Knowledge Spillovers Within China's System of Research Institutes

By

Renai Jiang, Daniel L. Tortorice, and Zhaohui Xuan

December 2019

COLLEGE OF THE HOLY CROSS, DEPARTMENT OF ECONOMICS
FACULTY RESEARCH SERIES, PAPER NO. 19-10*

Department of Economics and Accounting
College of the Holy Cross
Box 45A
Worcester, Massachusetts 01610
(508) 793-3362 (phone)
(508) 793-3708 (fax)

<https://www.holycross.edu/academics/programs/economics-and-accounting>

* All papers in the Holy Cross Working Paper Series should be considered draft versions subject to future revision. Comments and suggestions are welcome.

Knowledge Spillovers Within China's System of Research Institutes*

Renai Jiang[†], Daniel L Tortorice[‡] and Zhaohui Xuan⁺

December 2019

Abstract

We use a novel, fifteen-year panel dataset on China's system of research institutes to examine the determinates of knowledge production, the role external factors play in increasing research productivity, and the extent to which distance mitigates these spillovers. We find robust evidence that knowledge inputs like R&D personnel increase patenting and publication. External R&D spending in the institute's province and the institute's industry knowledge stock spill over into increased knowledge production. We find that being located on average farther from institutes engaged in similar research reduces the impact of these spillovers. These results have important policy implications as China attempts to increase economic activity away from the coast and aims to improve the productivity of its research institute sector.

JEL Code: O31, O32, O33, L2

Keywords: China, R&D, Research Policy, Knowledge Spillovers

* We gratefully acknowledge the assistance of the Ministry of Science and Technology (MOST) in allowing access to the data and answering our questions on the data. Ranchao (Audrey) Yang provided superb research assistance. We thank Gary Jefferson and members of the Holy Cross Department of Economics for many helpful conversations and suggestions. Any errors are our own.

[†] School of Economics and Finance, Xi'an Jiaotong University, Xi'an 710061, China, jiangrenai@xjtu.edu.cn

[‡] Department of Economics and Accounting, College of the Holy Cross, 1 College St, Worcester MA, 01610, dtortori@holycross.edu, 1 508 793 3873

⁺ Institute of S&T Statistics and Analysis, Chinese Academy of Science and Technology for Development, Beijing 100038, China, xuanzh@casted.org.cn

1 Introduction

A large literature in economics has studied the role that research and development (R&D) can play in promoting economic growth. This literature is specifically relevant in China a country with robust economic growth and levels of R&D comparable to high-income countries like the United States and those of the European Union.

However, detailed empirical studies of R&D and knowledge production are somewhat limited, especially in China. These limitations stem from the difficulty of measuring knowledge as an output, the diffuse nature of R&D as it is carried out by many different entities, and the external effects of knowledge production which can be difficult to account for without explicit connections between research entities.

We address these difficulties by exploiting a novel data set on China's system of roughly 4,000 research institutes. These data are administrative data collected by the Chinese Academy of Science and Technology for Development (CASTED) a division of the Ministry of Science and Technology (MOST) for the purpose of evaluating the large population of research institutes. The data are a panel beginning in 1998 and ending in 2012 and contain very detailed information about each research institute. For example, we observe: annual revenue broken down by type; number and type of employees, equipment and inputs in the production process again disaggregated by type, number of patents applied for and granted by type, number of scientific papers published, the industry in which the institute specializes, and the administrative region code corresponding to the research institute's location.

We exploit these data to examine the factors contributing to knowledge production within China's research institute sector. We examine the role various inputs have in increasing knowledge production. Additionally, we focus on the role of external factors in influencing the productivity of the research institutes, an effect we refer to as the spillover effect. In analyzing these factors, we distinguish between two types of spillover variables. External spillover variables capture activity outside the research institute system, internal spillover variables capture activity within the research institute system

Specifically, for our external factors we use measures of economic activity, R&D spending, and patenting within the research institute's industry and province. For our internal factors we use measures of revenue and R&D expenditure among other research institutes within the research institutes industry and province.

