The impact of modularity on knowledge spillovers:
Role of IPR regime and Cross-border Inventions

Ashwin Iyenggar (1521001)
ashwin.iyenggar15@iimb.ernet.in

January 6, 2017

Abstract

I define modularity as a construct to classify patents and investigate two relationships: first, that between the intellectual property rights (IPR) score of invention locations and modularity of patents invented there, and second, that between the occurrence of cross border inventions and modularity of patents invented. I argue that the levels of modularity may partially explain the variation in local spillovers of knowledge across the various locations across the world. I conclude showing support for the influence of modularity on local spillovers.

Keywords: Modularity, Weak and Strong IPR Regions, Spillovers, Patent Citations

1 Introduction

For long, scholars have highlighted the agglomeration characteristics of economic regions (Marshall, 2009). In the last three decades, scholars have demonstrated through numerous studies that patent citations provide us with a paper trail of evidence for the existence the knowledge spillovers in economic regions (Almeida and Kogut, 1999; Jaffe et al., 1993), the effects of inventor mobility (e.g., Almeida and Kogut (1999)), of Intellectual Property Rights regime of locations (e.g., Zhao (2006)) and of the role of international geography (e.g., Singh (2007)) on knowledge spillovers. Since the extent of knowledge spillovers is observed in practice to be highly heterogenous across locations, firms and legal regimes, the question of the causal mechanisms leading to knowledge spillovers remains largely unresolved, despite the enormous progress made so far. In this article, I intend to explore the modularity of patents invented as a potential mechanism influencing the extent of local knowledge spillovers.

The investigation of potential mechanisms behind local spillovers is both interesting and important. Given the wide disparity in the extent of knowledge spillovers across locations, across firms and across IPR regimes it is intriguing to a researcher to find the mechanisms that may lie behind such a phenomenon. A specific flavor of this question is the investigation of the spillover effects of patenting in emerging countries, or those known to have weaker IPR regimes. Specifically, do multinational firms that develop patentable technologies in emerging countries create spillover effects in the host country talent pool, or do the benefits remain localized to within multinational companies (MNCs)? From a human capital policy perspective, it is valuable to understand the impact of allowing MNCs dominate the patenting process in emerging markets on the quality of the talent pool in the host country. Does a significant group of local inventors develop? Is this affected by the strength of the IPR regime in the host country? Patents data allows us to ask and try and answer this question.

Modularity may be seen as either an attribute of usage, or as an attribute of invention. A patent that is used (cited) by several patents belonging to distinct and different patent technology classes maybe seen as modular by virtue of it being able to be plugged into multiple diverse applications. Alternatively a patent that is constructed with few dependencies may also be seen as being modular by virtue of its capacity to be developed standalone, or with minimal intervention from other modules. For the purposes of my study, I use a definition of modularity that captures both the effects above.

Prior literature has looked at knowledge spillovers within geographic regions (e.g., Jaffe et al. (1993)) as well as across regions (e.g., Singh (2007)). For the purposes of my study, I refer to local knowledge spillovers as those that occur within an adjacent geographical area. This is in keeping with my objective of trying to understand the local impact of inventing activity by multinationals in emerging nations.

This article therefore, attempts to answer the following questions. First, does the modularity of patents developed across country borders differ from the modularity of patents that are developed within a country? Second, does the modularity of patents differ by the strength of the IPR regime of the inventor location? Finally, does the modularity of patents affect the extent of local knowledge spillovers?

I am interested in understanding the spillover effects of patenting in emerging countries. Specifically, do multinational firms that develop patentable technologies in emerging countries end up creating spillover effects in the host country talent pool, or do the benefits remain localized to within MNCs. My focus in on human capital, and I wish to understand the impact of a allowing MNC's dominate the patenting process in emerging markets on the quality of the talent pool in the host country. Specifically, is there a significant group of local inventors who develop? Do they then move around to cross-pollinate to other firms? Or do domestic firms get completely left out. Does the talent pool demonstrate a clear MNC vs Domestic company division, i.e., to employees who have previously worked at MNCs move only to other MNCs, and those who had previously worked at Domestic Firms only continue to do so? I may try and understand this from patent data though this may be a large question to be asked for all roles. But the specific context of highly skilled, technologically intensive inventive roles are definitely interesting to look at.

