Heterogeneity in Knowledge Flows of Regions: Impact on Invention Quality

ABSTRACT

Using patent citations as measures of knowledge flows, we explore if the different types of knowledge flows in a region affect the quality of inventions originating in that region. We leverage a database of worldwide urban centers obtained from remote sensing data to provide a consistent definition for regions worldwide. As some other recent research in the area (**?**), our preliminary results find inconclusive evidence for a clear positive effect of local knowledge flows on invention output of regions, or the globalization of R&D on the invention output of regions. Our work here suggests that much more research is required to distill any stylized facts on the implications of geography and firm organization on invention outcomes.

Draft

Results presented here are still at a preliminary stage. Kindly do not circulate.

Heterogeneity in Knowledge Flows of Regions: Impact on Invention Quality

Empirical studies in the literature on economic geography have used patent citations to suggest that knowledge spillovers are geographically localized (**??????**), and that innovation is more spatially concentrated that production (**?**). On the other hand, patent citations have also been used in the international business literature to demonstrate that multi-national firms (MNCs) gain from the cross-border flow of knowledge and the globalization of R&D (**???**). This leads to the question of what indeed is the net effect of geography on innovation outcomes of firms and regions. Firms seek to improve innovation outcomes as a way of building a sustainable competitive advantage. Understanding how prior knowledge flows (across firms and geographies) affect innovation outcomes would therefore be valuable to managers in their quest for superior performance.

In this study, we examine how the nature of knowledge spillovers or flows in a region affect the quality of the inventions generated in the region. Specifically, we look at knowledge flows spanning region boundaries and firm boundaries, and ask if and to what extent prior knowledge flows between and within regions and firms affects innovation quality of patents originating in a region. In this study we use citation data of patents applied for between 2001 and 2012 to empirically estimate the relationship between knowledge flows of a region and quality of inventions from that region. We are aware of no prior studies that have examined the effects of the geographical characteristics of knowledge flows on invention outcomes of regions. Our work therefore seeks to contribute to the innovation literatures in both economic geography and international business.

Correction The next section reviews the theory from the geography of innovation and innovation in international business. We then define a framework to classify knowledge flows in a region and motivate our work by demonstrating how different regions fare differently in terms of knowledge flows. We build on the theory and this framework to propose our hypotheses. We then propose our research design, and follow it up with a description of our data and methods. Our preliminary results are then presented, followed by a discussion of the results. We conclude with limitations, next steps and open questions for further research.

Draft

THEORY AND HYPOTHESES

? drew our attention toward two types of knowledge spillovers - "rent spillovers" that are rival and excludable knowledge flows that occur from the exchange of goods, and "pure or idea-creating spillovers" that are non-rival and non-excludable knowledge that is typically associated with R&D activity. While rent spillovers may be characterized as a private good, pure spillovers exhibit characteristics of a public good (**?**). This dual nature of knowledge is additionally complicated by the fact that knowledge flows may be largely invisible (**?**), and only sometimes do knowledge flows leave a paper trail in the form of patent citations (**?**). In the following sections, we review the literature in economic geography and international business **Correction** cognizant of these fundamental limitations relating to the nature and measurement of knowledge.

Economic geography of innovation

Correction - What is the purpose of this section Scholars of economics and strategy have for long recognized that clusters and agglomeration economies play an important role in fostering innovation (**??**). Agglomeration economies may arise due to labor pooling advantages, economies of specialization of local suppliers, and knowledge spillovers (**??**). Several studies have used patent citations to demonstrate that knowledge flows are localized (**??**). After controlling for the geographic distribution of production, **?** demonstrated that innovation exhibits a pronounced tendency to cluster spatially. **?** found that new product introductions were more geographically concentrated than patents and that they derived important inputs from universities and industrial R&D.

Regions, however vary in their inventive output (**?**) and the nature of knowledge flows in a region may be only one source of this variation within regions. Location however has been suggested to matter more at the earliest stage of the industry life cycle, with life-cycle production costs featuring more prominently in mature goods and industries. **?** suggest that early stages of industry life cycles are characterized by the importance of tacit knowledge, and may hence exhibit higher a propensity to cluster to benefit from knowledge spillovers.

Draft

Multiple reasons have been attributed for the localized nature of knowledge spillovers. First, costs of collaboration are lower due to geographic proximity. Second, proximity has been argued to promote serendipitous and chance encounters that create the environment for new solutions. Third, diversity within a region has also been suggested to benefit innovative output (**?**). Scholars have suggested that related variety (**??**) may benefit in the generation of new ideas and that diversity between industrial activities may benefit the transfer of ideas (**?**). Additionally, scholars have suggested that the tacit nature of knowledge, institutions and regional innovation systems may all contribute to the localization of knowledge flows (**????**).

Correction - Summarize, the reader is left hanging

Globalization of innovation

Regional innovation systems have also been suggested as influencing MNC decisions to locate subsidiaries (**?**). However, scholars have also considered organizational capabilities (**?**), property rights, technological sourcing (**?**), higher product ownership and interdependence (**?**), and access to highly skilled human capital. On the other hand, political borders have been found to constrain the flows of knowledge (**?**), but mobility of inventors helps firms improve innovation outcomes (**?**).

Correction - Literature review is not sufficient. What is the point that you are making?

Inter-firm knowledge flows and geography

The review of the literature in the economic geography and international business traditions suggests both positive and negative effects of localization of knowledge spillovers on innovation outcomes. One the one hand agglomerations may seem to benefit firms through access to tacit knowledge, while on the other hand it may also increase the number of competitors who are specialized in the same market segment (**?**). MNC subsidiary location choices seem to be affected by local factors, as well as relative ties with local firms and MNC affiliates(**?**). **Correction - Write more**. In the following section, we present a simple framework with which to address this phenomenon.

Draft

Effects of knowledge flows on quality of inventions

We formalize the discussions in the previous sections by categorizing all knowledge flows incorporated into an invention along two dimensions: first, whether the knowledge flows among inventors are local to a geographical region or not, and second, whether knowledge flows are within the boundary of the firm or not. The motivation behind defining such a structure is to both keep it simple, as well as to focus on aspects that firms and governments may have an ability to influence. Given the nuanced and subtle nature of knowledge flows, and the difficulty in measuring them, firms may be better served with simple and actionable innovation strategies (**?**).

Our classification allows us to analyze knowledge flows in four mutually exclusive, but collectively exhaustive categories as illustrated in Figure 1. We next describe each quadrant and discuss how the category of knowledge flow in that quadrant can affect invention quality. Within the context of this framework, we ask what is the net effect of each category of knowledge flow on invention quality, and which category of flow has the largest effect on invention quality.

Insert FIGURE 1 about here.

The top left quadrant, labelled an "Independent Research Center" captures those knowledge flows that reflect competence building. Since these knowledge flows are both within the region and within the firm, these flows represent local search on two dimensions (within firm and within region). Thus, while the competence that is being built up by the Independent Research Center can be expected to have a positive effect on invention quality, the localness of the search on both dimensions may have a negative effect on invention quality. Figure 2 depicts the knowledge flows for this category (percentage of backward citations from this region that are to the same firm or assignee and same region) across time for five regions: Bangalore, Beijing, Tel Aviv-Yafo, Boston and San Jose (core of "Silicon Valley"). While our empirical analysis covers all the major regions of the world, we chose these five regions as illustrative examples. We note that both San Jose and Boston report a higher proportion of knowledge flows within the same firm in the same region,

Draft

while Bangalore and Tel Aviv-Yafo have the lowest proportion (fewer than 1%) of their citations from the same firm within the same region.

Insert FIGURE 2 about here.

The quadrant on the bottom left, labelled "Cluster" captures knowledge spillovers within a region. Here firms may be seen as performing local search on one dimension (within regions) but not the other (within firms). Figure 3 depicts the knowledge flows for this category across time for the same five regions. San Jose clearly stands out from the rest, suggesting a higher amount of across firm flows of knowledge in Silicon Valley, a result consistent with several prior studies(**?**).