We find that inputs are generally helpful in increasing knowledge production. Institutes with more personnel or equipment and those that use more intermediate inputs receive more revenue and produce more knowledge as measured by patents and scientific papers. Additionally, we find that when these inputs are science and technology (S&T) inputs the research institutes are more productive in terms of receiving more revenue and when these inputs are R&D inputs the institutes publish more scientific papers. Importantly, these results are robust to allowing for institute fixed effects that control for unobservable (time invariant) heterogeneity across the research institutes.

We find robust evidence of positive spillovers to research institute productivity. Based on our external spillover variables we find that when the research institute's industry grows as a fraction of the output in its province it also tends to generate more revenue. We also find that when the research institute's province spends more on R&D or patents more, research institutes also patent more and publish more scientific papers.

For our internal spillover variables, we find that research institutes generate more revenue, controlling for observable inputs and institute fixed effects, when other research institutes that are in both their province and sector generate more revenue. We also find that these institutes patent more and publish papers more when research institutes in their province and sector have spent more on R&D in the past.

A novel aspect of our paper is knowing the location of the research institutes. We use this information to create a measure of how far away each institute is from related institutes. Using this measure, we demonstrate that the positive spillovers from similar research institutes diminish with increased distance from them.

The results in our paper have important policy implications for research initiatives in China. Specifically, they point to difficulties in developing more remote regions and the importance of

promoting knowledge transmission among more distant research institutes. They also reveal inefficiencies in the research institute sector that the government can work to ameliorate.

Our paper relates to three substantial literatures: the literature on knowledge spillovers, the literature on agglomeration economies, and the China-specific research on R&D. We will discuss each literature briefly and contrast our paper with these previous works.

It has long been recognized that the production of knowledge – new ideas, inventions or innovations, differs from the production of other goods and services. Specifically, novel innovations depend more strongly on the work of others in the same field and therefore the productivity of one institution is more likely to impact positively the productivity of another related institution. An early example of this recognition is Griliches (1991) who surveyed the current work on knowledge spillovers. He found rates of return to R&D of between 20-50% after accounting for spillovers. Bernstein and Nadiri (2002) document substantial R&D spillovers across manufacturing firms. Jaffe (1986) found that R&D expenditure by similar firms increased own patenting activity and Jaffe et. al. (1993) found that inventors were more likely to cite patents from inventors that were close to them geographically. This result is echoed by Thompson (2006) who shows that inventors are more likely to cite local inventors than patent examiners. However, Thompson and Fox-Kean (2005) find that while inventors are more likely to cite inventions from their own country, there are insignificant effects at a more disaggregated level. Recent surveys, e.g. Wieser (2005) and Bloom et al. (2018) continue to find an important role for knowledge spillovers in productivity as do meta-analyses like Neves and Sequeira (2018) and Ugar et al. (2020).

University research is also expected to spill over into for-profit R&D activity (Qiu et. al 2017). Jaffe (1989) finds that geographic overlap between research universities and private R&D labs, increase lab productivity. Anselin et al (1997) document larger knowledge production in states where universities and R&D intensive firms are more co-located. Kantor and Whalley (2014) show that endowment shocks increase labor income in the non-education sector.

Foreign direct investment can also be a channel of knowledge spillovers. Branstetter (2000) examines the role that FDI played in facilitating knowledge spillovers to and from Japanese firms

investing in the US. Cheung and Lin (2004) show that FDI is associated with increased patenting in China at the provincial level as does Zhang (2017).

A small number of papers have looked at knowledge spillovers in China. Kuo and Yang (2008) document a correlation between regional R&D spending and regional economic growth rates. Hu et al. (2005) showed that R&D spending by private enterprises can facilitate technology transfer.

Knowledge spillovers explain agglomeration, the tendency for economic activity to be clustered in specific geographic areas. Audretsch and Feldman (1996) show that innovative sectors are more clustered than other economic activity. Ellison et al. (2010) find that knowledge spillovers contribute to industry agglomeration as do Buzzard and Carlino (2013) for R&D labs. Carlino and Kerr (2015) provide an extensive survey of this literature and note importantly, based partially on the evidence of Rosenthal and Strange (2001), that spillovers attenuate with distance.

