Heterogeneity in knowledge flows across regions: Investigating patterns and mechanisms

Ashwin Iyenggar and Sai Yayavaram

February 23, 2017

Abstract

Using patent citations as measures of knowledge flows, we explore if the different types of knowledge flows in a region affects the quality of the inventions originating in that region. Using a consistent definition of urban centers for all locations worldwide, we find that local knowledge flows do not seem to impact inventive quality.

Introduction

Scholars of economics and strategy have for long recognized that clusters and agglomeration economies play an important role in fostering innovation (Marshall, 2009; Porter, 1990). Agglomeration economies arise due to labour pooling advantages, economies of specialization of local suppliers, or knowledge spillovers (Krugman, 1991). Patent citations have traditionally been used as a measure of knowledge flows (Jaffe et al., 1993) although recent work raises questions about the efficacy of using patent citations as a measure of knowledge flows (Arora et al., 2017).

Several studies have, however used patent citations to demonstrate that knowledge flows are localized (Alcácer and Gittelman, 2006; Almeida and Kogut, 1999; Jaffe et al., 1993). Regions, however vary in their inventive output (Agrawal et al., 2014). Prior theory does not guide us about how the nature of knowledge flows in a region may affect the quality of the innovations generated in the region. In the current article we use patent citation data for patents applied for between 2001 and 2012 to empirically estimate this relationship.

The rest of the paper is organized as follows. The next section defines a framework to classify knowledge flows in a region and motivates our work by demonstrating how different regions fare differently. We then describe our data and methods in the next section. Our preliminary results are then presented, followed by a discussion of the results. We conclude with next steps and open questions.

Effects of knowledge flows on quality of inventions

We categorize all knowledge flows along two dimensions: first, whether the knowledge flows among inventors are geographically localized or not, and second, whether the knowledge flows are within the boundary of the firm or not. This classification allows us to see knowledge flows

	Same Region	Different Region
Same Assignee	Independent Research Center	Geographic Diversification
Different Assignee	Cluster	Diffusion

Figure 1: Categorizing inventions by matching inventor regions and assignees

in four mutually exclusive, but collectively exhaustive categories as illustrated in Figure 1.

The top left quadrant, labelled an "Independent Research Center" captures those that reflect competence building. Since the knowledge flows both locally and within the firm, this may be seen to represent local search on two dimensions (within firm and within region). Figure 2 depicts the normalized knowledge flows for this category across time for five regions.

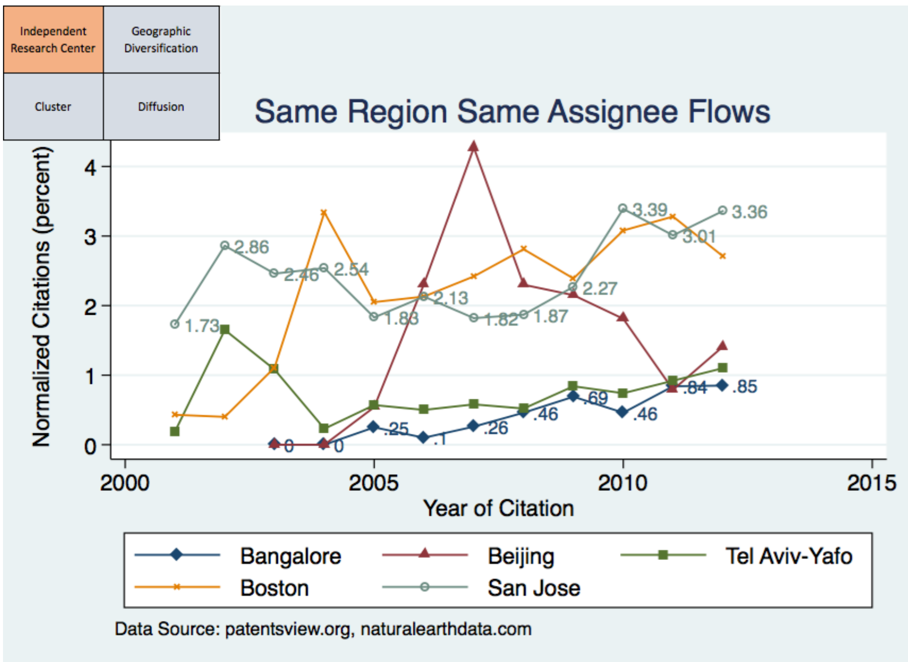

Figure 2: Flows within regions and within assignees

The quadrant on the bottom left, labelled "Cluster" captures knowledge spillovers. Here firms may be seen as performing local search on one dimension (within regions) but not the other (within firms). Figure 3 depicts the normalized knowledge flows for this category across time for five regions.

