Innovation and Development

ISSN: 2157-930X (Print) 2157-9318 (Online) Journal homepage: https://www.tandfonline.com/loi/riad20

Economic geography of innovation in India: an empirical investigation

Aasheerwad Dwivedi & Aakanksha Arora

To cite this article: Aasheerwad Dwivedi & Aakanksha Arora (2019): Economic geography of innovation in India: an empirical investigation, Innovation and Development, DOI: 10.1080/2157930X.2019.1679952

To link to this article: https://doi.org/10.1080/2157930X.2019.1679952

Published online: 08 Nov 2019.

INNOVATION AND DEVELOPMENT
https://doi.org/10.1080/2157930X.2019.1679952

Routledge
Taylor & Francis Group

Economic geography of innovation in India: an empirical investigation

Aasheerwad Dwivedi[a] and Aakanksha Arora[b]

[a]CSRD, Jawaharlal Nehru University, New Delhi, India; [b]Department of Economic Affairs, Ministry of Finance, Government of India, New Delhi, India

ABSTRACT
Theory points toward geographical clustering of economic activity as an important determinant of innovation. A stylized fact about the Indian economy is the geographic concentration of both economic activity and innovation. The paper first analyses the spatial pattern of patent applications across Indian districts. Considering innovation to be a complex and collaborative process, this paper investigates the relevance of economic geography for innovation for all Indian districts. We estimate agglomeration economies by creating Herfindahl index, localization index and Access index for all districts. Using the count data model, we estimate the effect of agglomeration economies, knowledge spillovers along with other variables including R&D expenditure, human capital, institution and infrastructure in creating innovation. The results show a strong and statistically significant effect of the agglomeration variables on innovative activity in a district. R&D expenditures in own district and neighbouring districts also have a positive relationship with the number of patent applications. Moreover, institutions, infrastructure and local socio-economic conditions do matter for innovative activity.

KEYWORDS
Innovation; patents; agglomeration; knowledge spillover; infrastructure index; institution index

1. Introduction

There is a general consensus among academicians and policy makers that technology plays an important role in the growth and development of any region. From Veblen's (1915) work on German industrialization to neoclassical growth theories, technology has played an important role in explaining economic growth. Neoclassical theories prior to Romer, such as Solowik (1956) considered technological progress as exogenous. Knowledge was treated as a public good, which can be produced and transferred anywhere, leading to the prediction of long-term convergence. However, this view that technological catch-up by the late-comers is natural was challenged by 'technology gap theory of economic growth' (Abramovitz 1979, 1986). Furthermore, the advocates of 'new-growth theory' brought further primacy to the role of technology in the growth and development in the mainstream economic theory (Romer 1986, 1990; Lucas 1988; Howitt and Aghion 1998).

CONTACT Aasheerwad Dwivedi aasheerwad.bhu@gmail.com, aasheerwad.dwivedi@gmail.com

Technological progress is driven by innovation and so it is essential to understand how innovative activity takes place. Innovation occurs in dynamically diverse, spatially concentrated and imperfectly competitive space, and hence it can only be analysed by abandoning the conventional assumption of perfectly competitive markets and constant returns to scale (Baldwin and Martin 2004; Krugman 1991). It is a stylized fact that innovation activities are highly concentrated, sometimes even more concentrated than the economic activity (Florida 1994). The latest literature on innovation recognizes innovation to be a dynamic, localized, interactive and complex process (Kumar and Joseph 2006; Prakash Pradhan 2011). One crucial theoretical framework, which builds on the premise that innovation is an outcome of non-linear, interactive and learning process, is 'Innovation System' (Lundvall 1992, 2016; Freeman 1987).

Over the last 40 years, the remarkable success of industrial clusters, such as the Silicon Valley, the southern California electronic industry, and information technology industry in Bengaluru and their outstanding performance in innovative activity, has brought the locational dynamics of innovation at the forefront of policies for development. The successes of these clusters are based on taking advantages of agglomeration benefit supported by better network linkages, infrastructure and supporting social relations as well (Lawson 1997; Marshall and Marshall 1920). This could also be the reason behind increasing localization of innovative activity, where network between local firms and institutions leads to innovation (Lundvall and Borrás 1998).

Regional Innovation Systems (RIS), as a perspective, provide the analytical framework to understand the dynamics of innovation at regional level. There is no specific definition of RIS; however, broadly, it encompasses economic, social and institutional contexts under which innovation occurs (Doloreux and Parto 2005). One key take-away from the literature on RIS is the significance of proximity for interactive learning by sharing tacit and explicit knowledge (Prakash Pradhan 2011). In RIS, all aspects of learning and innovating are directly or indirectly linked to the concept of proximity, with spatial concentration being the most popular manifestation of this notion. It is widely recognized that not only creation of new ideas and innovation, but also the absorption of innovations generated in other regions is equally valuable for technological development, and in fact, the latter is more important in the context of a developing country like India (Marrocu, Paci, and Usai 2013). Emerging economies like India are mired by a high degree of regional inequality in terms of infrastructure provision, R&D investment, labour markets, institutions, etc. Hence, the nature and dynamics of innovation in a developing country can be quite different from that in the developed world. In fact, many researchers have argued that the RIS in developing and developed countries function differently (Lundvall et al. 2009). For instance, there are many micro-, small and medium enterprises in India, the characteristics of which may be very different as compared to large and mega firms in terms of the availability of credit, capacity for in-house R&D, hiring of skilled and technical people, etc. This may also mean that the firms in a specific region in India may be more affected by the spillover of knowledge and technological knowhow as compared to those in any region in the developed economies. Also, the cost of mobility of people may vary across countries depending on the transport infrastructure. Thus, one might expect that the innovative activity may be even more dependent on the characteristics of the location in developing economies.

There are numerous studies on the determinants of innovation at regional level for developed economies, but almost negligible for developing economies. This is also true for India, on which there are very few studies and most focus on sectoral or firm-level analysis (Prakash Pradhan 2011). This paper is an attempt to fill this gap in the literature. We use the regional production function approach to assess the role of agglomeration economies, R&D activity, infrastructure, institutions and other local characteristics in determining the innovative activity of Indian districts.

2. Literature review

The development of a region depends on its ability to produce new products and processes, for which tacit knowledge plays a crucial role (Pavitt 2002). In the world of Internet, codified knowledge and falling transportation costs, the creation of new knowledge still largely depends on the availability and use of tacit knowledge (Maskell and Malmberg 1999). The difficulty of exchange of tacit knowledge over long distances makes it a key determinant for the geography of innovative activity. It also makes innovation a social process, where knowledge flow via interaction between different players like research organizations, firms, public agencies, etc. becomes very crucial. This is also the basic idea behind the popular 'innovation systems' literature, where innovation is path dependent and is the outcome of an interactive process (Lundvall 1992, 2016; Freeman 1987). The literature on 'Learning regions', 'Learning through interacting' and 'Learning economy thesis' has conveyed that the transmission of tacit knowledge depends on face to face interaction between players who share some common features like language, conventions, institutions, etc. (Lundvall and Johnson 1994; Asheim 1996; Lundvall 2000). These common features are really important for knowledge flow, which also make spatial proximity essential for the effective creation and transmission of tacit knowledge. The success of innovative clusters, districts or regions further reinforce this idea of territorial knowledge sharing which leads to innovation and better economic outcomes.