Our paper differs substantially from previous works on knowledge spillovers. Almost all work has been done in developed countries, specifically the US and European countries. Additionally, the research has focused on for-profit entities like manufacturing firm, R&D labs, or research universities. Little work has been done on a system of government funded research institutes whose main purpose is knowledge production. Furthermore, despite the emphasis in China on research and development, almost all current economic analyses of knowledge production in China have been performed at a substantially aggregated level.

A noticeable body of work establishes the key role that R&D and science and technology play in the Chinese economy. Xie and Freeman (2018) argue that Chinese authors account for almost 20% of scientific journal output. Zhong and Yang (2007) discuss various Chinese reforms and policies to promote R&D and linkages between enterprises and government research institutes. Liu et al. (2011) document that these policies have become multifaceted and more decentralized. Cao et al. (2019) suggest further reforms to improve Chinese R&D output. Motohashi and Yun (2007) show that manufacturing firms outsourced more S&T activity over time from 1996 to 2002 and Chen et al. (2019) document collaboration between Chinese Academy of Science researchers and industry.

Research institutes are specifically mentioned in Zhong and Yang (2007) along with the work of Conroy (1982) who provides a history of the sector. Hu and Matthews (2008) argue that

universities are the key driver of knowledge production but document a growing role for the research institutes after a series of reforms intended to make the research institutes more adaptive to industry needs (see Jiang et al. 2016).

Our paper differs from other work on Chinese research institutes by being substantially more quantitative. First, we work with very detailed data: a panel of 4,000 institutes over 15 years. Second, we have a large number of input and output variables and we are able to directly examine the role distance plays in attenuating spillovers.

The remainder of the paper proceeds as follows. Section 2 describes the data, section 3 describes our empirical strategy, section 4 presents the results, and section 5 concludes.

2 Data

2.1 Key Variables

For our analysis we use data collected by the Chinese Academy of Science and Technology for Development (CASTED) a division of the Ministry of Science and Technology (MOST). As the umbrella ministry charged with the evaluation of China's expansive research institute sector, MOST has authorized CASTED to administer an extensive and annual survey of China's roughly 4,000 research institutes. As the research institutes are required by both law and regulation to respond to the survey in a complete and timely manner, the data are typically comprehensive and reliable. CASTED has graciously allowed us direct access to this data to conduct analysis at their offices in Beijing, China.

The data are a panel of research institutes that begin in 1998 and extend through 2012. The panel is not balanced and we observe both institute entrance into and exit from the database. The vast database contains over 300 variables on the research institutes. A complete documentation of the database would be misplaced here, instead we describe the variables that we use in our analysis.

We have various measures for the research institutes that we can loosely describe as output. The first of these measures is total revenue, which represents income from all sources. Total revenue is comprised of three components which we refer to as science and technology (S&T) revenue,

production and management (P&M) revenue and other revenue. S&T revenue represents revenue from activities related to science and technology. These activities can range from income from licensing patents to research grants the institute receives. One subset of S&T revenue is government grant revenue, which represents research funding the institute receives from the local, provincial or national government. However, research institutes can also receive S&T income from technological development, licensing, consulting and other services which we refer to as non-government grant revenue. Finally, P&M revenue captures revenue the research institute receives from the production and sale of products.

Unlike a manufacturing firm or a similar type of business, output of a research institute is more difficult to measure. Indeed, their end goal is often the production of knowledge which is both hard to measure and hard to value. To capture knowledge output we use data on patenting and publishing by the research institutes. We have data on number of patents applied for and also the number of patents granted. CASTED also collects data on the number of these patents which are invention patents, a more novel level of innovation, compared to a utility patent which is more of a process innovation or adaptation of an already established invention. We also have data on the number of research papers published by the institutes' employees.