The quadrant on the top right, labelled as "Geographic Diversification" captures local search

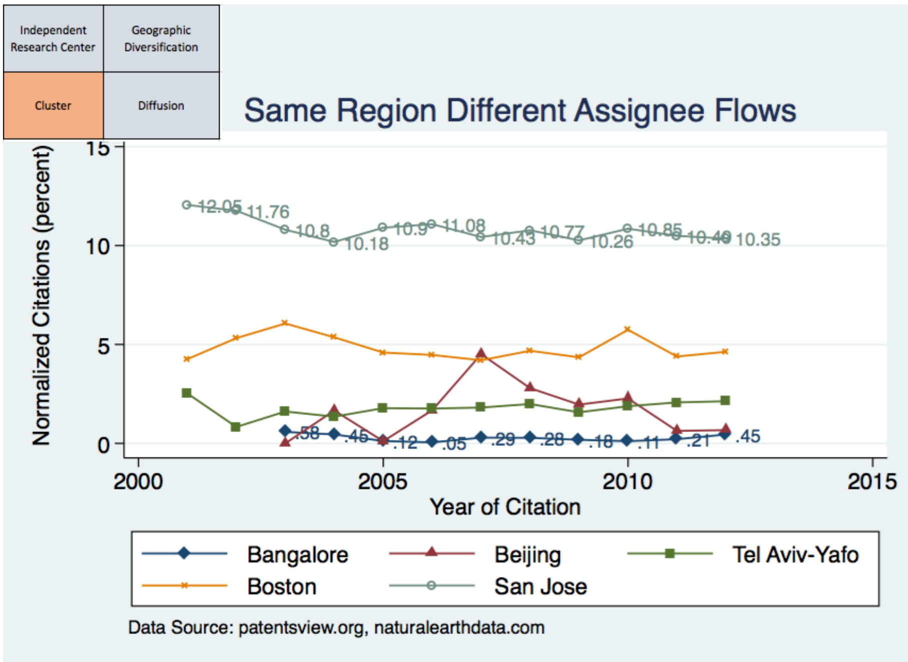

Figure 3: Flows within regions and across assignees

on the dimension of the firm (across geographies) but not across regions. Figure 4 depicts the normalized knowledge flows for this category across time for five regions.

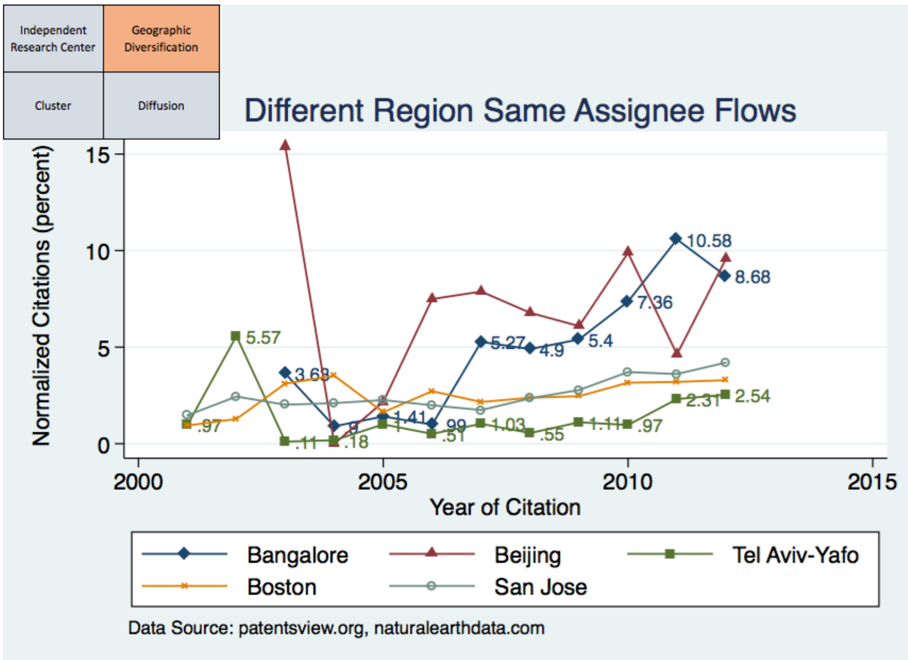

Figure 4: Flows across regions and within assignees

Finally, the bottom right quadrant labelled "Diffusion" captures high exploration and the development of a global pipeline. Figure 5 depicts the normalized knowledge flows for this category across time for five regions.