How do the patenting patterns of inventors whose first patent was co-invented with someone living abroad compare with that of inventors whose first patent was co-invented with someone living in the same location.

I am interested in understanding the impact of cross-national inventions on subsequent patenting career of inventors in weak IPR (or emerging) countries. Do they end up patenting more from the weak IPR country or do they eventually move to a strong IPR country. Do they develop more local inventors or do they patent more with international collaborators?

Zhao (2006) argues that competing firms may have a lower ability to imitate when the value of technology is highly dependent on the proprietary firm's internal resources. We contend that that assumption is weakened significantly if the knowledge of this technology is codified and published, and when published work on related technologies cite a common prior art. Zhao (2006) herself cites Kogut and Zander (1993) suggesting that difficult to codify knowledge lends itself to more efficient transfer within the firm. Drawing on Cohen et al. (2000), we thereby conclude that firms would have a greater incentive to keep such highly dependent technology developed in weaker IPR countries secret, rather than make this knowledge public. While Zhao (2006) highlights the strength of internal firm linkages in identifying and appropriating the knowledge generated in weak IPR location subsidiaries, she does not emphasize the importance of secrecy. Indeed, she uses patenting data from weaker IPR location offices to substantiate her hypothesis, which we believe actually weakens her argument for the strength of internal firm linkages.

2 Theory

Lay the hypotheses here

Hypothesis 1: I hope I can have something to go

3 Research Design

3.1 Modularity

I construct my measure of modularity based interactions between the different patent sub-classes. Since each of the interactions between patent sub-classes may introduce a new interaction, I model interactions on a binomial function. Specifically, when `subclass` represents the number of distinct patent sub-classes, I define `interaction(subclass)` as follows:

$$interaction(subclass) = \begin{cases} 1 & : subclass \leq 2 \\ \binom{subclass}{2} & : subclass > 2 \end{cases}$$

I would expect, from a user perspective that the more number of contexts in which the patent is valuable, the higher should be the modularity. If `modularity` represents my measure of the modularity of the patent, and `usage contexts` represents the number of distinct contexts where the patent is found valuable, I should expect the following relationship to hold:

$$\texttt{modularity} \propto \texttt{usage contexts}$$

Similarly, from an inventor perspective, the more the number of contexts that the patent is built on, the lower should be the modularity. A patent that is developed without citing any other patents is an extreme case of highest modularity, while one that requires to be built upon several `source contexts` is properly understood as being less modular. The relationship between `source contexts` and `modularity` is therefore an inverse one as depicted below.

$$\texttt{modularity} \propto \frac{1}{source\ contexts}$$

Using the principles above, I therefore develop the following definition of modularity.

$$\texttt{modularity} = \frac{interaction(subclass_{\text{patent}})}{interaction(subclass_{\text{cited}})}$$

By the definition above, a patent that cites no patents (and hence has $subclass_{\text{cited}} = 0$) but is itself assigned to 4 sub-classes (and hence has $subclass_{\text{patent}} = 4$) will have a raw modularity score of $\frac{\binom{4}{2}}{1} = 6$. If the patent itself had been assigned onto to 2 sub-classes, the raw modularity score would have been just 1. Therefore, the more the number of patent sub-classes a patent is assigned to, the higher its modularity score (by a square term). A similar but inverse relationship would hold for sub-classes arising out of cited patents. Here, I take a set union of patent sub-classes assigned to each cited patent, and use that count to determine the value of the `interaction` function.

3.2 IPR Classification

A review of the academic literature surrounding the construction of IPR indexes indicated that there were several, as was also evident in Zhao (2006) constructing a composite measure for the purposes of her article. Lesser (2010) provides an alternative, composite scoring system that includes the following components: protectable subject matter, membership in convention, enforcement, administration and duration of protection. I have therefore used the scores generated by Lesser (2010) for the purposes of this study. The extensive table of IPR scores in presented in Table 7. The listing has several countries for which scores have not been provided. However none of the top patenting nations were among them, and I therefore chose to go along with this scale.