Our second set of variables are our input measures. We use data on the total number of employees at the research institute. We can decompose employment into three type of employees: production and management personnel, science and technology personnel and other personnel. We also have data on a subset of S&T personnel, the number of R&D personnel. We know the amount spent on S&T related equipment and intermediate inputs (e.g. material inputs, conference expenses) used in any production activity the institute is engaged in. Finally, we have a catch all total expenditure variable that captures all spending including wages for personnel, equipment costs and the costs of intermediate inputs. We use data on total expenditure overall and total expenditure broken up into S&T expenditure, R&D expenditure and P&M expenditure.

Finally, we know some important institute specific information which we exploit. Institutes are classified by industry based on the most likely commercial applications of their research output.

Institute are grouped into 45 industry categories.¹ Additionally, we have location data on these institutes. Specifically, we know the administrative region code in which they are located. We have six-digit administrative region codes which correspond to a finely disaggregated geographic area. These areas can be district within a populous city or cities or counties in less dense areas.

Specifically, the rule for the six-digit administrative region codes is as follow: the first two numbers represent provinces (autonomous regions, municipalities directly under the Central Government), the third and fourth numbers represent regions (cities, prefectures, alliances), and the fifth and sixth numbers represents counties (districts, cities, banners). For example, in 440106, 44 represents Guangdong province, 01 represents Guangzhou city, 06 represents the district of Tianhe with in Guangzhou city. Similarly, in 440222, 44 represents Guangdong province, 02 represents Shaoguan city within Guangdong province, 22 represents the Shixing county with in Shaoguan city.

Based on the R&D expenditure and location data, we then calculate two important variables for our analysis. The first is the research institute's knowledge stock and the second is a measure of a research institutes distance from its relevant knowledge stock.

2.2 Calculating a knowledge stock

Investment in R&D may not have an immediate payoff in terms of increased revenue, patenting or publication. In fact, the effects of R&D may take many years to result in increased output. Additionally, R&D creates a stock of knowledge that can increase productivity for many years to come. To address both the lagged impact of R&D expenditure and its potential persistent affects, consistent with the literature, e.g. Bernstein and Nadiri (2002), Hall et al. (2010), we rely on a knowledge stock variable which we calculate from the data.

Specifically, for each research institute $\{i\}$ we can calculate a knowledge stock at time $\{t\}$ as:

¹ At times it is useful to aggregate these industries by sector. The sectors we use are: Agriculture, Mining, Low-tech, Med-low tech, Med-high tech, High-tech, High-tech service, Sanitation, Electric and Gas production, and Other. The appendix documents how we assign industries to sectors.

$$K_{i,t}^{R\&D} = \sum_{j=0}^{\infty} (1 - \delta)^j R\&D \text{ Expenditure}_{i,t-j}$$

We use a depreciation rate of $\delta=0.15$ based on the review of the literature from Hall et al (2010). Since our R&D expenditure data begin in 1998, we use the perpetual inventory method to calculate an initial knowledge stock. This method assumes that R&D expenditure in the past grew at the rate g (where g is the average growth rate of R&D in the sample) and depreciated at the rate δ in the past as well.²

2.3 Calculating the distance between institutes

We know the administrative code where the institute is located. We can therefore calculate the distance between any two research institutes (assuming the institute is centered in the region corresponding to the administrative code). First, we map administrative code to GPS coordinates (latitude and longitude in degrees) using Google maps. Then we can calculate the distance between two institutes $\{d_{i,j}\}$ using the Haversine formula [Sinnott (1984)]. For this method let:

$$a = \sin^2(\Delta\phi/2) + \cos \phi_1 \cdot \cos \phi_2 \cdot \sin^2(\Delta\lambda/2)$$

$$c = 2 \cdot \text{atan2}(\sqrt{a}, \sqrt{1-a})$$

$$d = R \cdot c$$

where ϕ is latitude in radians, λ is longitude in radians, and R is the radius of Earth (6,371km).

We convert degrees to radians using the following formula: Radians = Degrees* π /180.

