variation at the firm level

Correlated with channels of diffusion and associated with higher firm value:

1. Correlated with trade (at the industry level), FDI (at the firm level) and immigration (at the county level).
2. Patents with higher proximity to advanced countries are more valuable.
3. Firms which file these patents are more productive.

Large variation at the firm level:

1. About 70% of the variation at the firm level.
2. Larger firms have patents which are more proximate to foreign countries.

Table of Contents

Measure of spillovers at the firm level

Measure is based on textual similarity between two patents.

1. Proximity ($Proximity_p^c$) at the patent level:

$$Proximity_p^c = \frac{\text{Avg Similarity to c patents}|t-1 \ and \ t-5}{\text{Avg Similarity to US patents}|t-1 \ and \ t-5}$$

2. Backward similarity calculated with two word combinations. [TF-IDF + dot product]
3. Proximity ($Proximity_f^c$) at the firm level:

$$Proximity_f^c = Avg(Proximity_p^c|p \in f)$$

Example Patent - MATTEL INC - ‘Robotic Humanoid’ more frequent in JP patents

- **Title** : Reconfigurable toy
- **Abstract** : A **reconfigurable toy** which may be used to selectively simulate a **robotic humanoid figure** and a **rock configuration**. The toy has head, arm and leg members which are **rotatably coupled** to a torso member. Each member has a **surface simulating** part of a **robotic humanoid figure** and a **rock simulating** surface. The toy may be folded into a compact configuration by rotating the members until they are positioned so that only the **rock simulating surfaces** of the members are visible and a **rock configuration** is simulated such as a boulder resting on a supporting surface. The toy may then be unfolded or reconfigured to simulate a **robotic humanoid** figure supported on two leg members.

bigram	Frequency	# patents	weight	relative overall weight (JP vs US)
robotic_humanoid	33	14	0.269	8.432
humanoid_figure	30	17	0.214	2.525
rock_configuration	17	1	0.114	0.0
mating_edges	24	236	0.073	0.115
surface_simulating	17	22	0.065	0.0
reconfigurable_toy	13	15	0.041	0.0
head_retainer	13	15	0.041	0.0
hollow_torso	12	15	0.035	0.0
semicircular_apertures	11	13	0.030	0.0
rock_simulating	9	3	0.026	Inf
arm_retainer	10	19	0.023	NA
rotatably_coupled	12	907	0.012	0.2919
simulating_surface	6	6	0.010	0.0

Table of Contents

Trade (esp. Imports) are associated with proximity of innovations, US industry x Country level

Specification: $Proximity_{o,d,t} = \delta_o + \delta_d + \delta_t + log(imports_{o,d,t}) + \epsilon_{o,d,t}$

	$Proximity_{o,d,t}$		
$\log(\text{imports}_{o,d,t})$	0.062***		0.062***
	(0.006)		(0.006)
$\log(\text{exports}_{o,d,t})$		0.025***	0.013
		(0.009)	(0.008)
R^2	0.265	0.168	0.268
N	4,019	3,584	3,474
Time FE	Y	Y	Y
Country FE	Y	Y	Y
Industry FE	Y	Y	Y

Standard errors are robust.

Immigration is associated with proximity of inventions, US County x Country Level

- Specification: $Proximity_{o,d,t} = \delta_o + \delta_d + \delta_t + log(immigrants^{flow}_{o,d,t}) + \epsilon_{o,d,t}$

	$Proximity_{o,d,t}$ (using text)		(using citations)
	(1)	(2)	(3)
$log(immigrants_{o,d,t})x0.01$	–0.002***	0.001***	–0.007
	(0.000)	(0.000)	(0.005)
R^2	0.090	0.254	0.275
N	113,613	113,613	113,613
Distance	N	Y	Y
Time FE	Y	Y	Y
County FE	Y	Y	Y
Country FE	Y	Y	Y

Standard Errors are clustered by state.

Firms with shareholders in foreign countries file patents with higher foreign proximity, US Firm x Country Level

	$Proximity_{f,t,c}$			
$shareholding_{i,c,t}$	0.249***	0.256***	0.313***	0.325***
	(0.007)	(0.007)	(0.029)	(0.029)
R^2	0.004	0.006	0.026	0.101
N	190,564	190,564	190,564	190,564
Time FE	N	Y	Y	Y
Country FE	N	N	Y	Y
Firm FE	N	N	N	Y

Table of Contents

More foreign proximate patents are more valuable

Specification: $log(Valuation_{p_t}) = \alpha_0 + \beta \sum_j Proximity_{i,p_t} \; T_{j,t} + \chi_p + \epsilon_{i,t}$

	$log(Valuation_p)$		
$log(Proximity_p^{non-US})$	0.018***	0.047***	
	(0.004)	(0.003)	
$log(Similarity_p^{non-US})$			0.096***
			(0.005)
$log(TotalSimilarity_p)$			–0.103***
			(0.006)
R^2	0.238	0.688	0.688
N	1,176,564	1,201,660	1,201,660
Year FE	Y	Y	Y
Class FE	Y	Y	Y
Industry FE	Y	Y	Y
Firm FE	N	Y	Y

Standard Errors are robust.