Insert FIGURE 3 about here.

The quadrant on the top right, labelled as "Geographic Diversification" captures local search on the dimension of the firm (across geographies) but not across regions. Innovations that are built on knowledge from several regions can be expected to benefit from the diversity of knowledge across regions. Yet, as in the previous quadrant, there is localness along the dimension of firm and such localness can have a negative effect on invention quality (**?**). Figure 4 depicts the knowledge flows for this category across time for the five regions. We note that Bangalore and Beijing have a relatively higher proportion of knowledge flows from same assignees in different locations, thus confirming the role of these regions as R&D outposts of multinational firms.

Insert FIGURE 4 about here.

Finally, the bottom right quadrant labelled "Diffusion" captures high exploration along both dimensions, indicating the development of a global pipeline (**?**). Figure 5 depicts the knowledge flows for this category across time for the five regions. We note that Bangalore, Beijing and Tel

Draft

Aviv-Yafo have a higher level of knowledge flows from other firms in other regions compared to Boston and San Jose, which is to be expected given that the absolute level of innovative activity in these emerging hotspots is still lower compared to that in Boston and San Jose.

Insert FIGURE 5 about here.

As can be seen from the preceding discussion, prior theory suggests both positive and negative effects for each of these four categories of knowledge flows and it is not clear what the net effect will be on invention quality. This also suggests that prior theory does not provide guidance on which category of knowledge flows will have the highest effect on invention quality. Since theory does not provide us with an answer, we rely on empirical analysis to inform us on the net effect of each category of knowledge flow on invention quality and which category has the highest effect on invention quality.

DATA AND METHODS

We use patent citations data from the U.S. Patent Office (USPTO) as provided by patentsview.org. The USPTO provides information on assignee, the location of inventors, and year of patent application. We use this information to identify which patents belong to which region in a given year. Since there can be more than one inventor on a patent, a patent can belong to more than one region. We compare the assignee and inventor location of each patent with each assignee and inventor of each backward citation to identify the category of knowledge flow (e.g., same assignee, different location) indicated by the backward citation. Additionally, to map location data of inventors from USPTO to regions, we use urban centers data for worldwide locations from Natural Earth Data that uses remote sensing data to determine urban agglomerations (a process developed in **?**). While it has been common practice to use Metropolitan Statistical Areas (MSA) for analyses related to economic geography in the U.S., an equivalent measure is unavailable for the rest of the world. For

Draft

comparability and consistency, we choose to use the urban centers definitions from Natural Earth Data for all regions both within U.S. and outside U.S.

Our unit of analysis is the region-year.

Applicant citations and Examiner Citations

Much of the prior work analyzing patent citations had pooled all citations made irrespective of the source (i.e., cited by applicant, or cited by examiner, or cited by other) due to lack of systematic classification of the citation source in the USPTO database. However, systematic classification of patent citations is available from 2001 onwards. Scholars have even called for more granular analysis based on citation source. For example, **?** suggest that inferences about inventor knowledge using pooled citations may suffer from bias or overinflated significance levels. On the other hand scholars have argued that patents cited by the examiner may not represent knowledge flows at all. To be consistent with our objective of measuring knowledge flows without strong assumptions, we do not restrict ourselves to those citations categorized as 'cited by applicant' or those categorized as 'cited by examiner' . We use three data sets - a) the pooled data set of all applicant and examiner citations (this is our primary data set), b) the data set of only examiner citations, and c) the data set of only applicant citations. This decision has the additional effect of limiting our period of analysis to citing patents applied for after the year 2000 since the data on which citations were added by examiners is available for patents from only 2001 onwards.

Citations Received

Correction - Revert this to counting all citations to date, and justify it We restrict our sample to patents applied for between 2001 - 2012, but citations received till 2015. Our default strategy for measuring patent citations is to count citations received until and including the fifth annual anniversary of the patent application. This has the effect of biasing against patent applications from 2011 and 2012. This problem may be resolved by updating the dataset to the version updated to March 2017. For paucity of time, the dataset was not updated in time for the submission of this

Draft

paper.

Dependent Variable

Correction - Revert this to counting all citations to date, get rid of fifth annual anniversary
Our primary dependent variable is the total count of citations by patents belonging to a region-year. However, for all models we also report results for total non-self citations received as the dependent variable. Citations are counted as those received up to and including the fifth annual anniversary of the patent filing or until 2015, whichever is earlier. In the section on additional results, we also provide results with an alternative measure of citations received with the dataset on applicant only citations. Here we count all citations received till 2015 rather than truncate after the fifth annual anniversary of patent application filing date.

Independent Variables

Building on the framework depicted in Figure 1, our independent variables are the percentage of backward citations made to each of the four categories: those to a) same region, same assignee, b) same region, different assignee, c) different region, same assignee, and d) different region, same assignee.

Percentage of backward citations made to same region and same assignee We compute this variable as follows. For each region, for each year, we compute the ratio of the total number of backward citations made to same region and same assignee in that year to the total number of backward citations made in that year.

Percentage of backward citations made to same region and different assignee We compute this variable as follows. For each region, for each year, we compute the ratio of the total number of backward citations made to same region and different assignee in that year to the total number of backward citations made in that year.

Draft

Percentage of backward citations made to different region and same assignee We compute this variable as follows. For each region, for each year, we compute the ratio of the total number of backward citations made to different region and same assignee in that year to the total number of backward citations made in that year.

Percentage of backward citations made to different region and different assignee We compute this variable as follows. For each region, for each year, we compute the ratio of the total number of backward citations made to different region and different assignee in that year to the total number of backward citations made in that year.

For any given region-year, we would therefore have the sum of each of the four independent variables to add to 1 (one). **Correction - Percentage and fraction, both terminologies are used. Stick to one**

Control Variables

We control for the total number of citations made in the region-year, the total number of patents in the region-year, the size of the patent pool in the region-year, as well as the percentage of patents in region-year in each technology subcategory as defined by **?**. The patent pool is the total number of patents that belong to a region up to the previous year. It is important to control for the size of the patent pool as regions that have a larger patent pool (such as San Jose) will have more patents that can be cited and can therefore have larger within region spillovers as compared to a region that has only a small patent pool. We include region fixed effects and year dummies in all regression models so as to control for region level and year specific effects, if any. Since our dependent variable is a count variable, we use negative binomial regression analysis with fixed effects.

Table **??** displays the correlation of our primary variables in the dataset of applicant and examiner citations. Summary statistics for the variables in the dataset of applicant and examiner citations is provided in Table **??**. Similar statistics for the applicant only citations dataset and the examiner only citations dataset have been provided in the section on additional results.

Draft

Insert TABLE 1 about here.

Insert TABLE 2 about here.

Insert TABLE 3 about here.

Insert TABLE 4 about here.

Draft

RESULTS

The preliminary results from our analysis for the dataset of applicant and examiner cited patents are presented in Table **??**. Models 1-3 report results with the dependent variable as the total number of citations received for all regions worldwide, U.S. locations and non-U.S. locations respectively. Models 4-6 report results with the dependent variable as the number of non-self citations received. The initial observation is that results seem to vary depending on the chosen dependent variable, though the signs of the coefficients are the same. Across the models we find that knowledge flows under cluster and geographical diversification have a significant impact on the quality of inventions as measured by the number of citations received. However, the differences in the results between the two dependent variables suggests that there is an underlying effect that the current model is not capturing.

Let us set aside the inconsistency in results between the two measures of the dependent variable, and just focus on the results for total citations received. H1 suggested a negative effect for within firm, within region flows on invention quality. The results however suggest a positive effect and significant effect in models 1 and 3. H1 is therefore not supported. H2a and H2b suggested that potentially opposite claims for higher cluster knowledge flows. The results are mixed across

the samples, suggesting perhaps that different effects are at play under different conditions. H3 predicted a positive effect for the globalized diversification driven knowledge flows. The results here are consistent with our hypotheses, across all models and across datasets. Finally, we had predicted a non-effect for the diffusion knowledge flow on invention performance. Again the results for the total citations received dependent variable hold.