2.4 External Data

As part of the analysis we also consider spillovers from outside the research institute sector. For these data we rely on the China Statistical Yearbook on Science and Technology for provincial

² This assumption implies that $K_{i,1998}^{R\&D} = R\&D\text{Expenditure}_{i,1998} \frac{1+g}{g+\delta}$.

data on R&D expenditure and patenting.³ We use data from the China Statistical Yearbook for data on provincial value added in industry and regional gross domestic product.⁴

We obtain data from the China Statistical Yearbook for data on the size of the provincial labor force in 1998-2010, and the Provincial Statistical Yearbook for data in 2011-2012.⁵

3 Econometric Framework

For our econometric analysis we model each research institute's output with a Cobb-Douglas production function. Output will depend on labor inputs, intermediate inputs, and spillovers from other related research institutes.

Specifically, research institute $\{i\}$ has output at time $\{t\}$ which is given by a production function:

$$Y_{i,t} = A_{i,t} L_{i,t}^{\alpha} X_{i,t}^{\beta} [K_{i,t}^o]^{\gamma}$$

Here $L_{i,t}$ is labor input, $X_{i,t}$ are other inputs, and $K_{i,t}^o$ are inputs and production of other related institutes which, we call the spillover variable.

Taking logs, we arrive at the following estimation equation:

$$y_{i,t} = \eta_i + \lambda_t + \alpha l_{i,t} + \beta x_{i,t} + \gamma k_{i,t}^o + v_{i,t}$$

here we are implicitly assuming that the productivity level can be broken up into a common year effect, a time invariant institute fixed effect, and an unobservable institute-year specific shock. To interpret these coefficients, note that α is the elasticity of output to labor input, β is the elasticity

³ For example, the China Statistical Yearbook on Science and Technology 2013, pp7, Table 1-7 contains the data on Intramural Expenditure on R&D by region (2012) and pp185, Table 7-7 contains data on Three Kinds of Domestic Patent Applications Accepted by Region (2012).

⁴ China Statistical Yearbook 2013, Table 2-16, Regional Gross Domestic Product by Industrial Sector. Since industrial sector does not match exactly with sector in our research institute data, we match institutes to industrial sector by choosing the closest available option.

⁵ China Statistical Yearbook 2011, Table 4-4, Employee Labor Force in Three Industries by Region.

of output to other inputs and γ is the elasticity of output to the inputs and output of other related institutes.

Since we are interested in the impact that distance has on the ability of a research institute to receive positive spillovers from related institutes, we consider the interaction of the spillover variable with distance. This modification then gives our main estimation equation:

$$y_{i,t} = \eta_i + \lambda_t + \alpha l_{i,t} + \beta X_{i,t} + \gamma k_{i,t}^0 + D_{i,t} * k_{i,t}^0 + D_{i,t} + v_i,$$

In this formulation, the spillover variable is calculated as

$$k_{i,t}^0 = \sum_{j \neq i}^N x_{j,t}$$

where $x_{j,t}$ is an input or output measure for a related research institute. (The concept of related research institutes will be defined below.) The spillover variable is the sum across all related research institutes, of a specific input or output variable, examples of which will be given below.

We then calculate for each research institute its distance from related research institutes by taking the average physical distance weighted by the input/output measure for the related research institutes.

$$D_{i,t} = \frac{1}{X_t} \sum_{j \neq i}^N d_{i,j} x_{j,t}$$

where $X_t = \sum_{j \neq i}^N x_{j,t}$

Next, we describe the various measures we use for outputs, inputs and spillover variables. For this section it is helpful to explain the analysis in two parts. The first part consists of regressions that use monetary variables like revenue as an output measure; the second part consists of regressions that use count variables like patents which are a more direct measure knowledge output.

For our more economic output measures we consider total revenue, S&T revenue, P&M revenue and government grant revenue. For our labor input variables, we use total personnel, the fraction of total personnel that are S&T personnel and the fraction that are P&M personnel. For our other input variables, we use S&T equipment and S&T intermediate inputs.⁶

For our spillover variables we consider spillover variables both external to the system of research institute and internal to the system of research institute. Our first external variable is total R&D expenditure within the research institute's province. To control for the large variation in size across provinces we divide this number by the size of the provincial labor force. Our second external spillover variable is total industry sales/RGDP by province. For each research institute this variable is the fraction of the domestic product of its province that comes from output in its industry. Our first internal variable is total revenue, where we take as the relevant comparison group other research institutes in the same industry. Our second internal spillover is again total revenue but here the comparison group is research institutes within the same sector and province as the research institute.