In theory, we find the following explanation for this peculiar phenomenon of concentration of innovative activity. First, despite the availability of codified knowledge, the circulation of new knowledge and knowledge spillovers remains localized (Marrocu, Paci, and Usai 2013); second, these clusters also attract highly educated and motivated workforce both for the challenging nature of work and the quality of life provided. Storper and Venables (2004) very creatively call this phenomenon leading to concentration as 'buzz'. Interestingly, the concept of RIS, which appeared first in the early 1990s, was inspired by the success and popularity of clusters. In fact, some scholars see RIS as *expost* rationalization of the success of clusters (Lundvall et al. 2009).

The interconnected role of geographical proximity and local infrastructure on regional innovative capacity can be analysed by using the Knowledge Production Function (KPF) approach at the district level. This was, indeed, proposed by Jaffe (1989). KPF was first introduced by Griliches (1979, 1984) to study the relationship between innovative output and input at firm level. Jaffe (1989) modified the KPF given by Griliches to account for spatial and product dimensions and this modification has changed the traditional unit of analysis from a firm to a geographical unit.

In order to understand the role of geography in economic and innovative activity, empirical research has classified agglomeration economies into either localization

economies or urbanization economies, as was first noted by Loesch (1954). Localization economies are external to a firm, but internal to an industry within a geographic region, whereas urbanization economies refer to the diversity of economic activity within a region. This classification pertains to varying composition of economic activity in a region and hence will have different implications for innovation. Localization might lead to better provision of industry-specific complementary assets which foster both concentration and innovation (Glaeser et al. 1992). Some scholars relate urbanization economies to the scale effect, external to industries, but internal to a geographic unit, which leads to the formation of cities as well (Lucas 1993). Jacobs (1969) gave a rather compelling explanation of urbanization economies by tracing it to the exchange of complementary knowledge across diverse firms and economic agents within a spatial unit.

Many scholars of different schools of thought have worked using the KPF by augmenting its traditional formulation with inter-regional factors, infrastructure variables, human capital, agglomeration effect, etc. Jaffe's work was further extended by many scholars who provided evidence in favour of local externalities both within and across the regions of USA (Acs, Audretsch, and Feldman 1992; Anselin, Varga, and Acs 1997). Similar studies have also been done using the data of European Union and most of these find that both internal factors and spillover from the nearby region are key determinants of innovative performance (Acosta et al. 2009; Buesa, Heijs, and Baumert 2010; Tappeiner, Hauser, and Walde 2008). Some researchers have found evidence in support of localized knowledge spillover originating mainly from academic research (Anselin, Varga, and Acs 1997; Fritsch and Slavtchev 2007; Kantor and Whalley 2009). There is support for the positive impact of both concentration and diversification on the innovative capacity of a region (Feldman and Audretsch 1999; Mukim 2012). Some other studies specifically focus on impact of R&D expenditure in a region on the innovative capacity in neighbouring regions (Ponds, van Oort, and Frenken 2010; Fischer and Varga 2003a, 2003b). Most studies have found a positive impact of higher R&D spending in neighbouring areas, although this impact decays with an increase in distance (Bottazzi and Peri 2003; Bode 2004). The most popular measure of innovation used in the literature is the number of patents in any given area. Depending on the availability of data, scholars have either used the number of patents granted or the number of patent applications (Almeida and Kogut 1997; Jaffe 1989; Jaffe, Trajtenberg, and Henderson 1993; Mukim 2012). Innovation systems literature also emphasizes that local science and technology institutions like universities, colleges, laboratories, etc. are extremely important factors for developing innovative capacity (Arbo and Benneworth 2007). Not only these kinds of institution but others like judiciary and law and order too can be pivotal in creating innovative culture (Subramanian 2007; Hughes 2006).

Most empirical studies are based on the data of USA or European economies. There are very few studies that have examined this subject for developing countries, and none in the Indian context barring Mukim (2012) at the district level and Pradhan (2014) at the state level. Geography came out as the key driver of innovation in India in both the studies (Mukim 2012; Pradhan 2014). The foregoing discussion can be summarized by saying that innovative activity of any region could be dependent on agglomeration forces, local stock of knowledge, knowledge spillover, local infrastructure and institutions.

3. Continuous patent blocks: patents in India

We use patent applications filed as a proxy for the outcome of the inventive process. Data on patent applications are taken from the Indian Patent Office (IPO). The exact location of the applicant is only contained in the weekly journals of IPO, which we use to get the data for the patents filed in various districts and its International Patent Classification (IPC) code. If the district was not given, pin code from the address of the inventor was used to identify the district (first name if there are multiples). Only domestic applicants are included.

The total number of patent applications filed increased from around 5000 in 1999–2000 to 45,444 in 2016–2017. Although the applications from residents have increased steadily over the last decade, the increase is only marginal and the increase in total applications is mostly due to higher applications by foreigners (Figure 1). The share of patents granted to Indian firms increased, and hovered about 20% since the mid-1990s, barring a couple of years where their share increased abruptly. The abrupt increase in the share of Indian patents in 1999–2000 was due to a decline in foreign patents, which could have been due to the amendment of Indian Patent Act 1970 in the same year. It is only in the middle 2000s, the share of patents filled by Indians consistently started rising, and became 30% in 2016–2017.

There is a huge inter-district variation in patent applications. For instance, in 2012–2013, 1092 patents were filed in Mumbai, whereas in 370 districts (out of 626 districts) no applications were filed. Mumbai, Bengaluru, Delhi, Chennai, Hyderabad, Pune, Ahmadabad and Kolkata (as shown in dark red areas in Figure 2) alone account for about 65% of total patent applications, showing a very high level of spatial concentration of innovative activity in India. Another important feature is the presence of continuous patent blocks in

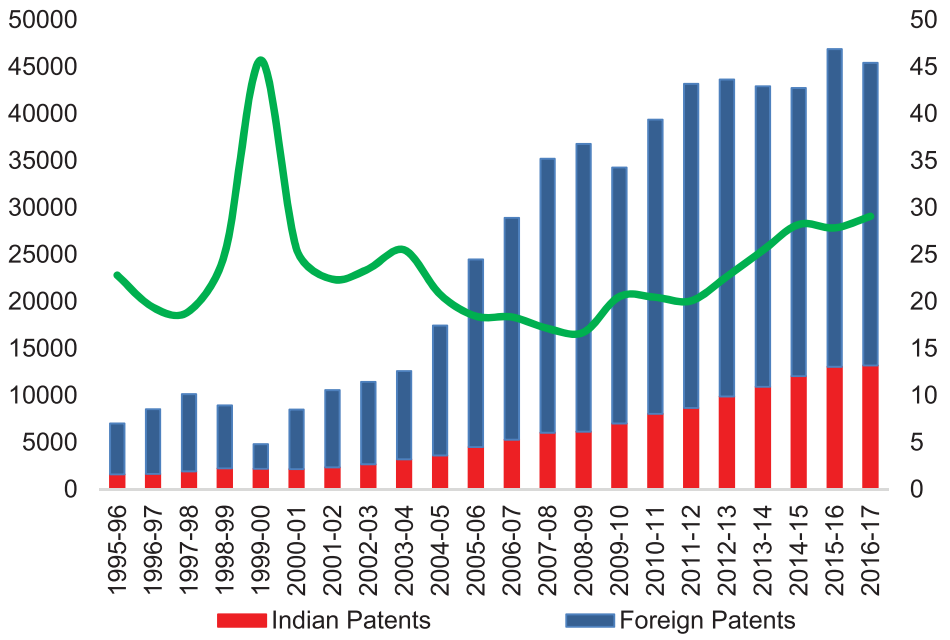

Figure 1. Patent applications in India. Source: IPO reports. (Colour online)