Insert TABLE 5 about here.

Additional Results

Table **??** displays the correlation of our primary variables in the dataset of examiner only citations. Summary statistics for the variables in the dataset of examiner only citations is provided in Table **??**.

Insert TABLE 6 about here.

Insert TABLE 7 about here.

Insert TABLE 8 about here.

Insert TABLE 9 about here.

The additional results from our analysis for the dataset of examiner only cited patents are presented in Table 10. Models 1-3 report results with the dependent variable as the total number of citations received for all regions worldwide, U.S. locations and non-U.S. locations respectively.

Draft

Models 4-6 report results with the dependent variable as the number of non-self citations received. The divergence between results for the two outcome measures of total citations received and non-self citations received are similar to that observed in Table **??** on the dataset with applicant and examiner citations.

Insert TABLE 10 about here.

Table 11 displays the correlation of our primary variables in the dataset of applicant only citations. Summary statistics for the variables in the dataset of applicant only citations is provided in Table 12.

Insert TABLE 11 about here.

Insert TABLE 12 about here.

Insert TABLE 13 about here.

Insert TABLE 14 about here.

The additional results from our analysis for the dataset of applicant only cited patents are presented in Table 15. Models 1-3 report results with the dependent variable as the total number of citations received for all regions worldwide, U.S. locations and non-U.S. locations respectively. Models 4-6 report results with the dependent variable as the number of non-self citations received. The divergence between results for the two outcome measures of total citations received and non-self citations received are similar to that observed in Table **??** on the dataset with applicant and examiner citations.

Draft

Insert TABLE 15 about here.

Insert TABLE 16 about here.

Insert TABLE 17 about here.

Insert TABLE 18 about here.

Insert TABLE 19 about here.

Draft

The additional results from our analysis for the dataset of applicant only cited patents are presented in Table 20. Models 1-3 report results with the dependent variable as the total number of citations received for all regions worldwide, U.S. locations and non-U.S. locations respectively. Models 4-6 report results with the dependent variable as the number of non-self citations received. The divergence between results for the two outcome measures of total citations received and non-self citations received are similar to that observed in Table **??** on the dataset with applicant and examiner citations.

Insert TABLE 20 about here.

The additional results from our analysis for the dataset of applicant only cited patents with citations received counted till 2015 are presented in Table 21. Models 1-3 report results with the dependent variable as the total number of citations received for all regions worldwide, U.S. locations and non-U.S. locations respectively. Models 4-6 report results with the dependent variable as

the number of non-self citations received. Interestingly here, we note that results between the two dependent variable measures are not particularly different. However, these set of measures seem to suggest different effects to that observed in Table 15 on the dataset with applicant only citations but where citations were counted until the fifth annual anniversary of the patent filing date. This suggests that we may need to investigate the time effect on citations more closely.

Insert TABLE 21 about here.

DISCUSSION

Knowledge flows are hard to measure, and knowledge is characterized by by opposing qualities of being tacit and coded, and of behaving like a public or a private good. The central challenge therefore seems to be to find a way to disentangle the opposing natures of knowledge. It is possibly this that is behind the different results we observe between the two dependent variable measures of total citations received and non-self citations received. In fact, the 2005 debate in the American Economic Review highlights how minor changes in measures can dramatically alter results when measuring knowledge and its effects (**??**). The question then is if there is an alternative means to measure knowledge flows more reliably, or if the problem lies with the research methodology used herein. **?** have suggested a regional compatibility index for capturing technological linkages. An alternate approach of understanding knowledge flows relies on the movement of people with the assumption that knowledge is embedded within an individual. While patent citations arguably mapped codified knowledge, mobility captures a measure of tacit knowledge (**?**). However even in this tradition, scholars have suggested opposite effects of knowledge flows. While **?** show that mobility patterns of star inventors directly the geographic patterns of knowledge spillovers, **?** illustrate that mobile engineers who join a firm with stronger path dependence are unlikely to build on the knowledge of their previous firms. While geography may be seen as providing a platform to organize these interactions amongst inventors, it seems that we may need to consider other dimensions

Draft

of geography so as to fully understand the effect on invention outcomes (**?**). Universities, research consortia, and other institutions may play a significant role in affecting invention outcomes.

It would seem appropriate to raise two measurement related issues here. First, despite well known issues in the use of patents (**??**) and in the use of patent citations to demonstrate localization of knowledge spillovers (**??**), much work in understanding knowledge spillovers has continued to rely on patent citations. While our analysis in this paper is singularly dependent on data of United States patents, we are careful to not make strong assumptions about patent citations reflecting underlying knowledge flows. Wherever possible, we consider competing hypotheses built on opposing assumptions about what patent citations may capture. Second, innovations are clearly not the same as inventions. The measure of an innovation lies in its acceptance in the marketplace. An invention may therefore only represent an early event in the innovation process. While we recognize that managers and policy makers are interested in innovation outcomes, we continue to be limited by data availability that allows us to only make claims about invention quality. Despite those caveats, pursuing the questions posed above can provide not just interesting trade-offs to explore, but indeed create opportunities for a larger research agenda on actionable strategies for firms seeking to improve their innovation outcomes.

Draft

LIMITATIONS

In addition to the customary disclaimers about using patent citations as measures of knowledge flows, we would like to highlight a couple of limitations that may stem from the methodology we adopted. While the use of patent citations as a measure of knowledge flows has been popular in the literature, this may nevertheless 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 region and across region knowledge flows.

CONCLUSION

While still at a preliminary stage, our analysis seem to suggest that knowledge flows are extremely sensitive to the nature of mesurement deployed. This casts a doubt on the widely accepted idea that local knowledge spillovers are an important source of agglomeration economies. On the other, it also brings us to bear on the important issue of measuring knowledge flows. A potential extension of the study could be to conduct empirical analysis at the level of firm-year rather than at the 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 with a more nuanced understanding of the factors affecting invention quality. Future studies could potentially examine other measures of invention outcomes such as breakthrough inventions. Finally, while our work suggests multiple effects of knowledge flows on invention quality, it is not quite as clear why this may be the case. We have attempted to provide some direction by building from the theory on both streams economic geography, and international business, but there is clearly a lot to be done. We hope that our current work spurs further research in this direction.