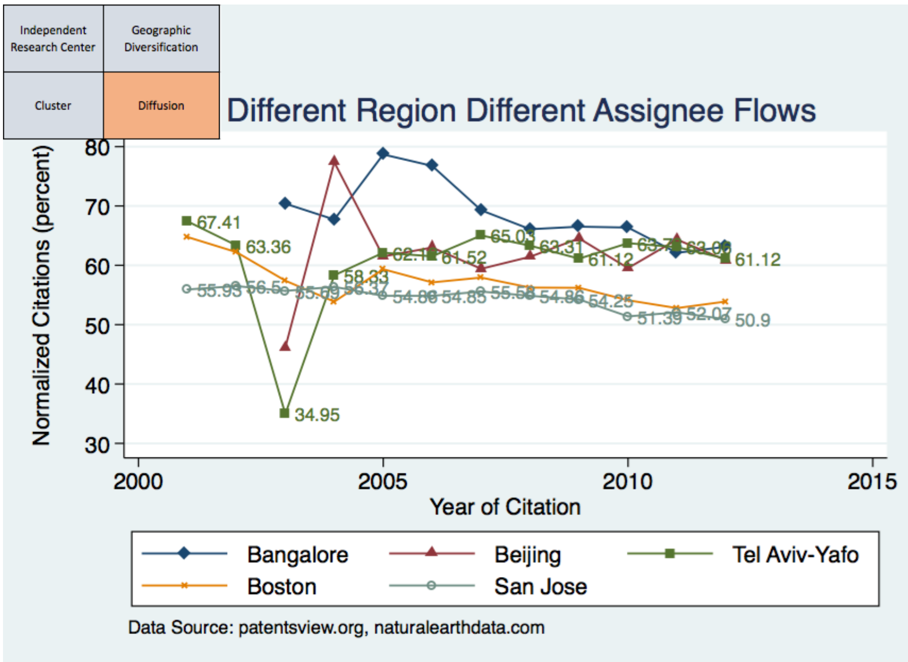

Figure 5: Flows across regions and across assignees

Within the context of this framework, we ask what is the net effect of each type of knowledge flow, and which type of flow will have the largest effect. Since theory does not provide us with an answer, we draw on the patent citations data to provides us some direction.

Data and Measures

We use patent citations data from the USPTO as provided by patentsview.org. Additionally to map inventors to regions, we use urban centers data for world wide locations from Natural Earth Data. While it has been common to use Metropolitan Statistical Areas (MSA) for analyses related to economic geography in the United States, an equivalent measure is unavailable for the rest of the world. For comparability and consistency, we choose to use the urban centers definitions from Natural Earth Data for all regions.

Our unit of analysis is the region-year. In order that we are consistent with our objective of measuring only applicant cited patents as flows of knowledge, we restrict ourselves to those patents categorized as 'cited by applicant'. This decision has the additional effect of limiting our period of analysis to citing patents applied for after the year 2000. We restrict our sample to patents applied for between 2001 - 2012, but citations received till 2015.

Our dependent variable is the total count of self citations received (till 2015) by patents belonging to region-year. Self citations are patent citations where the assignee id on both the citing patent and the cited patent are identical. Our independent variables are the share of citations made to each of the four categories in our defined framework: those to a) Same Region, Same Assignee, b) Same Region, Different Assignee, c) Different Region, Same Assignee, and d) Different Region, Same Assignee.

We control for Log (citations made total), Log (number of patents in Region-Year), Log (patent pool size in Region-Year), percentage of patents in region-year in each technology subcategory defined in Hall et al. (2001)

Results

The preliminary results from our analysis are presented in Table 1. Model 1 reports the results for all regions worldwide, while Model 2 and Model 3 report results for US locations and non-US locations respectively. We find across the three models, that local knowledge flows do not seem to have a significant impact on the quality of inventions as measured by the number of self-citations received.