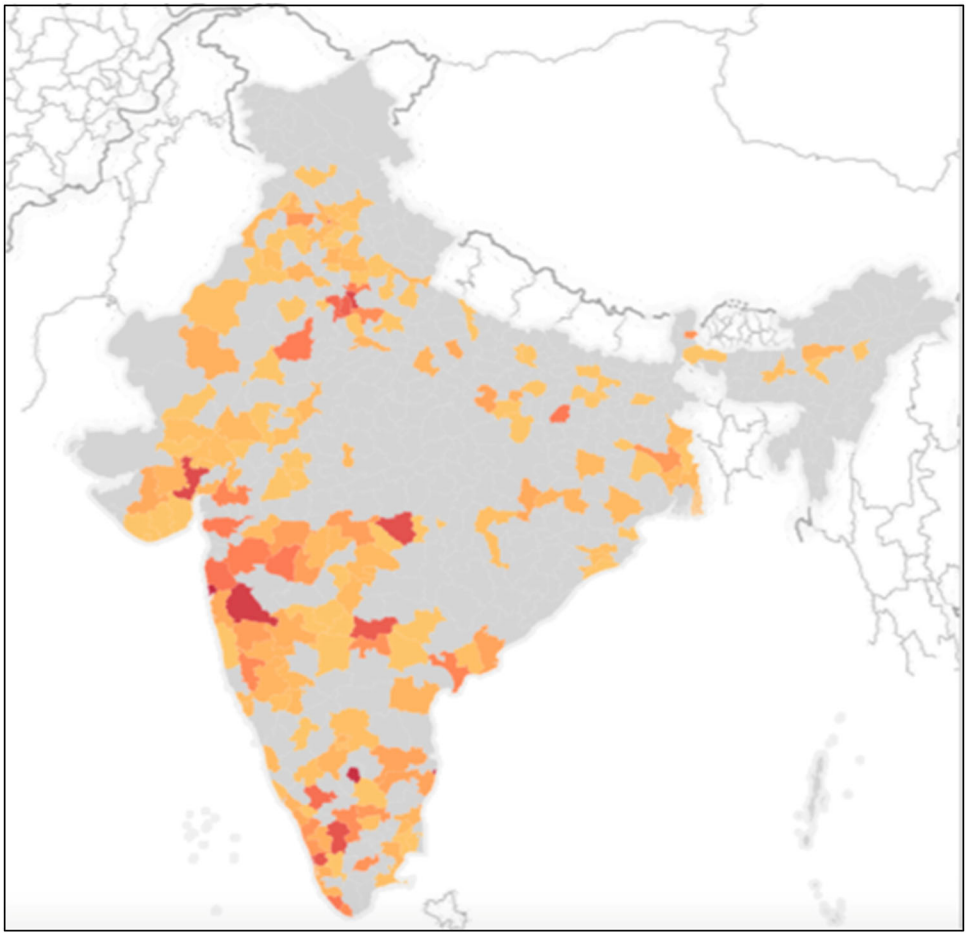

Figure 2. Spatial Pattern of Patent Applications filed in 2012–13. Source: IPO weekly reports. (Colour online)

Note: The area of Kashmir is empty because the data for those districts were not available.

the country with the above-mentioned districts as the core or centre. The dark red areas are the core areas with a concentration of maximum numbers of patent applications; and are surrounded by light red/orangish regions, reflecting that the core areas are surrounded by periphery areas which have less number of patents as compared to the core, but much higher than the rest of the country. The map clearly depicts that there are clusters of innovative activity and there is hardly any standalone district with a high number of applications.

The grey colour shows that no patent applications were filed from that district, and it is evident that northern (topmost), eastern and northeastern parts of India hardly have any patent applications. In the northern area, most of the innovative activity is clustered in and around Delhi and neighbouring areas. Central India also has very few areas with a high number of patent applications. The innovative activity is densely located in southern India, and within it, in the southwest region. The southwest cluster is around Mumbai

and Pune districts. Besides, there is heavy clustering in the western part of the country, which is around Ahmadabad district.

4. Empirical model and methodology

We use the extended version of the Griliches-Jaffe regional KPF to study the relationship between innovation and economic geography, using other local characteristics as control variables.

Doing this analysis at the firm level neglects the geographical and sectoral spillover of knowledge and other inputs. The link between knowledge inputs and innovative output becomes stronger with the aggregation of the unit of observation, from firm to industry to region. Some studies have used states (Jaffe 1989; Acs, Fitzroy, and Smith 1999) or Metropolitan Statistical Areas (Anselin, Varga, and Acs 1997, 2000), or districts (Keeble and Wilkinson 1999; Piergiovanni and Santarelli 2001) as their unit of study. State or nation may not be the right level of aggregation for our analysis as the agglomeration benefits like knowledge spillovers, common labour pool, lower transportation costs, bigger markets, etc. may not be captured fully either due to multiple levels of clustering within the state (or nation) or decay of these benefits in large geographical areas. Krugman (1991) emphasized that states aren't really the right geographical unit. Also, in the Indian context, it does not make much sense to compare states like Uttar Pradesh which has a population of 199 million and area of 241,000 km^2, with Chandigarh only 1,055,000 people and area of 114 km^2 and with city states such as Delhi. Thus, a better unit of observation for this study is a district.

Schumpeter (1939) pointed out innovation and inventions are not the same things. However, data constraints make it difficult to segregate both and hence the terms are used interchangeably. Most empirical studies measure technological innovation in one of the following three ways: by the inputs used in the innovation process, such as R&D expenditure; by intermediate outputs of the innovation effort, such as the number of patents; or by some final measure of innovative work, such as the count of new product announcements. Most empirical studies use patent data for measuring innovations, primarily due to the ease of data availability and also because patents are the outcome of innovative process. However, this approach is criticized on the grounds that it is just the first stage of innovation, i.e. invention. Nevertheless, the patent application represents the belief in economic value of new idea which is created by consuming valuable resources. Also, two important works (Mukim 2012; Pradhan 2014) in the Indian context studying innovation use patent application data. Considering all these things and data availability, we quantify innovative activity by the number of patent applications filed in a district.