Draft

FIGURE 1: Categories of knowledge flows

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

Draft

FIGURE 2: Knowledge flows within regions and within assignees

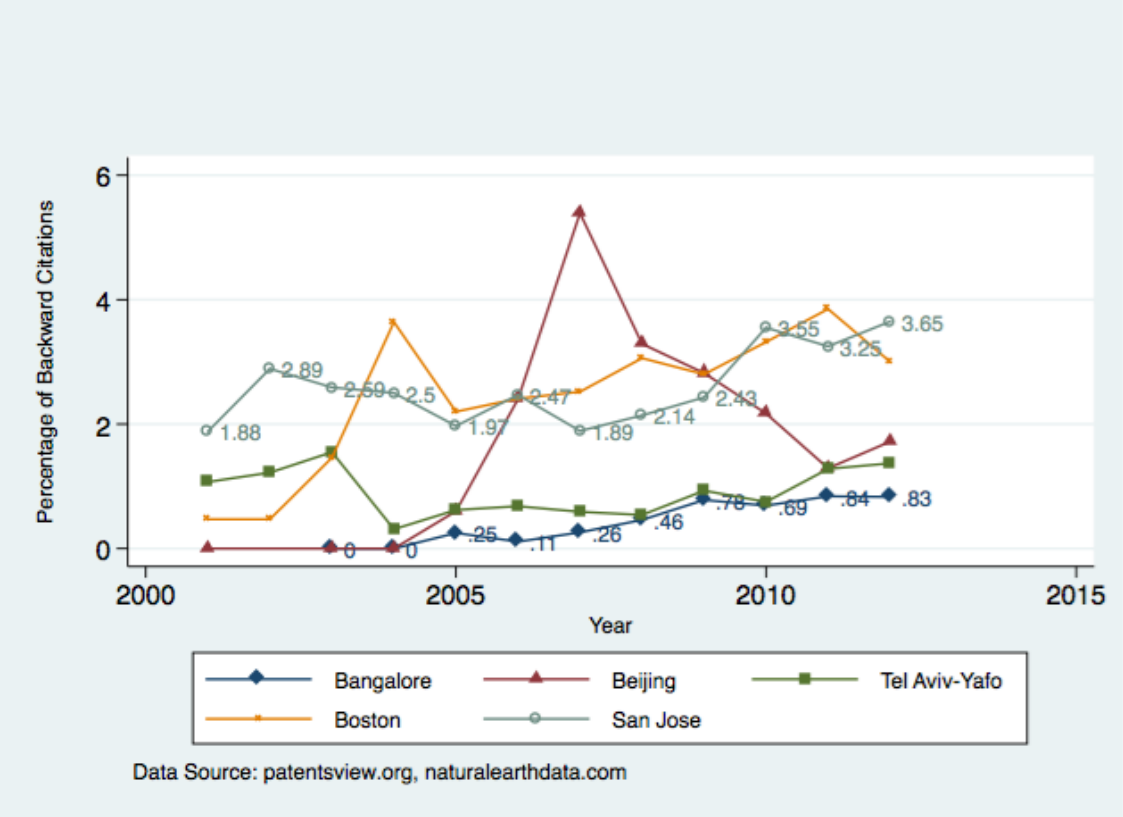

FIGURE 3: Knowledge flows within regions and across assignees

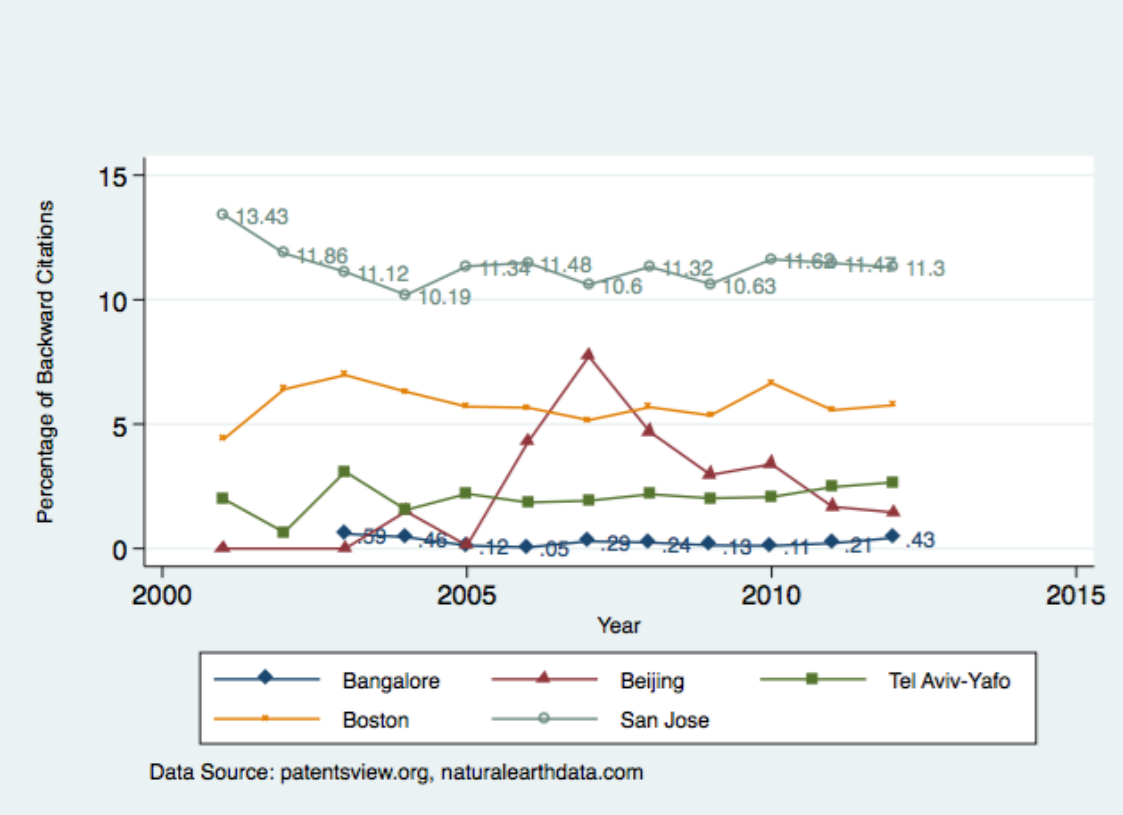

FIGURE 4: Knowledge flows across regions and within assignees

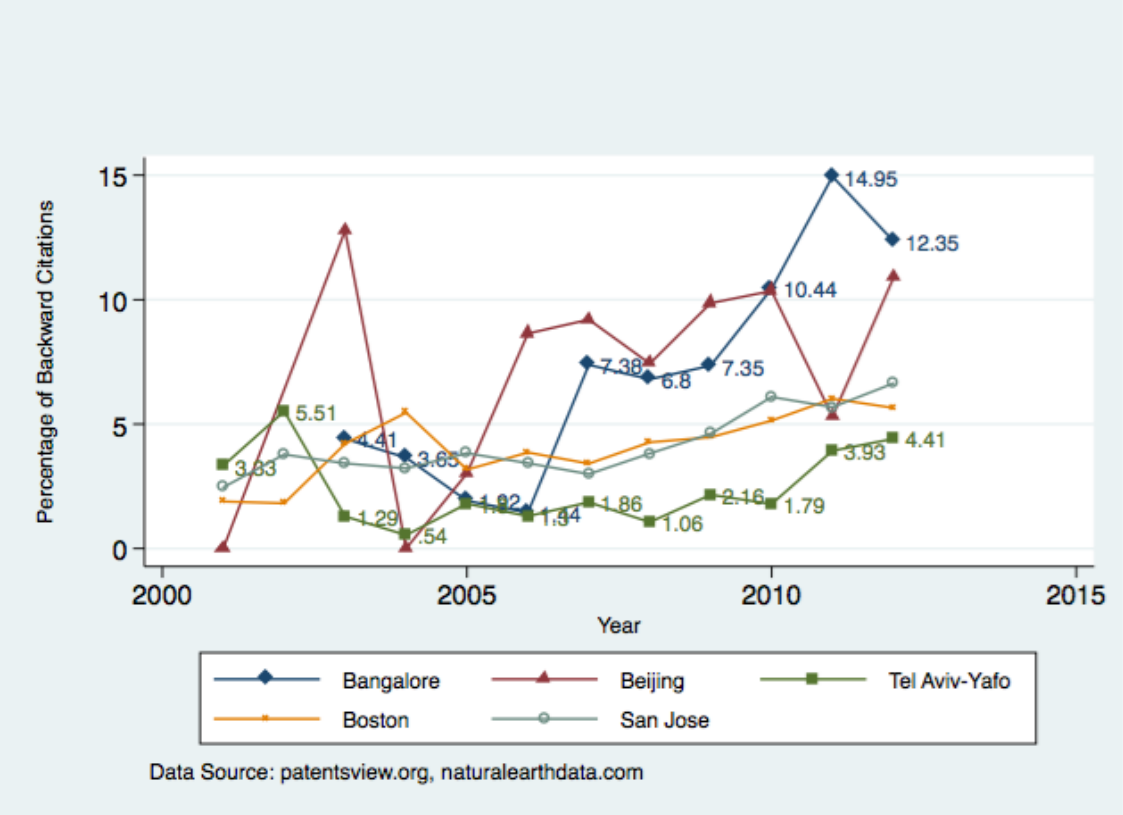

FIGURE 5: Knowledge flows across regions and across assignees

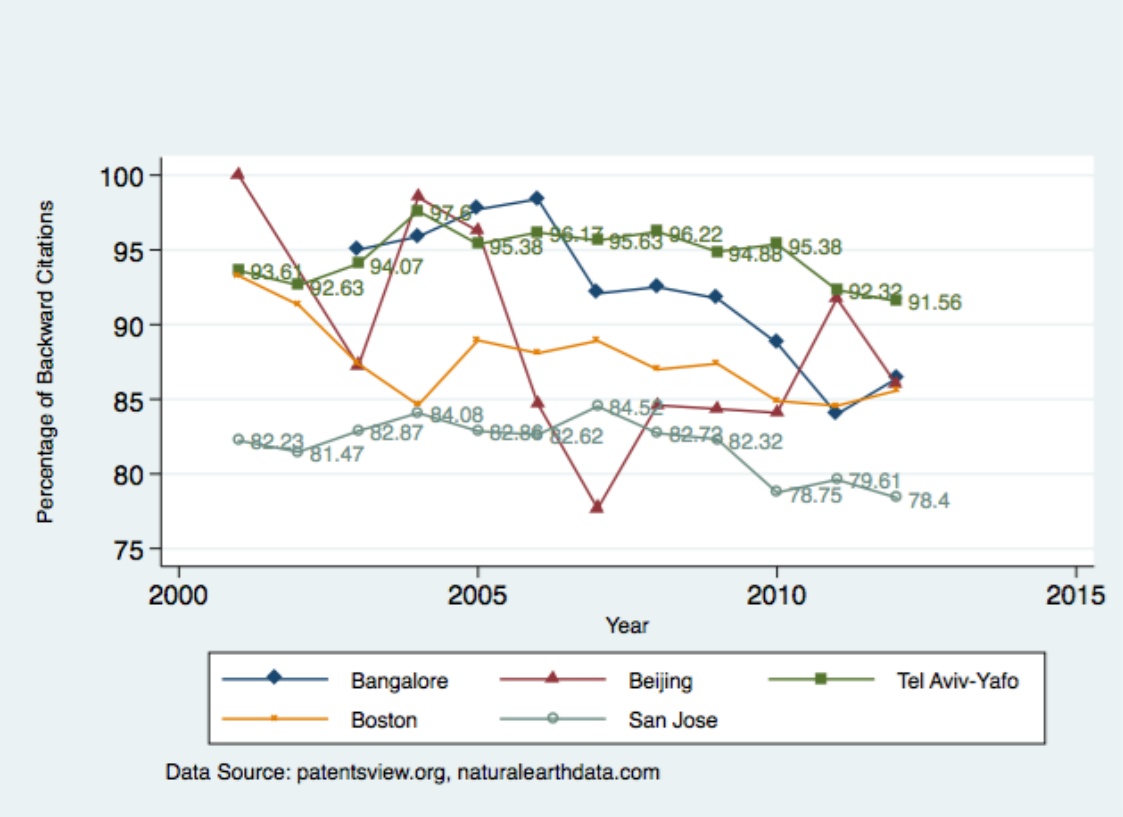

TABLE 1: Correlation table for all citations with DV as Total Citations Received

Variables	1	2	3	4	5	6	7	8	9	10	11	12
1. Total Citations Received	1.00											
2. Non-Self Citations Received	1.00	1.00										
3. Self Citations Received	0.88	0.83	1.00									
4. Share Citations Made[Same Region, Same Assignee]	0.08	0.07	0.10	1.00								
5. Share Citations Made[Same Region, Different Assignee]	0.16	0.15	0.17	0.11	1.00							
6. Share Citations Made[Different Region, Same Assignee]	-0.00	-0.01	0.01	0.18	-0.04	1.00						
7. Share Citations Made[Different Region, Different Assignee]	-0.07	-0.06	-0.09	-0.50	-0.28	-0.89	1.00					
8. Share Citations Made[Same Region]	0.16	0.15	0.18	0.77	0.71	0.10	-0.54	1.00				
9. Share Citations Made[Same Assignee]	0.02	0.02	0.04	0.49	-0.00	0.95	-0.96	0.35	1.00			
10. Log (Total Citations Made)	0.33	0.31	0.37	0.15	0.17	0.07	-0.16	0.22	0.11	1.00		
11. Log (Num Patents)	0.35	0.34	0.38	0.18	0.19	0.05	-0.16	0.25	0.11	0.94	1.00	
12. Log (Patent Pool Size)	0.31	0.30	0.34	0.20	0.23	0.04	-0.16	0.28	0.10	0.88	0.92	1.00

Draft

TABLE 2: Summary statistics for all citations with DV as Total Citations Received

Variable	Mean	Std. Dev.	N
Total Citations Received	1993.201	12786.711	17166
Non-Self Citations Received	1626.983	10883.951	17166
Self Citations Received	366.218	2214.878	17166
Share Citations Made[Same Region, Same Assignee]	0.017	0.038	17166
Share Citations Made[Same Region, Different Assignee]	0.013	0.034	17166
Share Citations Made[Different Region, Same Assignee]	0.068	0.101	17166
Share Citations Made[Different Region, Different Assignee]	0.901	0.119	17166