3.3 Data Source

I derive all patents data for this study from patentsview.org. The dataset considered is for all USPTO patents filed in the period 1976 to 2015. For the IPR Scores, I rely on the scores generated by Lesser (2010). For country definitions, I use the resources provided by Thematic Mapping. To determine if spillovers are local, I use a composite data source as described in the following. For locations in the United States, it has been standard to use Metropolitan Statistical Areas (MSA) for analyses related to economic geography. Such standardized data is unavailable for non-US locations. Urban areas are a close substitute for economic centers, and I therefore determine to use one such definition for non-US locations. My data source for MSA of US locations is the US census and that for urban areas for world wide locations is Natural Earth Data.
This automatically raises conflicting definitions for locations in the United States. So that the MSA definitions take precedence, I eliminated all data pertaining to US locations from the Natural Earth urban centers data and integrated this with the MSA information. With this I generated a single database of location information for economic centers around the world.

3.4 Dependent Variable

For the first two questions, I assess the impact of IPR regime and Cross-border inventions on modularity of patents. Since I have used a binomial measure for the raw value of modularity, I define my dependent variable for the first two questions to be the logarithm of the raw modularity score. The summary statistics for my primary quantitative variables is provided in Table 1.

For the third question on impact of modularity on local spillovers, I define as my dependent variable, `local spillover` which is a variable that takes a value between 0 and 1, with 0 representing no local spillover and 1 representing the maximum spillover. Therefore for patent$_i$

$$local\ spillover_i = \frac{Citations\ to\ Patents\ from\ Same\ Country_i}{Total\ Citations_i}$$

Table 1: Summary statistics

Variable	Mean	Std. Dev.	Min.	Max.	N
Log Modularity	1.224	2.331	-15.147	13.303	4252508
IPR(Inventor Location)	9.890	1.028	1.479	11.877	4249896
IPR(Assignee Location)	9.926	0.944	1.479	11.877	4240430
Local Spillover (0,1)	0.055	0.206	0	1	4252508

3.5 Explanatory Variables

3.5.1 IPR (Inventor Location) and IPR (Assignee Location)

While most patents have multiple inventors, and some patents also have multiple assignees, my question requires us to associate a single location to the inventor of a patent, and a single location for the assignee of the patent. For the inventor location, I tabulate the count of each of the regions that each inventor is a resident of at the time of the filing of the patent application. In doing so, I treat all inventors equally and allocate the most frequently occurring location as the location of the inventor for that patent. In case of a tie, I assign the location of the first inventor (given by the sequence number of the inventor on the patent) as the location of the inventor of the patent.

For the assignee location, I treat multiple assignees as having been granted separate patents. I do this since the number of patents with multiple assignees is small, and so as to not lose potentially valuable information.

3.5.2 Localized Invention

Patents for which the inventor resides in the same country as the assignee are marked as inventions that are localized. This is in contrast to those where the inventor and assignees are located in different countries. I capture this difference to identify the potential variation in modularity caused on account of cross-border collaborations in inventions.

3.5.3 Crossborder Invention

A patent in which the location country of the inventor differs from the country of the assignee is marked as a crossborder invention.

3.5.4 Dummy variables and Interaction Terms

I create additional dummy variables for countries with weak IPR regimes (those with an IPR score of ≤ 7). Additionally I interact various explanatory variables, dummy variables as has been presented in the results in Table 5, Table 6, and Table 4.

Table 2: T-test for modularity of patents by localized vs. crossborder inventions

	(1) Localized Invention
Log Modularity	-0.261***
	(-63.13)
N	4252508

t statistics in parentheses
* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

4 Results

4.1 Localized vs. Crossborder Inventions

The preliminary answer to my first question on the impact of the location on log(modularity) is provided by a t-test in Table 2. I see that localized inventions on an average score a log(modularity) value of about 0.26 higher than those scored by crossborder inventions. This result comes as a bit of a surprise as one might have expected that inventions being developed across borders may have been the ones with the least interaction with others. The result on 4.25 Million patents filed between 1975 and 2015 provides a statistically significant result that it is the crossborder inventions that are of lower modularity (or alternatively, of higher complexity).