Whilst patent applications data exist for many years, the data on district-level variables, which are taken from the National Sample Survey Organization (NSSO) and census, are available only for 2011–2012 in the last decade. Hence for this analysis, the choice of year of study is limited to 2012–2013. In this paper, we use patent data of 2012–2013 and all the independent variables for the year 2011–2012 to control for endogeneity in the model.

Anselin et al.(1997), Feldman and Audretsch (1999), Del Barrio-Castro and García-Quevedo (2005) and Ponds et al.(2010) have used count data model for similar analysis,

following which we also use count data model within KPF. We estimate the following two models:

$$\begin{aligned} Patents\ jt = {} & \beta 0 + \beta 1(R\ D\ jt - 1) + \beta 2(LI\ jt - 1) + \beta 3(HI\ jt - 1) \\ & + \beta 4(Access\ j1t - 1) + \beta 5(Accessj2t - 1) + \beta 6(General\ Education\ jt - 1) \\ & + \beta 7(Technical\ Education\ jt - 1) + \beta 8(Electricity\ jt - 1) + \beta 9(Banking\ jt - 1) \\ & + \beta 10(Institution\ indexjt - 1) + \beta 11(State\ FE\ jt - 1) + \beta 12(District\ FE\ jt - 1) \\ & + ejt \qquad Model\ I \end{aligned}$$

$$\begin{aligned} Patents\ jt = {} & \beta 0 + \beta 1(LI\ jt - 1) + \beta 2(HI\ jt - 1) + \beta 3(Access\ j1t - 1) \\ & + \beta 4(Access\ j2t - 1) + \beta 5(R\&D\ jt - 1) + \beta 6(Infrastructure\ index\ jt - 1) \\ & + \beta 7(Institution\ Indexjt - 1) + \beta 8(State\ FE\ jt - 1) + \beta 9(District\ FE\ jt - 1) \\ & + ejt \qquad Model\ II \end{aligned}$$

where j refers to the districts, R&D refers to research and development expenditure by the firms in the district in the last five years, localization index (LI) is a measure of concentration, Herfindahl index (HI) is a measure of industrial diversity in district, access indices refer to R&D activity in the proximate districts, general education is a measure of educated population in the district, technical education refers to people with technical skills, banking refers to proportion of households availing banking services in the district, electricity refers to the proportion of households having access to electricity in the district, and Infrastructure index and Institution Index are the district-level indices.

Data for economic geography indicators and infrastructure are taken from NSSO's 68th employment–unemployment survey and Census 2011. Access variable is created using data on R&D expenditure of industries for last five years from CMIE Prowess and the orthodromic distance is calculated using latitude and longitude data of districts available from Community Created Maps of India.[1] The effect of R&D spending has a gestation period, so it is used as a stock variable.

'Localization economies (LI_j)', an index of concentration, is measured as the proportion of sector k's employment in district j as a share of total employment of sector k in the country. If more than one sector exists in a district, then the employment in all those sectors is added and taken as a share of total employment of those sectors in the country. The higher this value, the higher the expectation of intra-industry concentration benefits in the region.

$$LI_j = \sum_k \frac{Employment\ kj}{Employment\ k}$$

'Herfindahl index (HI_j)', which measures economic diversity in the region is measured as the sum of squares of employment shares of all industries in district j. The largest value for Herfindahl index is 'one', when the entire regional economy is dominated by a single industry.

$$HIj = \sum \left(\frac{Employment\ kj}{Employment\ j} \right) \hat{2}$$

The third economic geography variable is 'access', which is defined as the sum of R&D

expenditure in the last five years in the neighbouring districts divided by the distance between the district j and the district in which R&D activity is taking place. The rationale of this indicator is because proximity to other firms and industries undertaking R&D activities has a positive externality due to the knowledge spillovers. We consider the districts within orthodromic distance of 200 km and then 200–500 km from district *j*. The idea behind taking the districts in two ranges is to assess how the spillover impacts of R&D change as the distance increases. The orthodromic distance, also called as great-circle distance, is the shortest distance between two points on the surface of a sphere measured along its surface. Since the districts have boundaries rather than a point, first, the centroids of the district boundaries are calculated and then they are used to find out the latitude and longitude and then the distance of each district from all other 625 districts is calculated. In the formula below, d_{mj} refers to the orthodromic distance between two districts, *j* (the one under study) and *m* (the neighbouring district).

$$Access\ j1 = \Sigma \frac{R\&D\ Expenditure\ m}{d_{mj}(d_{mj} < 200)}$$

$$Access\ j2 = \Sigma \frac{R\&D\ Expenditure\ m}{d_{mj}(200 < d_{mj} < 500)}$$

Apart from these economic geography variables, we use some control variables which, in theory, impact innovative activity in any region, such as stock of R&D expenditure, human capital endowments and infrastructure variables. It is particularly important to control for local inputs into innovative activity. In Model I, we include the share of the population with a higher education (defined as a high school degree or more) as a proxy for the general quality of human capital, as it is expected to have a positive impact on knowledge creation. We also include the proportion of the population that possesses a degree in a scientific/technical subject, as skilled workers endowed with a high level of human capital are expected to have a positive effect on the innovative activity. Furthermore, we use infrastructure variables including percentage of households with electricity connection and access to banking services. In Model II, we use a district infrastructure index created using Principal Component Analysis based on variables related to physical infrastructure like road, water; social infrastructure like school, hospital and financial infrastructure like banks and other credit institutions. The variables used here are scaled to area or population to account for vast differences within the districts. Institutions are expected to have a positive relationship with innovation. To explicitly capture the role of institutions, we create an index using Principal Component Analysis comprising nine variables which capture different aspects of institutions like Law & Order, judiciary, fiscal management, etc. The detailed list of indicators used for both indices is placed in the appendix.[2] A priori, all the explanatory variables in the model are expected to have a positive relationship with the dependent variable, barring HI which is expected to have a negative sign.

The broad features of the data are as follows: first, the mean number of patent applications per district is 9.04 and the standard deviation is 60.5, which means that the standard deviation is over 6.5 times the mean (Table 1); second, the top 13 districts account for about 70% of the total patents and 370 out of 626 (60%) districts have 0 patent application. The mean patent application in non-zero category is 22.1 and the standard deviation is 93.2, i.e. about two times the mean in the non-zero patent application districts.

Table 1. Basic characteristics of patent application data.

Variable	Observations	Mean	Std. Dev.	Min.	Max.
Patent count	626	9.04	60.5	0	1092
Patent count > 0	256	22.1	93.2	1	1092

Poisson distribution model could be applied, which requires the assumption of mean variance equality and our data do not fulfill this assumption (Table 1). Also, the frequency of zero patent count is in excess of what would be expected in the Poisson model. These characteristics imply that we need to factor in over-dispersion and excess of zeroes in the data while choosing a model.

Over-dispersion can be taken care using a negative binomial model, which does not require mean variance equivalence. We estimate the over-dispersion parameter from the negative binomial model, which is statistically different from zero[3] indicating over-dispersion in the data set.