Share Citations Made[Same Region]	0.03	0.053	17166
Share Citations Made[Same Assignee]	0.086	0.114	17166
Log (Total Citations Made)	5.331	2.463	17166
Log (Num Patents)	2.753	2.04	17166
Log (Patent Pool Size)	5.381	2.286	17166

Draft

TABLE 3: Correlation table for all citations with DV as Non-Self Citations Received

Variables	1	2	3	4	5	6	7	8	9	10	11	12
1. Total Citations Received	1.00											
2. Non-Self Citations Received	1.00	1.00										
3. Self Citations Received	0.88	0.83	1.00									
4. Share Citations Made[Same Region, Same Assignee]	0.08	0.07	0.11	1.00								
5. Share Citations Made[Same Region, Different Assignee]	0.16	0.15	0.17	0.11	1.00							
6. Share Citations Made[Different Region, Same Assignee]	-0.00	-0.01	0.01	0.19	-0.04	1.00						
7. Share Citations Made[Different Region, Different Assignee]	-0.07	-0.06	-0.09	-0.50	-0.29	-0.90	1.00					
8. Share Citations Made[Same Region]	0.16	0.15	0.18	0.77	0.72	0.10	-0.54	1.00				
9. Share Citations Made[Same Assignee]	0.02	0.02	0.04	0.49	-0.00	0.95	-0.96	0.34	1.00			
10. Log (Total Citations Made)	0.33	0.31	0.37	0.16	0.17	0.07	-0.16	0.23	0.12	1.00		
11. Log (Num Patents)	0.35	0.34	0.38	0.19	0.19	0.06	-0.16	0.26	0.11	0.94	1.00	
12. Log (Patent Pool Size)	0.31	0.30	0.34	0.21	0.23	0.04	-0.17	0.29	0.11	0.88	0.92	1.00

Draft

TABLE 4: Summary statistics for all citations with DV as Non-Self Citations Received

Variable	Mean	Std. Dev.	N
Total Citations Received	1998.552	12803.462	17120
Non-Self Citations Received	1631.355	10898.237	17120
Self Citations Received	367.197	2217.771	17120
Share Citations Made[Same Region, Same Assignee]	0.017	0.037	17120
Share Citations Made[Same Region, Different Assignee]	0.013	0.034	17120
Share Citations Made[Different Region, Same Assignee]	0.068	0.1	17120
Share Citations Made[Different Region, Different Assignee]	0.901	0.118	17120
Share Citations Made[Same Region]	0.03	0.053	17120
Share Citations Made[Same Assignee]	0.086	0.113	17120
Log (Total Citations Made)	5.339	2.461	17120
Log (Num Patents)	2.76	2.038	17120
Log (Patent Pool Size)	5.392	2.279	17120