Table 3: T-test for modularity of patents by strength of IPR regime of assignee country

	(1)
	Weak IPR Dummy
Log Modularity	-0.00533
	(-0.65)
N	4252508

t statistics in parentheses

* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

Table 4: Regression Results

Effects of IPR strength and Localization of Inventions on Modularity

VARIABLES	(1) log(modular)	(2) log(modular)	(3) log(modular)	(4) log(modular)	(5) log(modular)	(6) log(modular)
IPR(Inventor Location)	0.02		0.07***		0.05**	0.06***
	(0.04)		(0.02)		(0.02)	(0.02)
IPR(Assignee Location)		0.00	-0.06*		-0.03	0.01
		(0.04)	(0.03)		(0.04)	(0.02)
Localized Invention				0.26***	0.23**	0.94**
				(0.10)	(0.11)	(0.42)
IPR(Inventor Location) * Localized Invention						-0.07
						(0.05)
Constant	0.99***	1.19***	1.13***	0.98***	0.90**	0.26
	(0.36)	(0.33)	(0.35)	(0.05)	(0.39)	(0.30)
Observations	4,249,896	4,240,430	4,239,225	4,252,508	4,239,225	4,239,225
R^2	0.00	0.00	0.00	0.00	0.00	0.00

Assignee location clustered standard errors in parentheses

*** p<0.01, ** p<0.05, * p<0.1

4.2 Strong vs. Weak IPR location of Assignee

Table 3 presents the results from a t-test of log(modularity) by the weak ipr dummy variable. The table suggests that weak and strong ipr locations do not differ significantly in the log(modularity) of their patents. This answers my second question.

Table 5: Additional Regression Results

Effects of Localized and Crossborder Inventions in Weak and Strong IPR Locations on Modularity

VARIABLES	(1) log(modular)	(2) log(modular)	(3) log(modular)	(4) log(modular)	(5) log(modular)
Crossborder Invention in Strong IPR Location	-0.26***			-0.25***	-0.14***
	(0.10)			(0.08)	(0.02)
Crossborder Invention in Weak IPR Location	-0.12	-0.09			-2.26***
	(0.16)	(0.16)			(0.76)
Localized Invention in Weak IPR Location		0.01	0.25***		-2.02***
		(0.14)	(0.08)		(0.74)
Localized Invention in Strong IPR Location			0.26**	0.02	
			(0.10)	(0.15)	
Constant	1.25***	1.22***	0.98***	1.23***	3.25***
	(0.13)	(0.13)	(0.05)	(0.05)	(0.05)
Observations	4,252,508	4,252,508	4,252,508	4,252,508	4,252,508
R^2	0.00	0.00	0.00	0.00	0.01

Assignee location clustered standard errors in parentheses

Model 5 uses fixed effects for Inventor location and Assignee location

*** p<0.01, ** p<0.05, * p<0.1

4.3 Regressions on log(modular)

Table 5 lays out the results from my primary regressions. As was suggested by the t-test in Table 3, models 1 and 2 indicate a statistically non-significant effect of inventor and assignee location IPR scores on log (modularity). However Inventor location IPR shows up as statistically significant in models 3, 5 and 6 when multiple variable or interactions are included. This indicates that the impact of Inventor location IPR is not quite conclusive. This is therefore further investigated with the results from Table 5 where the effect of inventor location IPR is considered alongside the localization dimension of the invention. I find there, that both localized and crossborder inventions in weak IPR locations signify a lower level of modularity than localized inventions in strong IPR locations.

The strongest support from Table 4 is for the effect of localized inventions on log(modularity). Across models 4, 5, and 6 in Table 4, I find that local inventions are on an average more modular than crossborder inventions.