Excess zeroes can be tackled using a two-stage process. In the first stage, logit or probit model is used to distinguish between zero and positive counts, and then in the second stage we use a zero-truncated Poisson or a zero-truncated negative binomial model for positive counts. If the two processes turn out to be the same, then we have the standard count model. The problem of excess zeroes can also be accounted for by assuming that the data come from two separate populations, one where the number of patents is zero, and another where the population has a Poisson distribution. The distribution of the outcome is then modelled in terms of two parameters – the probability of average patent application is always zero is one group and not always zero in another.

5. Results

We use Incidence Rate Ratio (IRR) for all models to interpret the results. The regression coefficients in count data models are differences between the logs of expected count. But the difference of two logs is also equal to their quotients; therefore, we could also interpret the parameter estimate as the log of the ratio of expected counts. IRR values can be understood as follows: if the value of an independent variable increases by a percentage point, the rate ratio for the count of patent application would be expected to increase or decrease by the factor of IRR. An IRR equal to 1 implies no change, less than 1 implies a decrease and more than 1 implies an increase in the ratio. Tables 2 and 3 present the IRR for different techniques used to estimate our models.

HI, which captures the degree of industrial diversity, has the expected (negative) relationship. Less than one value of IRR for HI implies higher employment concentration by one industry or lower industrial diversity has a negative impact on innovative activity in a district and vice versa. Localization index, which captures the effect of concentration has IRR greater than one and is highly significant across both models, which is evident towards positive association between agglomeration and innovation. Overall, both these economic geography indicators show that agglomeration economies have a positive relationship with innovation. Other geographic variable in the model is access variable, which captures knowledge spillover has IRR greater than one and significant in most cases. Ideally, the IRR for access variable for distance between 200 and 500 km should be smaller than

Table 2. IRRs of Model 1.[a]

Variables	(1) Poisson	(2) NB	(3) ZIP	(4) ZINB	(5) Poisson	(6) NB	(7) ZIP	(8) ZINB	(9) Poisson	(10) NB	(11) ZIP	(12) ZINB
Patent												
HI	0.81***	0.59***	0.93***	0.70*	0.79***	0.53***	0.92***	0.62***	0.80***	0.53***	0.93***	0.61***
	(0.02)	(0.06)	(0.02)	(0.09)	(0.02)	(0.06)	(0.02)	(0.08)	(0.02)	(0.06)	(0.02)	(0.08)
LI	1.38***	2.29***	1.25***	1.48***	1.35***	2.17***	1.23***	1.35***	1.37***	2.18***	1.24***	1.34***
	(0.02)	(0.23)	(0.02)	(0.13)	(0.02)	(0.23)	(0.02)	(0.12)	(0.02)	(0.23)	(0.02)	(0.12)
Access 200	0.98	1.42***	0.97	1.27*	1.01	1.63***	0.99	1.52***	1.00	1.63***	0.98*	1.52***
	(0.01)	(0.14)	(0.01)	(0.12)	(0.01)	(0.18)	(0.01)	(0.17)	(0.01)	(0.18)	(0.01)	(0.17)
Access 200–500	1.01***	1.24***	1.02	1.13***	1.04***	1.27***	1.03**	1.16***	1.03***	1.26***	1.02*	1.15***
	(0.01)	(0.04)	(0.01)	(0.05)	(0.01)	(0.05)	(0.01)	(0.04)	(0.01)	(0.05)	(0.01)	(0.04)
R&D Stock	1.22***	2.13***	1.27***	2.13***	1.22***	2.21***	1.27***	2.19***	1.22***	2.22***	1.27***	2.20***
	(0.01)	(0.41)	(0.01)	(0.38)	(0.01)	(0.43)	(0.01)	(0.38)	(0.01)	(0.43)	(0.01)	(0.39)
Electricity	2.74***	1.93***	1.77***	1.48	2.57***	1.58***	1.75***	1.16	2.64***	1.61***	1.81***	1.18
	(0.12)	(0.22)	(0.09)	(0.23)	(0.12)	(0.18)	(0.09)	(0.17)	(0.12)	(0.18)	(0.09)	(0.17)
Banking	1.52***	1.25**	1.48***	1.17	1.59***	1.50***	1.50***	1.42**	1.57***	1.49***	1.48***	1.41**
	(0.03)	(0.15)	(0.04)	(0.17)	(0.04)	(0.19)	(0.04)	(0.21)	(0.04)	(0.18)	(0.04)	(0.21)
General Education	1.94***	1.26**	1.92***	1.41***	2.05***	1.39***	1.96***	1.64***	2.02***	1.39***	1.94***	1.65***
	(0.04)	(0.12)	(0.05)	(0.16)	(0.05)	(0.14)	(0.05)	(0.21)	(0.05)	(0.14)	(0.05)	(0.21)
Technical Education	1.19***	1.27	1.18***	1.16	1.16***	1.03	1.18***	0.97	1.17***	1.03	1.19***	0.97
	(0.01)	(0.14)	(0.01)	(0.11)	(0.01)	(0.11)	(0.01)	(0.09)	(0.01)	(0.12)	(0.01)	(0.09)
Institution index	1.01**	0.84	1.07*	0.93								
	(0.03)	(0.11)	(0.03)	(0.13)								
State F.E.					Y	Y	Y	Y				
District F.E.									Y	Y	Y	Y
Constant	0.39	0.17	1.27***	0.91***	1.14**	0.64*	3.05***	1.03	1.26***	0.68*	3.28***	1.09
	(0.27)	(0.13)	(0.33)	(0.32)	(0.07)	(0.16)	(0.19)	(0.29)	(0.07)	(0.16)	(0.20)	(0.29)
AIC	7529.83	2095.14	6272.06	2044.70	7560.44	2102.26	6286.44	2043.88	7579.07	2102.66	6293.18	2043.39
BIC	7578.47	2148.2	6369.34	2146.39	7609.19	2155.44	6383.93	2145.80	7627.75	2155.84	6390.68	2145.31
Log Likelihood	−3753.9	−1035.6	−3114.0	−999.3	−3769.2	−1039.1	−3121.2	−998.94	−3778.5	−1039.3	−3124.5	−998.6
Observations	615	615	615	615	621	621	621	621	621	621	621	621

Robust standard error in parentheses.
***$p < .01$, **$p < .05$, *$p < .1$
[a]Other models and combinations were also tried. Results will be made available on demand.