Firms with higher foreign proximity innovations are more productive

- Specification:

$$log(TFP_{i,t}) = \delta_t + \delta_i + \beta \sum_j Proximity_{i,j,t} \ T_{j,t} + \epsilon_{i,t}$$

- Long Differences (as in Bloom et al. (2015))

$$\Delta^5 log(TFP_{i,t}) = \Delta^5 \delta_t + \beta \Delta^5 \sum_j Proximity_{i,j,t} \ T_{j,t} + \Delta^5 \epsilon_{i,t}$$

	$\Delta log(TFP_{i,t})$		
$\Delta Proximity_{i,t+2}$	0.082***	0.077***	
	(0.023)	(0.024)	
$\Delta Similarity^{nonUS}_{i,t+2}$			0.083**
			(0.039)
$\Delta TotalSimilarity_{i,t+2}$			–0.092**
			(0.046)
R^2	0.020	0.020	0.019
N	17,086	16,508	17,086

Variance Decomposition: Large chunk of variation at the firm level

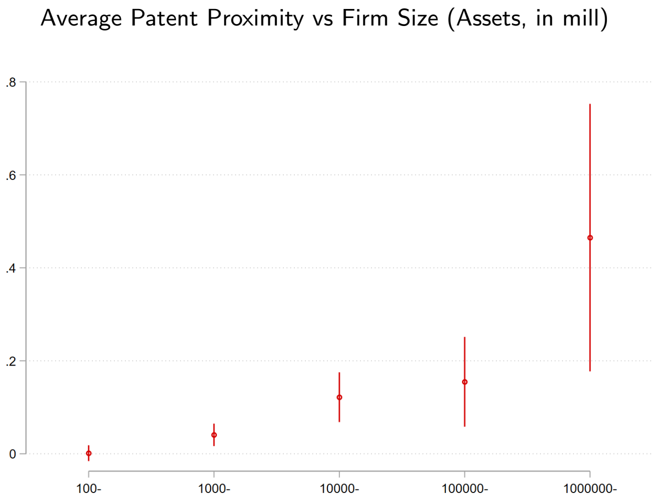
Table: Variance Decomposition

Fixed Effects	R^2
Country x industry x time	29%
Firm + Country x industry x time	34%
Firm x year + Country x industry x time	41%
Firm x Country + Country x industry x time	64%

Industry is NAICS4.

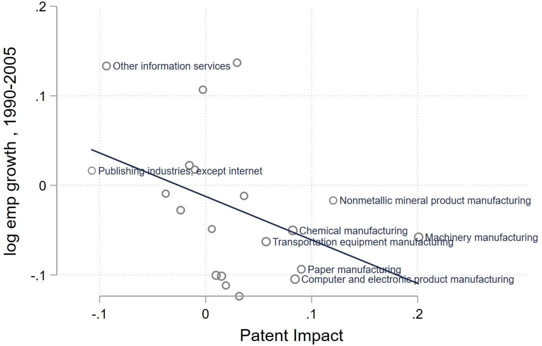
- About one-fourth of the variation at firm x country level.

Large firms file patents more proximate to foreign innovation



Direction of Technology Flows: Industries with innovation closer to Japanese patents lost employment

- Japanese patent proximity at the industry level: Avg. $Proximity_{i,JP,1980s}$



Industries with at least 500 patents in the sample.

Industry Exposure to Labor market impact

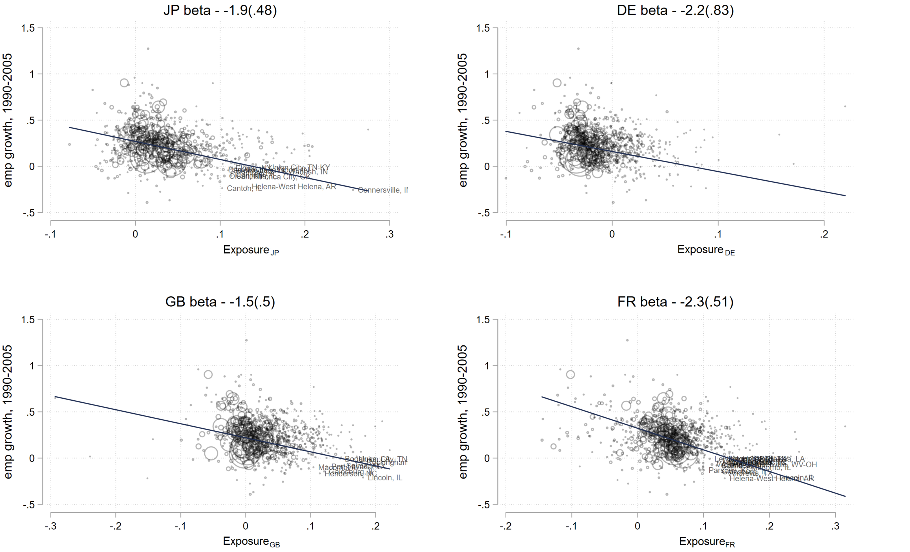
- Regression specification - Bartik Shift Share:

$$\Delta log(Emp)_{i,90-05} = \beta_0 + \beta_c \sum_j ShareEmp_{j,i,1990} * \Delta Proximity_{j,c} + \epsilon_{i,90-05}$$

  i is industry, j is a CBSA (urban area), c is country.

- Identification assumption: $Proximity_{j,c} \perp \epsilon_{i,90-05}$.
  Lead-lag textual relationships between foreign patents and US industries are uncorrelated with local level shocks in employment.

CBSAs exposed to innovation from ageing economies lost employment



Conclusion + In Progress

- A text-based measure of knowledge diffusion: variation in similarity across countries.
- Measure of knowledge diffusion correlated with trade, immigration and FDI.
- Firms with higher diffusion from advanced countries are more productive and file more valuable patents.
- Large variation at the firm level and evidence of firm-level frictions.
- WIP: A quantified model of diffusion to quantify the level of convergence across and within countries.