Draft

TABLE 5: NB Regression Analysis of Invention Quality for all Citations

	(1) Total Citations Received	(2) Total Citations Received	(3) Total Citations Received	(4) Non-Self Citations Received	(5) Non-Self Citations Received	(6) Non-Self Citations Received
Share Citations Made[Same Region, Different Assignee]	-0.516	0.0657	-1.308	-0.604	0.189	-1.419
	(0.054)	(0.859)	(0.001)	(0.028)	(0.617)	(0.000)
Share Citations Made[Different Region, Same Assignee]	-0.220	-0.0596	-0.563	-0.559	-0.166	-0.920
	(0.273)	(0.837)	(0.036)	(0.008)	(0.584)	(0.001)
Share Citations Made[Different Region, Different Assignee]	-0.588	-0.149	-1.065	-0.579	0.0762	-1.049
	(0.001)	(0.545)	(0.000)	(0.001)	(0.766)	(0.000)
Log (Total Citations Made)	0.168	0.162	0.174	0.120	0.105	0.138
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
Log (Num Patents)	0.615	0.673	0.668	0.650	0.714	0.696
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
Log (Patent Pool Size)	-0.104	-0.287	-0.113	-0.0703	-0.219	-0.101
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
Constant	0.470	1.314	0.568	0.501	1.002	0.588
	(0.009)	(0.000)	(0.019)	(0.007)	(0.000)	(0.023)
Observations	17166	6373	10793	17120	6371	10749
Groups	1929	601	1328	1914	600	1314
Sample	All Locations	U.S. Locations	Non-U.S. Locations	All Locations	U.S. Locations	Non-U.S. Locations

p-values in parentheses
All models include region fixed effects, year dummies and technology subcategory controls

Draft

TABLE 6: Correlation table for examiner only data set with DV as Total Citations Received (distance calculated)

Variables	1	2	3	4	5	6	7	8	9	10	11	12
Total Citations Received	1.00											
Non-Self Citations Received	1.00	1.00										
Self Citations Received	0.92	0.90	1.00									
Share Citations Made[Same Region, Same Assignee]	0.06	0.06	0.08	1.00								
Share Citations Made[Same Region, Different Assignee]	0.16	0.15	0.16	0.06	1.00							
Share Citations Made[Different Region, Same Assignee]	-0.00	-0.00	0.01	0.25	-0.02	1.00						
Share Citations Made[Different Region, Different Assignee]	-0.06	-0.06	-0.08	-0.58	-0.28	-0.89	1.00					
Share Citations Made[Same Region]	0.14	0.14	0.16	0.81	0.63	0.18	-0.62	1.00				
Share Citations Made[Same Assignee]	0.02	0.02	0.04	0.59	0.00	0.93	-0.96	0.46	1.00			
Log (Total Citations Made)	0.39	0.38	0.43	0.11	0.13	0.03	-0.11	0.17	0.07	1.00		
Log (Num Patents)	0.40	0.39	0.43	0.14	0.14	0.04	-0.12	0.19	0.09	0.96	1.00	
Log (Patent Pool Size)	0.35	0.34	0.38	0.16	0.17	0.04	-0.14	0.22	0.09	0.88	0.92	1.00

Draft

TABLE 7: Summary statistics for examiner only data set with DV as Total Citations Received (distance calculated)

Variable	Mean	Std. Dev.	N
Total Citations Received	581.998	3264.788	16464
Non-Self Citations Received	490.539	2829.21	16464
Self Citations Received	91.459	478.033	16464
Share Citations Made[Same Region, Same Assignee]	0.023	0.051	16464
Share Citations Made[Same Region, Different Assignee]	0.013	0.038	16464
Share Citations Made[Different Region, Same Assignee]	0.069	0.112	16464
Share Citations Made[Different Region, Different Assignee]	0.895	0.14	16464
Share Citations Made[Same Region]	0.036	0.065	16464
Share Citations Made[Same Assignee]	0.092	0.134	16464
Log (Total Citations Made)	4.352	2.162	16464
Log (Num Patents)	2.855	2.018	16464
Log (Patent Pool Size)	5.495	2.249	16464

Draft

TABLE 8: Correlation table for examiner only data set with DV as Non-Self Citations Received (distance calculated)

Variables	1	2	3	4	5	6	7	8	9	10	11	12
Total Citations Received	1.00											
Non-Self Citations Received	1.00	1.00										
Self Citations Received	0.92	0.90	1.00									
Share Citations Made[Same Region, Same Assignee]	0.06	0.06	0.09	1.00								
Share Citations Made[Same Region, Different Assignee]	0.16	0.16	0.16	0.06	1.00							
Share Citations Made[Different Region, Same Assignee]	-0.00	-0.00	0.01	0.25	-0.02	1.00						
Share Citations Made[Different Region, Different Assignee]	-0.07	-0.06	-0.09	-0.58	-0.28	-0.89	1.00					
Share Citations Made[Same Region]	0.14	0.14	0.16	0.81	0.63	0.18	-0.62	1.00				
Share Citations Made[Same Assignee]	0.02	0.02	0.04	0.59	0.00	0.93	-0.96	0.46	1.00			
Log (Total Citations Made)	0.39	0.38	0.43	0.12	0.13	0.04	-0.11	0.17	0.08	1.00		
Log (Num Patents)	0.40	0.39	0.43	0.15	0.14	0.04	-0.13	0.20	0.09	0.96	1.00	
Log (Patent Pool Size)	0.35	0.34	0.38	0.17	0.17	0.05	-0.14	0.23	0.10	0.88	0.92	1.00

Draft

TABLE 9: Summary statistics for examiner only data set with DV as Non-Self Citations Received (distance calculated)

Variable	Mean	Std. Dev.	N
Total Citations Received	583.91	3269.985	16410
Non-Self Citations Received	492.153	2833.721	16410
Self Citations Received	91.756	478.791	16410
Share Citations Made[Same Region, Same Assignee]	0.023	0.05	16410
Share Citations Made[Same Region, Different Assignee]	0.013	0.038	16410
Share Citations Made[Different Region, Same Assignee]	0.069	0.111	16410
Share Citations Made[Different Region, Different Assignee]	0.895	0.139	16410
Share Citations Made[Same Region]	0.036	0.065	16410
Share Citations Made[Same Assignee]	0.092	0.133	16410
Log (Total Citations Made)	4.36	2.161	16410
Log (Num Patents)	2.864	2.015	16410
Log (Patent Pool Size)	5.508	2.241	16410

Draft

TABLE 10: NB Regression Analysis of Invention Quality for Examiner Citations Only (Distance Calculated)

	(1) Total Citations Received	(2) Total Citations Received	(3) Total Citations Received	(4) Non-Self Citations Received	(5) Non-Self Citations Received	(6) Non-Self Citations Received
Share Citations Made[Same Region, Different Assignee]	-0.555	-0.987	-0.114	-0.364	-0.570	-0.0850
	(0.010)	(0.001)	(0.716)	(0.114)	(0.060)	(0.803)
Share Citations Made[Different Region, Same Assignee]	-0.466	-0.582	-0.290	-0.497	-0.389	-0.453
	(0.002)	(0.003)	(0.174)	(0.002)	(0.071)	(0.054)
Share Citations Made[Different Region, Different Assignee]	-0.944	-0.705	-0.953	-0.677	-0.274	-0.796
	(0.000)	(0.000)	(0.000)	(0.000)	(0.110)	(0.000)
Log (Total Citations Made)	0.255	0.200	0.299	0.248	0.179	0.300
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
Log (Num Patents)	0.586	0.680	0.589	0.572	0.685	0.567
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
Log (Patent Pool Size)	-0.112	-0.274	-0.120	-0.0863	-0.252	-0.0974
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
Constant	1.200	2.443	0.821	0.913	2.067	0.607
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.002)
Observations	16464	6246	10218	16410	6244	10166
Groups	1839	596	1243	1820	595	1225
Sample	All Locations	U.S. Locations	Non-U.S. Locations	All Locations	U.S. Locations	Non-U.S. Locations

p-values in parentheses
All models include region fixed effects, year dummies and technology subcategory controls