Table 5 looks at the impact of IPR strength and localization jointly. For model 5, I apply country level fixed effects for both inventor location as well as assignee location. Here, I find further evidence to support my initial finding that crossborder inventions are less modular (alternatively, more complex) than are localized inventions in weak IPR locations. This is despite the fact that localized inventions in weak IPR locations themselves are less modular than those

Table 6: Regression Results

Effect of Modularity on Local Spillovers		
	(1)	(2)
VARIABLES	Local Spillover	Local Spillover
Log Modularity	0.0050***	0.0049***
	(0.0013)	(0.0013)
IPR(Inventor Location)		0.0005
		(0.0009)
IPR(Assignee Location)		0.0027
		(0.0017)
Localized Invention		-0.0044
		(0.0486)
IPR(Assignee Location) * Localized Invention		0.0023
		(0.0059)
Constant	0.0487***	-0.0001
	(0.0156)	(0.0156)
Observations	4,252,508	4,239,225
R^2	0.0032	0.0043

Robust standard errors in parentheses

*** p<0.01, ** p<0.05, * p<0.1

in strong ipr locations to begin with. I may therefore conclude that multinationals conduct their R&D across the world and in countries with weak IPR not because that research is more modular. Given the higher complexity of crossborder patenting activity, I may conjecture that the crossborder collaborations are due to key competencies located globally rather than due to ease of offshoring.

4.4 Regressions on Local Spillovers

In Table 6, I present my results for my investigation on whether modularity affects local spillovers[1]. Models 1 and 2 demonstrate that there is indeed a positive effect of log (modularity) on local spillovers across my sample of 4.2 million patents. When coupled with my findings in the previous section, I have some support for the effect of patenting activity by multinational firms in weak IPR locations on local spillovers. I discuss this in the following section on conclusions.

5 Conclusions

I started this study attempting to understand if I could use the mechanism of patent modularity to explain the heterogeneity in knowledge spillovers across locations, firms and IPR regimes. My study found that IPR regimes do not by themselves seem to directly affect the modularity of patents invented in those locations. However, I find strong evidence for the fact that multinationals (or

[1] Given that Local Spillover takes values between 0 and 1, it might have been appropriate to use a different estimation method than OLS

crossborder inventions, where the inventor and assignee are located in different countries) on an average file for more complex patents. My third finding was that modularity of patent work was positively correlated with local spillovers. Putting results two and three suggests therefore, that crossborder patenting performed in both weak IPR and strong IPR locations are unlikely to generate high local spillovers for those locations.

References

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

Cohen, W. M., Nelson, R. R., and Walsh, J. P. (2000). Protecting their intellectual assets: Appropriability conditions and why u.s. manufacturing firms patent (or not). Working Paper 7552, 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.

Kogut, B. and Zander, U. (1993). Knowledge of the firm and the evolutionary theory of the multinational corporation. *Journal of International Business Studies*, 24(4):625–645.

Lesser, W. (2010). Measuring intellectual property strength and effects: An assessment of patent scoring systems and casuality. *J. Bus. Entrepreneurship & L.*, 4:345.

Marshall, A. (2009). Principles of economics: unabridged eighth edition.

Singh, J. (2007). Asymmetry of knowledge spillovers between mncs and host country firms. *Journal of International Business Studies*, 38(5):764–786.

Zhao, M. (2006). Conducting r&d in countries with weak intellectual property rights protection. *Management Science*, 52(8):1185–1199.

Table 7: Countries and their IPR scores (Lesser, 2010)

Country	IPR Score
Afghanistan	
Albania	4.7682
Algeria	2.7608
Angola	1.8734
Anguilla	
Antigua and Barbuda	
Argentina	5.4684
Armenia	4.4032
Aruba	
Australia	11.1872
Austria	9.4024
Azerbaijan	3.1358
Bahamas	
Bahrain	5.7736
Bangladesh	2.3664
Barbados	
Belarus	3.2344
Belgium	9.6096
Belize	
Benin	
Bermuda	
Bhutan	4.9300
Bolivia	4.2752
Bosnia and Herzegovina	2.9580
Botswana	6.2666
Brazil	5.2612
British Virgin Islands	
Brunei Darussalam	5.4230
Bulgaria	5.3598
Burkina Faso	3.5496
Burma	
Cambodia	1.9720
Cameroon	2.1692
Canada	11.1872
Cayman Islands	