Table 3. IRRs of Model 2.[a]

Variables	(1) Poisson	(2) NB	(3) ZIP	(4) ZINB	(5) Poisson	(6) NB	(7) ZIP	(8) ZINB	(9) Poisson	(10) NB	(11) ZIP	(12) ZINB
Patent												
HI	0.39***	0.40***	0.48***	0.54***	0.35***	0.37***	0.45***	0.48***	0.35***	0.36***	0.45***	0.47***
	(0.00)	(0.04)	(0.01)	(0.06)	(0.01)	(0.04)	(0.01)	(0.05)	(0.01)	(0.04)	(0.01)	(0.05)
LI	1.38***	2.47***	1.17***	1.60***	1.23***	2.13***	1.11***	1.46***	1.24***	2.13***	1.11***	1.46***
	(0.01)	(0.26)	(0.01)	(0.14)	(0.01)	(0.23)	(0.01)	(0.14)	(0.01)	(0.23)	(0.01)	(0.14)
Access 200	1.19***	1.95***	1.12***	1.49***	1.23***	2.21***	1.14***	1.67***	1.23***	2.23***	1.14***	1.69***
	(0.01)	(0.19)	(0.01)	(0.13)	(0.01)	(0.23)	(0.01)	(0.15)	(0.01)	(0.24)	(0.01)	(0.16)
Access 200 to500	1.14***	1.26***	1.06***	1.10**	1.14***	1.27***	1.06***	1.11**	1.13***	1.26***	1.05***	1.11**
	(0.01)	(0.05)	(0.01)	(0.05)	(0.01)	(0.06)	(0.01)	(0.05)	(0.01)	(0.06)	(0.01)	(0.05)
R&D Stock	1.35***	2.26***	1.37***	2.48***	1.33***	2.45***	1.35***	2.54***	1.32***	2.46***	1.35***	2.55***
	(0.01)	(0.39)	(0.01)	(0.42)	(0.01)	(0.42)	(0.01)	(0.42)	(0.01)	(0.42)	(0.01)	(0.42)
Infra index	1.16***	0.94	1.20***	1.03	1.14***	0.92	1.18***	1.02	1.14***	0.92	1.18***	1.03
	(0.01)	(0.08)	(0.02)	(0.11)	(0.01)	(0.08)	(0.02)	(0.11)	(0.01)	(0.08)	(0.02)	(0.11)
Institution index	1.64***	1.40***	1.48***	1.15								
	(0.03)	(0.15)	(0.03)	(0.15)								
State fixed Effect					Y	Y	Y	Y				
District fixed Effect									Y	Y	Y	Y
Constant	1.16***	0.58***	1.94***	1.22***	1.55***	0.70*	4.59***	1.36	1.72***	0.75	4.91***	1.44
	(0.30)	(0.11)	(0.35)	(0.13)	(0.07)	(0.13)	(0.21)	(0.32)	(0.07)	(0.14)	(0.22)	(0.32)
AIC	10387.9	2145.2	8281.6	2100.4	10643.19	2141.778	8503.69	2092.783	10707.9	2143.35	8526.72	2094.46
BIC	10422.7	2184.4	8351.4	2174.6	10678.14	2181.092	8573.58	2167.043	10742.8	2182.67	8596.61	2168.72
Log likelihood	−5185.9	−1063.6	−4124.8	−1033.2	−5313.6	−1061.8	−4235.8	−1029.4	−5346	−1062.7	−4247.3	−1030.2
Observations	578	578	578	578	583	583	583	583	583	583	583	583

Robust standard error in parentheses.
***$p < .01$, **$p < .05$, *$p < .1$
[a]Other models and combinations were also tried. Results will be made available on demand.

Table 4. Fit from all the variants of Model 1 (with district fixed effects).

Variable	Mean	Std. Dev.	Min.	Max.
Data	0.5910	0.4920	0	1
Poisson	0.3264	0.3518	0	0.9916
NB	0.3887	0.3456	0	0.9886
ZINB	0.3927	0.3779	0	.9992
ZIP	0.3913	0.3855	0	.9987

within 200 km. The results for Model 2 confirm this; however are mixed for Model 1. This could be due to the fact that different types of infrastructure are reinforcing in nature and the infrastructure index in Model 2 is based on this logic, which may explain the more robust results for Model 2.

Stock of R&D expenditure in the district has IRR greater than one which remains significant in both the models. In Model 1, IRR is greater than one and significant for general education and technical education as expected. Electricity connection and banking services, which represent infrastructural facilities in districts, have IRR greater than 1 in almost all cases. In Model 2, we have used the infrastructure index instead of individual infrastructure variables and it also has IRR greater than one and is highly significant (Table 3). Furthermore, IRR coefficient of Institution index is greater than 1 and significant in almost all cases in Model 2.[4] The results confirm that institutions, infrastructure and the socio-economic conditions also play an important role in determining innovative capabilities of a region.

Although all the regressors have been lagged, there could remain endogeneity concerns that could bias the coefficients. One potential problem could also be omitted variable bias; however, the introduction of district or state fixed effects would effectively control for the effect of time-invariant unobservable variables at the level of the district or state. Hence we introduced the state and district time fixed effects and find the coefficients and IRR of agglomeration, R&D expenditure, infrastructure and institution variables still remain highly significant. This shows that the results are robust and that even accounting for the differences between districts and states, the agglomeration economies explain higher innovation in districts. We also want to highlight the point that the coefficients of other variables when Institution index is used are almost the same as that when we use the state or district fixed effect. This point towards the fact that the state/district fixed effect have successfully captured almost the entire effect of institutions as well; though by using a specific index for institutions, we are able to project the crucial role of institutions.

We use Likelihood ratio (LR), Bayesian Information Criteria (BIC) and Akaike's Information Criteria (AIC) to select the best-fit model. The model with highest value (closest to zero as it is a negative number) of LR and the lowest value of AIC and BIC is considered the best model (Tables 2 and 3). On the basis of these criteria, negative binomial model

Table 5: Fit from all the variants of Model 2 (with district fixed effects)

Variable	Mean	Std. Dev.	Min.	Max.
Data	0 .5910	0.4920	0	1
Poisson	0.1466	0.1957	0	0.7638
NB	0.3038	0.2871	0	0.9274
ZINB	0.3081	0.3312	0	0.9963
ZIP	0.2428	0.3138	0	0.9760

and zero-inflated negative binomial model are better, specifically, zero-inflated negative binomial fits the best (Tables 4 and 5).

6. Conclusion

One important conclusion of the paper is the significance of agglomeration for innovation. We find that not only agglomeration economies but also other locational characteristics like institutions, infrastructure and socio-economic endowment also play an important role in promoting innovative activity. We also find that not only local generation of knowledge but also its spillover are also crucial for higher innovative output. Perhaps, this holds the key for absorbing and then innovating on the knowledge transferred from the neighbouring regions. These results reiterate the importance of geography in promoting innovation.

There is high concentration of innovative activity in India. Huge spatial disparity is found in the Indian districts in terms of innovative activity. There are continuous regions of high innovative activity, with core areas surrounded by periphery areas. The core is always the important city of the region, which highlights the significance of cities for innovation.

Innovation is crucial for growth, and so its determinants become relevant for policy formulation. As the results show, the potential to absorb spillovers of knowledge depends on the level of infrastructure and human capital in that region, so improving these becomes very important. This brings to light the role of state in innovation in a region, as the state is heavily involved in the provisioning of basic infrastructure, basic education etc. The positive externalities of the development work in terms of higher innovation should be accounted for by the state while taking the decision of investing in these. Equally important implication is that policy for regional development should be framed keeping an eye on how to benefit from the agglomeration economies in that region. For example, Governments can have policies which promote economic clusters, making provision of targeted infrastructure easier, or governments could put in place incentives for firms to spend on research-related activities, because it not only helps the firm but also has positive externality on other firms.