Draft

TABLE 11: Correlation table for applicant only data set with DV as Total Citations Received (distance calculated)

Variables	1	2	3	4	5	6	7	8	9	10	11	12
1. Total Citations Received	1.00											
2. Non-Self Citations Received	0.99	1.00										
3. Self Citations Received	0.79	0.70	1.00									
4. Share Citations Made[Same Region, Same Assignee]	0.05	0.04	0.08	1.00								
5. Share Citations Made[Same Region, Different Assignee]	0.14	0.13	0.13	0.08	1.00							
6. Share Citations Made[Different Region, Same Assignee]	-0.02	-0.03	0.01	0.19	-0.04	1.00						
7. Share Citations Made[Different Region, Different Assignee]	-0.04	-0.04	-0.07	-0.51	-0.33	-0.88	1.00					
8. Share Citations Made[Same Region]	0.13	0.12	0.14	0.72	0.75	0.10	-0.57	1.00				
9. Share Citations Made[Same Assignee]	-0.00	-0.01	0.03	0.51	-0.01	0.94	-0.94	0.33	1.00			
10. Log (Total Citations Made)	0.26	0.23	0.33	0.10	0.11	0.07	-0.13	0.14	0.10	1.00		
11. Log (Num Patents)	0.40	0.38	0.38	0.13	0.15	0.02	-0.11	0.19	0.06	0.70	1.00	
12. Log (Patent Pool Size)	0.36	0.34	0.35	0.14	0.19	0.01	-0.12	0.22	0.06	0.69	0.94	1.00

Draft

TABLE 12: Summary statistics for applicant only data set with DV as Total Citations Received (distance calculated)

Variable	Mean	Std. Dev.	N
Total Citations Received	1762.906	8670.540	8947
Non-Self Citations Received	1423.084	7358.189	8947
Self Citations Received	339.822	1754.63	8947
Share Citations Made[Same Region, Same Assignee]	0.016	0.042	8947
Share Citations Made[Same Region, Different Assignee]	0.016	0.045	8947
Share Citations Made[Different Region, Same Assignee]	0.065	0.11	8947
Share Citations Made[Different Region, Different Assignee]	0.903	0.132	8947
Share Citations Made[Same Region]	0.032	0.064	8947
Share Citations Made[Same Assignee]	0.081	0.125	8947
Log (Total Citations Made)	4.891	2.397	8947
Log (Num Patents)	3.807	1.928	8947
Log (Patent Pool Size)	6.586	2.02	8947

Draft

TABLE 13: Correlation table for applicant only data set with DV as Non-Self Citations Received (distance calculated)

Variables	1	2	3	4	5	6	7	8	9	10	11	12
1. Total Citations Received	1.00											
2. Non-Self Citations Received	0.99	1.00										
3. Self Citations Received	0.79	0.70	1.00									
4. Share Citations Made[Same Region, Same Assignee]	0.05	0.04	0.08	1.00								
5. Share Citations Made[Same Region, Different Assignee]	0.14	0.14	0.13	0.09	1.00							
6. Share Citations Made[Different Region, Same Assignee]	-0.02	-0.03	0.01	0.20	-0.04	1.00						
7. Share Citations Made[Different Region, Different Assignee]	-0.05	-0.04	-0.08	-0.51	-0.33	-0.88	1.00					
8. Share Citations Made[Same Region]	0.13	0.12	0.14	0.73	0.75	0.10	-0.57	1.00				
9. Share Citations Made[Same Assignee]	-0.00	-0.01	0.03	0.51	-0.01	0.94	-0.94	0.34	1.00			
10. Log (Total Citations Made)	0.26	0.23	0.33	0.11	0.11	0.08	-0.14	0.15	0.11	1.00		
11. Log (Num Patents)	0.40	0.38	0.38	0.14	0.15	0.02	-0.12	0.20	0.07	0.69	1.00	
12. Log (Patent Pool Size)	0.36	0.35	0.35	0.15	0.19	0.02	-0.13	0.23	0.07	0.69	0.94	1.00

Draft

TABLE 14: Summary statistics for applicant only data set with DV as Non-Self Citations Received (distance calculated)

Variable	Mean	Std. Dev.	N
Total Citations Received	1780.181	8711.249	8860
Non-Self Citations Received	1437.058	7392.873	8860
Self Citations Received	343.123	1762.906	8860
Share Citations Made[Same Region, Same Assignee]	0.016	0.042	8860
Share Citations Made[Same Region, Different Assignee]	0.016	0.043	8860
Share Citations Made[Different Region, Same Assignee]	0.065	0.108	8860
Share Citations Made[Different Region, Different Assignee]	0.903	0.13	8860
Share Citations Made[Same Region]	0.032	0.063	8860
Share Citations Made[Same Assignee]	0.08	0.123	8860
Log (Total Citations Made)	4.914	2.394	8860
Log (Num Patents)	3.836	1.913	8860
Log (Patent Pool Size)	6.622	1.993	8860

Draft

TABLE 15: NB Regression Analysis of Invention Quality for Applicant Citations Only (Distance Calculated)

	(1) Total Citations Received	(2) Total Citations Received	(3) Total Citations Received	(4) Non-Self Citations Received	(5) Non-Self Citations Received	(6) Non-Self Citations Received
Share Citations Made[Same Region, Different Assignee]	0.150	0.113	0.148	0.261	0.238	0.274
	(0.405)	(0.713)	(0.506)	(0.114)	(0.413)	(0.206)
Share Citations Made[Different Region, Same Assignee]	0.365	0.568	0.284	0.175	0.312	0.143
	(0.020)	(0.018)	(0.168)	(0.248)	(0.174)	(0.493)
Share Citations Made[Different Region, Different Assignee]	0.178	0.316	0.0835	0.0835	0.183	0.0659
	(0.170)	(0.102)	(0.631)	(0.497)	(0.317)	(0.704)
Log (Total Citations Made)	0.0198	0.0132	0.0252	0.0144	0.00658	0.0184
	(0.000)	(0.025)	(0.000)	(0.000)	(0.233)	(0.002)
Log (Num Patents)	0.783	0.848	0.830	0.780	0.824	0.843
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
Log (Patent Pool Size)	-0.117	-0.285	-0.113	-0.0903	-0.213	-0.116
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
Constant	-1.192	0.102	-1.553	-1.071	0.0237	-1.522
	(0.000)	(0.688)	(0.000)	(0.000)	(0.924)	(0.000)
Observations	8947	3798	5149	8860	3782	5078
Groups	1309	518	791	1281	512	769
Sample	All Locations	U.S. Locations	Non-U.S. Locations	All Locations	U.S. Locations	Non-U.S. Locations

p-values in parentheses
All models include region fixed effects, year dummies and technology subcategory controls

Draft

TABLE 16: Correlation table for other citations only data set with DV as Total Citations Received (distance calculated)

Variables	1	2	3	4	5	6	7	8	9	10	11	12
1. Total Citations Received	1.00											
2. Non-Self Citations Received	1.00	1.00										
3. Self Citations Received	0.92	0.87	1.00									
4. Share Citations Made[Same Region, Same Assignee]	0.05	0.04	0.08	1.00								
5. Share Citations Made[Same Region, Different Assignee]	0.10	0.10	0.12	0.07	1.00							
6. Share Citations Made[Different Region, Same Assignee]	-0.01	-0.01	-0.00	0.16	-0.05	1.00						
7. Share Citations Made[Different Region, Different Assignee]	-0.04	-0.03	-0.06	-0.49	-0.27	-0.89	1.00					
8. Share Citations Made[Same Region]	0.10	0.09	0.13	0.77	0.69	0.08	-0.53	1.00				
9. Share Citations Made[Same Assignee]	0.01	0.00	0.02	0.49	-0.02	0.94	-0.96	0.34	1.00			
10. Log (Total Citations Made)	0.27	0.26	0.32	0.10	0.13	0.02	-0.09	0.16	0.05	1.00		
11. Log (Num Patents)	0.27	0.26	0.31	0.17	0.14	0.04	-0.13	0.21	0.09	0.84	1.00	
12. Log (Patent Pool Size)	0.24	0.22	0.27	0.18	0.18	0.03	-0.14	0.24	0.09	0.79	0.93	1.00