Country	IPR Score
Central African Republic	1.9720
Chad	1.5776
Chile	9.2152
China	6.1586
Colombia	6.2572
Congo	1.8734
Cook Islands	
Costa Rica	6.8388
Cote d'Ivoire	2.0706
Croatia	5.8528
Cuba	
Cyprus	7.2526
Czech Republic	6.4444
Democratic Republic of the Congo	3.8260
Denmark	11.7788
Djibouti	
Dominica	
Dominican Republic	
Ecuador	3.7822
Egypt	2.7608
El Salvador	3.3524
Equatorial Guinea	
Estonia	9.1166
Ethiopia	2.6622
Fiji	
Finland	11.3844
France	10.3984
French Guiana	10.3984
Gabon	2.8594
Gambia	2.8594
Georgia	4.9106
Germany	10.4970
Ghana	4.5904
Greece	5.4878
Greenland	
Guadeloupe	
Guam	

Country	IPR Score
Guatemala	3.3524
Guernsey	
Guinea	1.7748
Guinea-Bissau	
Guyana	2.5636
Haiti	1.7748
Honduras	3.2100
Hong Kong	8.0852
Hungary	7.6376
Iceland	10.1912
India	4.0974
Indonesia	4.5018
Iran (Islamic Republic of)	1.7748
Iraq	1.4790
Ireland	9.6290
Isle of Man	
Israel	8.6236
Italy	6.8488
Jamaica	2.9580
Japan	10.2012
Jersey	
Jordan	6.5430
Kazakhstan	2.6622
Kenya	3.7822
Korea, Democratic Republic of	
Korea, Republic of	7.1640
Kuwait	4.0426
Kyrgyzstan	3.4864
Lao People’s Democratic Republic	1.9720
Latvia	6.0500
Lebanon	2.4650
Lesotho	
Liberia	3.0566
Libyan Arab Jamahiriya	
Liechtenstein	
Lithuania	7.4404
Luxembourg	8.8302

Country	IPR Score
Macau	
Madagascar	2.9580
Malawi	3.2538
Malaysia	5.1820
Mali	2.7608
Malta	
Mauritania	2.4650
Mauritius	5.3244
Mexico	4.8668
Monaco	
Mongolia	3.4072
Montenegro	
Morocco	5.8628
Mozambique	2.4650
Namibia	4.4370
Nepal	2.2678
Netherlands Antilles	11.3844
Netherlands	11.3844
New Caledonia	
New Zealand	11.8774
Nicaragua	5.0740
Niger	2.8594
Nigeria	3.2100
Northern Mariana Islands	
Norway	10.1912
Oman	7.0360
Pakistan	4.1074
Palau	
Palestine	
Panama	5.2164
Papua New Guinea	2.0706
Paraguay	3.6836
Peru	5.3892
Philippines	4.1074
Poland	7.5390
Portugal	8.3278
Puerto Rico	

Country	**IPR Score**
Qatar	7.6470
Republic of Moldova	4.1218
Reunion	
Romania	6.3558
Russia	4.0332
Saint Barthelemy	
Saint Kitts and Nevis	
Saint Lucia	
Saint Pierre and Miquelon	
San Marino	
Saudi Arabia	4.2398
Senegal	2.9580
Serbia	4.4470
Seychelles	
Sierra Leone	2.1692
Singapore	11.6802
Slovakia	7.0460
Slovenia	8.3716
Solomon Islands	
South Africa	7.2432
Spain	8.6236
Sri Lanka	3.0566
Sudan	1.4790
Suriname	3.6482
Svalbard	
Swaziland	4.2946
Sweden	11.6802
Switzerland	11.4830
Syrian Arab Republic	3.5596
Taiwan	7.2626
Tajikistan	1.9720
Thailand	4.0974
The former Yugoslav Republic of Macedonia	
Togo	2.7608
Trinidad and Tobago	5.1626
Tunisia	5.8528

Country	**IPR Score**
Turkey	6.9474
Turkmenistan	
Turks and Caicos Islands	
Uganda	2.4650
Ukraine	3.7822
United Arab Emirates	6.4090
United Kingdom	10.2012
United Republic of Tanzania	2.5636
United States Virgin Islands	10.0040
United States	10.0040
Uruguay	8.2192
Uzbekistan	3.6388
Venezuela	3.6144
Vietnam	4.2752
Yemen	2.0706
Zambia	2.9580
Zimbabwe	2.9142