To make innovation policy more effective, policy makers must also consider the geography of innovation apart from usual focus on subsidy on R&D expenditure. A correct choice of policy instrument would require the policymaker to realize the systemic bottlenecks like poor infrastructure, inadequate skill, absence of knowledge transfer institutions, interaction difficulties, etc. These challenges could be addressed by realizing that innovation is a complex process, where interaction of different players in a given region plays very crucial role. Another crucial factor for fostering innovation environment is to shift the locus of policy making from central to regional level.

Our study is not without limitations, though. First, we have used the data of patent applications, not patents granted. This is a useful measure in the absence of district-level time series data on patents granted, but may not capture the innovative capacity of a region, strictly speaking. Secondly, we have used data on R&D expenditure by firms to capture the knowledge spillover but R&D expenditure in a region is also done by institutions such as universities and government organizations.

Notes

1. www.projects.datameet.org/maps/districts/
2. Principal component analysis technique is used to arrive at the weights to create the infrastructure index. In this method, the sum of squared loadings of each factor is maximized and this is done with the objective that the factor arrived at from PCA explains the maximum possible variation in the data.
3. Value of α is 2.67 (2.06–3.06) in Model 1 and 3.29 (2.65–4.10) in Model 2.
4. For robustness check, we use the Public Affairs Index created by Public Affairs Centre and Governance Index created by Mundle et al. (2012) in a research paper titled 'The Quality of Governance: How Have Indian States Performed?'. The results hold when we use either of these indices to capture institution as well.

Acknowledgment

We are beholden to Dr A. J. C. Bose, Shri Ram College of Commerce and Prof. Arup Mitra, Institute of Economic Growth, Delhi for their unflinching encouragement. We are very grateful to the two anonymous reviewers and the editor of the journal for their valuable feedback and suggestions.

Disclosure statement

No potential conflict of interest was reported by the authors.

References

Abramovitz, M. 1979. "Rapid Growth Potential and Its Realisation: the Experience of Capitalist Economies in the Postwar Period." In *Economic Growth and Resources*, 1–51. London: Palgrave Macmillan.

Abramovitz, M. 1986. "Catching Up, Forging Ahead, and Falling Behind." *The Journal of Economic History* 46 (2): 385–406.

Acosta, M., D. Coronado, M. D. León, and M. Á. Martínez. 2009. "Production of University Technological Knowledge in European Regions: Evidence from Patent Data." *Regional Studies* 43 (9): 1167–1181.

Acs, Z. J., D. B. Audretsch, and M. P. Feldman. 1992. "R & D Spillovers and Recipient Firm Size." *The Review of Economics and Statistics* 76 (2): 336–340.

Acs, Z. J., F. R. Fitzroy, and I. Smith. 1999. "High Technology Employment, Wages and University R&D Spillovers: Evidence from US Cities." *Economics of Innovation and new Technology* 8 (1-2): 57–78.

Almeida, P., and B. Kogut. 1997. "The Exploration of Technological Diversity and Geographic Localization in Innovation: Start-up Firms in the Semiconductor Industry." *Small Business Economics* 9 (1): 21–31.

Anselin, L., A. Varga, and Z. Acs. 1997. "Local Geographic Spillovers Between University Research and High Technology Innovations." *Journal of Urban Economics* 42 (3): 422–448.

Anselin, L., A. Varga, and Z. J. Acs. 2000. "Geographic and Sectoral Characteristics of Academic Knowledge Externalities." *Papers in Regional Science* 79 (4): 435–443.

Arbo, P., and P. Benneworth. 2007. "Understanding the Regional Contribution of Higher Education Institutions: A literature Review", OECD Working Paper No. 9, OECD Publishing, Paris.

Asheim, B. R. T. 1996. "Industrial Districts as 'Learning Regions': A Condition for Prosperity." *European Planning Studies* 4 (4): 379–400.

Baldwin, R. E., and P. Martin. 2004. "Agglomeration and Regional Growth." In *Handbook of Regional and Urban Economics*. Vol. 4, 2671–2711. Elsevier.

Bode, E. 2004. "The Spatial Pattern of Localized R&D Spillovers: An Empirical Investigation for Germany." *Journal of Economic Geography* 4 (1): 43–64.

Bottazzi, L., and G. Peri. 2003. "Innovation and Spillovers in Regions: Evidence from European Patent Data." *European Economic Review* 47 (4): 687–710.
Buesa, M., J. Heijs, and T. Baumert. 2010. "The Determinants of Regional Innovation in Europe: A Combined Factorial and Regression Knowledge Production Function Approach." *Research Policy* 39 (6): 722–735.
Del Barrio-Castro, T. D., and J. García-Quevedo. 2005. "Effects of University Research on the Geography of Innovation." *Regional Studies* 39 (9): 1217–1229.
Doloreux, D., and S. Parto. 2005. "Regional Innovation Systems: Current Discourse and Unresolved Issues." *Technology in Society* 27 (2): 133–153.
Feldman, M. P., and D. B. Audretsch. 1999. "Innovation in Cities: Science-based Diversity, Specialization and Localized Competition." *European Economic Review* 43 (2): 409–429.
Fischer, M. M., and A. Varga. 2003a. "Production of Knowledge and Geographically Mediated Spillovers from Universities." *The Annals of Regional Science* 37 (2): 303–323.
Fischer, M. M., and A. Varga. 2003b. "Spatial Knowledge Spillovers and University Research: Evidence from Austria." *The Annals of Regional Science* 37 (2): 303–322.
Florida, R. 1994. "Agglomeration and Industrial Location: An Econometric Analysis of Japanese-affiliated Manufacturing Establishments in Automotive-related Industries." *Journal of Urban Economics* 36 (1): 23–41.
Freeman, C. 1987. *Technology Policy and Economic Performance*. London: Frances Pinter Publisher Ltd.
Fritsch, M., and V. Slavtchev. 2007. "Universities and Innovation in Space." *Industry and Innovation* 14 (2): 201–218.
Glaeser, E. L., H. D. Kallal, J. A. Scheinkman, and A. Shleifer. 1992. "Growth in Cities." *Journal of Political Economy* 100 (6): 1126–1152.
Griliches, Z. 1979. "Issues in Assessing the Contribution of Research and Development to Productivity Growth." *The Bell Journal of Economics* 10 (1): 92–116.
Griliches, Z. 1984. *R&D, Patents and Productivity*. Chicago: University of Chicago Press.
Howitt, P., and P. Aghion. 1998. "Capital Accumulation and Innovation as Complementary Factors in Long-run Growth." *Journal of Economic Growth* 3 (2): 111–130.
Hughes, Alan. 2006. *University-industry Linkages and UK Science and Innovation Policy*. Cambridge: Centre for Business Research, University of Cambridge.
Jacobs, J. 1969. *The Economy of Cities*. New York: Random House.
Jaffe, A. B. 1989. "Real Effects of Academic Research." *The American Economic Review* 79 (5): 957–970.
Jaffe, A. B., M. Trajtenberg, and R. Henderson. 1993. "Geographic Localization of Knowledge Spillovers as Evidenced by Patent Citations." *The Quarterly Journal of Economics* 108 (3): 577–598.
Kantor, S., and A. Whalley. 2009. "Do Universities Generate Agglomeration Spillovers? Evidence from Endowment Value Shocks (No. w15299)." National Bureau of Economic Research.
Keeble, D., and F. Wilkinson. 1999. "Collective Learning and Knowledge Development in the Evolution of Regional Clusters of High Technology SMEs in Europe." *Regional Studies* 33 (4): 295–303.
Krugman, P. 1991. "Increasing Returns and Economic Geography." *Journal of Political Economy* 99 (3): 483–499.
Kumar, N., and K. J. Joseph. 2006. "10. National Innovation Systems and India's IT Capability: Are There any Lessons for ASEAN Newcomers?." Asia's Innovation Systems in Transition, 227.
Lawson, Clive. 1997. "Territorial Clustering and High Technology Innovation: From Industrial Districts to Innovative Milieux," Centre for Business Research Working Paper 54. Cambridge University, Cambridge.
Loesch, A. 1954. *The Economics of Location: Translated from the Second Rev. German Ed. by William H. Woglom with the Assistance of Wolfgang F. Stolper*. New Haven: Yale University Press.
Lucas, R. E. Jr. 1988. "On the Mechanics of Economic Development." *Journal of Monetary Economics* 22 (1): 3–42.