Draft

TABLE 17: Summary statistics for other citations only data set with DV as Total Citations Received (distance calculated)

Variable	Mean	Std. Dev.	N
Total Citations Received	571.832	4349.983	13047
Non-Self Citations Received	448.739	3606.807	13047
Self Citations Received	123.093	828.482	13047
Share Citations Made[Same Region, Same Assignee]	0.02	0.05	13047
Share Citations Made[Same Region, Different Assignee]	0.016	0.043	13047
Share Citations Made[Different Region, Same Assignee]	0.078	0.126	13047
Share Citations Made[Different Region, Different Assignee]	0.887	0.148	13047
Share Citations Made[Same Region]	0.035	0.068	13047
Share Citations Made[Same Assignee]	0.097	0.142	13047
Log (Total Citations Made)	5.027	2.392	13047
Log (Num Patents)	3.297	1.964	13047
Log (Patent Pool Size)	5.983	2.121	13047

Draft

TABLE 18: Correlation table for other citations only data set with DV as Non-Self Citations Received (distance calculated)

Variables	1	2	3	4	5	6	7	8	9	10	11	12
1. Total Citations Received	1.00											
2. Non-Self Citations Received	1.00	1.00										
3. Self Citations Received	0.92	0.87	1.00									
4. Share Citations Made[Same Region, Same Assignee]	0.05	0.04	0.07	1.00								
5. Share Citations Made[Same Region, Different Assignee]	0.10	0.10	0.12	0.07	1.00							
6. Share Citations Made[Different Region, Same Assignee]	-0.01	-0.01	-0.00	0.16	-0.05	1.00						
7. Share Citations Made[Different Region, Different Assignee]	-0.04	-0.03	-0.06	-0.50	-0.27	-0.89	1.00					
8. Share Citations Made[Same Region]	0.10	0.09	0.13	0.77	0.69	0.08	-0.54	1.00				
9. Share Citations Made[Same Assignee]	0.01	0.00	0.02	0.49	-0.02	0.94	-0.96	0.35	1.00			
10. Log (Total Citations Made)	0.27	0.26	0.32	0.10	0.12	0.02	-0.09	0.15	0.05	1.00		
11. Log (Num Patents)	0.27	0.26	0.31	0.17	0.14	0.04	-0.13	0.21	0.09	0.84	1.00	
12. Log (Patent Pool Size)	0.24	0.23	0.27	0.18	0.18	0.03	-0.14	0.24	0.09	0.79	0.93	1.00

Draft

TABLE 19: Summary statistics for other citations only data set with DV as Non-Self Citations Received (distance calculated)

Variable	Mean	Std. Dev.	N
Total Citations Received	577.256	4370.279	12924
Non-Self Citations Received	453.01	3623.664	12924
Self Citations Received	124.246	832.331	12924
Share Citations Made[Same Region, Same Assignee]	0.02	0.05	12924
Share Citations Made[Same Region, Different Assignee]	0.016	0.043	12924
Share Citations Made[Different Region, Same Assignee]	0.078	0.125	12924
Share Citations Made[Different Region, Different Assignee]	0.887	0.148	12924
Share Citations Made[Same Region]	0.036	0.068	12924
Share Citations Made[Same Assignee]	0.097	0.142	12924
Log (Total Citations Made)	5.047	2.39	12924
Log (Num Patents)	3.322	1.955	12924
Log (Patent Pool Size)	6.016	2.101	12924

Draft

TABLE 20: NB Regression Analysis of Invention Quality for Other Citations Only (Distance Calculated)

	(1) Total Citations Received	(2) Total Citations Received	(3) Total Citations Received	(4) Non-Self Citations Received	(5) Non-Self Citations Received	(6) Non-Self Citations Received
Share Citations Made[Same Region, Different Assignee]	-1.176	-0.944	-1.351	-0.678	-0.674	-0.588
	(0.001)	(0.050)	(0.013)	(0.077)	(0.197)	(0.301)
Share Citations Made[Different Region, Same Assignee]	-0.238	-0.433	-0.106	0.0426	-0.0363	0.131
	(0.290)	(0.180)	(0.741)	(0.868)	(0.921)	(0.716)
Share Citations Made[Different Region, Different Assignee]	-0.607	-0.415	-0.720	0.0246	0.250	-0.0737
	(0.002)	(0.114)	(0.012)	(0.911)	(0.411)	(0.817)
Log (Total Citations Made)	0.228	0.186	0.212	0.164	0.113	0.156
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
Log (Num Patents)	0.554	0.586	0.646	0.610	0.633	0.726
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
Log (Patent Pool Size)	-0.0522	-0.101	-0.102	-0.00475	-0.0298	-0.0801
	(0.002)	(0.001)	(0.000)	(0.792)	(0.339)	(0.001)
Constant	0.254	0.452	0.342	-0.569	-0.396	-0.567
	(0.218)	(0.119)	(0.259)	(0.015)	(0.232)	(0.092)
Observations	13047	5618	7429	12924	5586	7338
Groups	1516	573	943	1484	566	918
Sample	All Locations	U.S. Locations	Non-U.S. Locations	All Locations	U.S. Locations	Non-U.S. Locations

p-values in parentheses
All models include region fixed effects, year dummies and technology subcategory controls

Draft

TABLE 21: Regression Analysis of Invention Quality* for Applicant Only Citations (from Feb 2017)

	(1) Total Citations Received	(2) Total Citations Received	(3) Total Citations Received	(4) Non-Self Citations Received	(5) Non-Self Citations Received	(6) Non-Self Citations Received
Share Citations Made[Same Region, Same Assignee]	-0.169	-0.104	-0.134	-0.232	-0.183	-0.316
	(0.218)	(0.623)	(0.454)	(0.105)	(0.392)	(0.123)
Share Citations Made[Same Region, Different Assignee]	-0.149	-0.191	-0.0919	0.0391	0.0945	0.0505
	(0.215)	(0.458)	(0.468)	(0.733)	(0.721)	(0.699)
Share Citations Made[Different Region, Same Assignee]	0.217	0.266	0.348	0.195	0.263	0.174
	(0.012)	(0.037)	(0.004)	(0.035)	(0.041)	(0.212)
Share Citations Made[Different Region, Different Assignee]	0.00828	0.0315	0.0179	0.00989	0.0203	0.00983
	(0.789)	(0.499)	(0.661)	(0.762)	(0.675)	(0.826)
Log (Total Citations Made)	0.0180	0.0198	0.0134	0.0150	0.0139	0.0107
	(0.000)	(0.001)	(0.014)	(0.000)	(0.023)	(0.075)
Log (Num Patents)	0.785	0.811	0.830	0.772	0.795	0.828
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
Log (Patent Pool Size)	-0.111	-0.227	-0.0944	-0.0453	-0.108	-0.0685
	(0.075)	(0.000)	(0.950)	(0.039)	(0.006)	(0.017)
Observations	9241	3885	5356	8879	3732	5147
Groups	1314	529	785	1199	478	721
Sample	All Locations	U.S. Locations	Non-U.S. Locations	All Locations	U.S. Locations	Non-U.S. Locations

p-values in parentheses
All regressions use negative binomial estimation on the region-year panel
* Citations received are counted all the way to end 2016 for which data is available
All models include region fixed effects, year dummies and technology subcategory controls

Draft