Discussion

Our preliminary findings in Table 1 suggest no evidence that local knowledge spillovers lead to higher quality inventions. On the other hand, building on knowledge from outside the firm and outside the region seems to improves invention quality

Limitations

While has been popular in the literature, as Arora et al. (2017) point out, using patent citations as a measure knowledge flows may be subject to error. Our definition of regions is dependent on the latitude/longitude assignment in the patentsview.org data and on the urban centers definition in the Natural Earth Data. Any systematic biases in the definition of regions can create biases in measures of within and outside region knowledge flows.

Conclusion

While still at a preliminary stage, our analysis seem to suggest that local knowledge flows as measured by patent citations may not have a significant effect on the quality of innovation produced in a region. A potential extension of the study could be to analyze at level of firm-year rather than region-year. We could additionally look at the additional dimension of technology (within and outside technological domain) in addition to those of within/outside region and within/outside firm. This may provide us a more nuanced understanding of the factors affecting innovation quality. Future studies could potentially examine other measures of invention outcomes such as breakthrough inventions. Finally, while our work suggests that local knowledge spillovers do not effect invention quality, it is not quite as clear why this may be the case and the mechanisms that underlie such an effect. We hope that our current work spurs further research in this direction.

Table 1: Effect of Nature of Citations Made on Self Citations Received*

	(1) Self Citations Received	(2) Self Citations Received	(3) Self Citations Received
Share Citations Made[Same Region, Same Assignee]	-0.0999	-0.324	0.175
	(0.733)	(0.475)	(0.651)
Share Citations Made[Same Region, Different Assignee]	-1.120	-1.718	-0.657
	(0.000)	(0.001)	(0.073)
Share Citations Made[Different Region, Same Assignee]	0.592	0.255	0.834
	(0.000)	(0.328)	(0.000)
Share Citations Made[Different Region, Different Assignee]	0.0226	-0.00949	0.0546
	(0.737)	(0.916)	(0.578)
Log(Total Citations Made)	0.0402	-0.00108	0.0537
	(0.000)	(0.926)	(0.000)
Log (Num Patents)	0.692	0.881	0.696
	(0.000)	(0.000)	(0.000)
Log (Patent Pool Size)	-0.0113	-0.187	-0.0109
	(0.719)	(0.001)	(0.796)
Constant	-2.745	-2.557	-2.637
	(0.000)	(0.000)	(0.000)
Observations	7979	3557	4422
Groups	1003	441	562
Sample	All Locations (UC)	US Locations (UC)	Non-US Locations (UC)

p-values in parentheses

* Results reported are from a priliminary analysis

All models include region fixed effects, year dummies and technology subcategory controls

UC - Urban Centers definition obtained from naturalearthdata.com

References

Agrawal, A., Cockburn, I., Galasso, A., and Oettl, A. (2014). Why are some regions more innovative than others? the role of small firms in the presence of large labs. *Journal of Urban Economics*, 81:149 – 165.

Alcácer, J. and Gittelman, M. (2006). Patent citations as a measure of knowledge flows: The influence of examiner citations. *Review of Economics and Statistics*, 88(4):774–779.

Almeida, P. and Kogut, B. (1999). Localization of knowledge and the mobility of engineers in regional networks. *Management Science*, 45(7):905–917.

Arora, A., Belenzon, S., and Lee, H. (2017). Reversed citations and the localization of knowledge spillovers. Working Paper 23036, National Bureau of Economic Research.

Hall, B. H., Jaffe, A. B., and Trajtenberg, M. (2001). The nber patent citation data file: Lessons, insights and methodological tools. Working Paper 8498, National Bureau of Economic Research.

Jaffe, A. B., Trajtenberg, M., and Henderson, R. (1993). Geographic localization of knowledge spillovers as evidenced by patent citations. *The Quarterly Journal of Economics*, 108(3):577–598.

Krugman, P. R. (1991). *Geography and trade*. MIT press.

Marshall, A. (2009). *Principles of Economics: Unabridged Eighth Edition*. Cosimo, Inc.

Porter, M. E. (1990). *The Competitve Advantage of Nations*. Free Press New York.