Lucas, R. E. Jr. 1993. "Making a Miracle." *Econometrica: Journal of the Econometric Society* 61 (2): 251–272.

Lundvall, B. A. 1992. *National Systems of Innovation: An Analytical Framework*. London: Pinter.

Lundvall, B. A. 2000. "Nation States and Economic Development. From National Systems of Production to National Systems of Knowledge Creating and Learning." In *Oxford Handbook of Economic Geography*, 353–372. Oxford: Oxford University Press.

Lundvall, B. Å. 2016. "Product Innovation and User-producer Interaction." In *The Learning Economy and the Economics of Hope*, 19–58. London: Anthem Press.

Lundvall, B. A., and S. Borrás. 1998. *The Globalising Learning Economy: Implications for Innovation Policy*. Luxembourg: European Communities.

Lundvall, B. Ä., and B. Johnson. 1994. "The Learning Economy." *Journal of Industry Studies* 1 (2): 23–42.

Lundvall, B.Å., J. Vang, K. J. Joseph, and C. Chaminade. 2009. "Innovation System Research and Developing Countries." In *Handbook of Innovation Systems and Developing Countries: Building Domestic Capabilities in a Global Setting*, edited by B. A. Lundvall, K. J. Joseph, C. Chaminade, and Jan Vang, 1–30. Cheltenham: Edward Elgar.

Marrocu, E., R. Paci, and S. Usai. 2013. "Knowledge Production Function and Proximities. Evidence from spatial Regression Models for the European Regions. " SEARCH Working Paper WP 4/1. January 2013. http://www. ub. edu/searchproject/wp-content/uploads/2013/01/WP-4.1. pdf.

Marshall, A., and M. P. Marshall. 1920. *The Economics of Industry*. London: Macmillan & Co.

Maskell, P., and A. Malmberg. 1999. "Localised Learning and Industrial Competitiveness." *Cambridge Journal of Economics* 23 (2): 167–185.

Mukim, Megha. 2012. "Does Agglomeration Boost Innovation? An Econometric Evaluation." *Spatial Economic Analysis* 7 (3): 357–380.

Mundle, Sudipto, Pinaki Chakraborty, Samik Chowdhury, and Satadru Sikdar. 2012. "The Quality of Governance: how Have Indian States Performed?" *Economic and Political Weekly* 47 (49): 41–52.

Pavitt, K. 2002. "Innovating Routines in the Business Firm: What Corporate Tasks Should They be Accomplishing?" *Industrial and Corporate Change* 11 (1): 117–133.

Piergiovanni, R., and E. Santarelli. 2001. "Patents and the Geographic Localization of R&D Spillovers in French Manufacturing." *Regional Studies* 35 (8): 697–702.

Ponds, R., F. van Oort, and K. Frenken. 2010. "Innovation, Spillovers and University–Industry Collaboration: An Extended Knowledge Production Function Approach." *Journal of Economic Geography* 10: 231–255.

Pradhan, Jaya Prakash. 2014. "The Geography of Patenting In India: Patterns and Determinants." *Metamorphosis* 13 (2): 29–43.

Prakash Pradhan, Jaya. 2011. "Regional Heterogeneity and Firms' R&D in India." *Innovation and Development* 1 (2): 259–282.

Romer, P. M. 1986. "Increasing Returns and Long-run Growth." *Journal of Political Economy* 94 (5): 1002–1037.

Romer, P. M. 1990. "Endogenous Technological Change." *Journal of Political Economy* 98 (5, Part 2): S71–S102.

Schumpeter, J. A. 1939. *Business Cycles*. Vol. 1, 161–174. New York: McGraw-Hill.

Solow, R. M. 1956. "A Contribution to the Theory of Economic Growth." *The Quarterly Journal of Economics* 70 (1): 65–94.

Storper, M., and A. J. Venables. 2004. "Buzz: Face-to-Face Contact and the Urban Economy." *Journal of Economic Geography* 4 (4): 351–370.

Subramanian, Arvind. 2007. "The Evolution of Institutions in India and its Relationship with Economic Growth." *Oxford Review of Economic Policy* 23 (2): 196–220.

Tappeiner, G., C. Hauser, and J. Walde. 2008. "Regional Knowledge Spillovers: Fact or Artifact?" *Research Policy* 37 (5): 861–874.

Veblen, T. 1915. "The Opportunity of Japan." *Journal of Race Development* 6 (1): 23–38.

Appendix

List of indicators used for creating district infrastructure index

1	Private schools per square kilometre (sq km)
2	Government schools per sq km
3	Total road per sq km
4	Hospitals per sq km
5	Government degree college per sq km
6	Private degree college per sq km
7	Government technical college per sq km
8	Private technical college per sq km
9	Water capacity from all sources per capita
10	Non-Agricultural Credit Society per sq km
11	Agricultural Credit Society per sq km
12	Public library per sq km
13	Banks per sq km

Note:

(1) Private and Government schools include primary, middle, secondary and senior secondary schools.
(2) Road include both pucca and kutcha road.
(3) Hospitals include Allopathic, alternative medicines, dispensaries, family welfare centres, maternity and child welfare clinics, nursing homes, Tuberculosis hospitals and nursing homes.
(4) Private and Government Degree College include only art colleges; only commerce colleges; only science colleges; art and science college; art and commerce college; art, science and commerce colleges and law colleges.
(5) Government and private technical colleges include medical colleges, engineering colleges, polytechnic colleges and vocational colleges.
(6) Library includes government library and public library.
(7) Banks include nationalized banks, private banks and cooperative banks.

List of indicators used for creating district infrastructure index

1	Percentage of police strength filled
2	Percentage of district court strength filled
3	Crimes against Scheduled Castes (SC) per lakh SC population
4	Cases pending in courts for more than five years as a ratio of completed cases
5	Fiscal deficit as a percentage of Gross State Domestic Product (GSDP)
6	Development expenditure as a ratio of GSDP
7	Own Tax Revenue
8	Infant mortality rate per lakh population
9	MPhil and PhD students as a ratio of population in age group 20–34