For our more knowledge-based output measure we use a variety of patenting measures: total patents applications, total patents granted, total utility patents, total invention patents and scientific papers published. Our input measures consist of total personnel, S&T personnel, the fraction of total personnel that is R&D personnel, S&T equipment, S&T intermediate inputs and the fraction of S&T inputs and equipment that are R&D equipment.

Again, we examine a variety of spillover variables both external and internal to the research institute system. For our external variables we use total R&D expenditure within the research institute's province scaled by the size of the provincial labor force. We also use total patents within the research institute's province scaled by provincial population. For our internal spillovers we rely on our knowledge stock variables. The first knowledge stock variable is based on R&D expenditure for the institute's industry. The second knowledge stock variable is based on R&D

⁶ Since inputs may not have an immediate impact on output, these input measures are calculated as the averages over the previous three years.

expenditure within the institute's industry and province. Finally, we also examine the impact of the number of research institutes within the research institute's industry and province.

Patenting and publishing are sporadic activities. In fact, many research institutes do not patent or publish in a given year. To account for this probability mass at zero we use a Poisson estimator. For this estimator the expected number of patents and papers is given by:

$$\lambda_{i,t} = \exp\{\eta_i + \lambda_t + \alpha l_{i,t} + \beta X_{i,t} + \gamma k_{i,t}^0 + D_{i,t} * k_{i,t}^0 + D_{i,t} + v_{i,t}\}$$

and the probability of observing a patenting or paper outcome $y_{i,t}$ is given by the Poisson pdf:

$$P(y_{i,t} | \lambda_{i,t}) = \frac{\lambda_{i,t}^{y_{i,t}} e^{-\lambda_{i,t}}}{y_{i,t}!}$$

We can then estimate the model by maximum likelihood.

4 Results

4.1 Summary Statistics

We begin with a descriptive analysis of the dataset. Table 1 lists the mean, standard deviation, minimum and maximum for the revenue, personnel, input, patent and paper production variables. In 1998 the average revenue across the sample of research institute was 8.6 million RMB. By 2012 it had climbed to 103 million. This change represents an average increase of 23% per year. In contrast, the Chinese economy grew at a rate of 8% per year over this period and inflation was 2%.⁷ These data show robust growth in the research institute sector. Additionally, there is a large amount of heterogeneity in revenue. The standard deviation of revenue reaches 1.6 billion RMB by 2012 and the largest research institutes have revenue in the tens of billions of RMB.

Next we look at the composition of revenue. Revenue can be split into revenue from S&T related activities and P&M (Production and Management) activities. We find that 85% of institute revenue

⁷ Inflation and real GDP growth come from the World Bank World Development Indicator and the Penn World Table 9.0 respectively. Data are obtained via Federal Reserve Economic Data (FRED).

comes from S&T activities and 15% come from P&M activities. These fractions are stable across time. Government grant revenue, a subset of S&T revenue, has increased from 55% of revenue on average in 1998 to 66% of revenue in 2012. This increase is consistent with the government's stated policy goal of concentrating the research institute sector towards basic research funded by the government and away from more commercial activities [see Jiang et al. (2016)]

Table 1 also reports summary statistics for personnel, the number of workers at each research institute. The average research institute employed about 140 workers in 1998 and about 150 workers in 2012. Combined with the large increase in revenue over this time period we see a surge in productivity, measured as revenue per worker. Not surprisingly, there is a large variance in the size of research institutes measured by the number of employees. The standard deviation of employment is 758 workers in 2012 and the largest research institutes employee over 10,000 workers.

While the number of workers has remained roughly constant over time the composition of the workers has changed noticeably. The fraction of employees primarily involved in S&T activities rose from 70% in 1998 to 76% in 2012. While the fraction involved in production and management fell from 19% to 12% over the same period. Similarly, the fraction of personnel that were classified as R&D personnel (an employee working on the most innovative activities) rose dramatically from 17% in 1998 to 41% in 2012.

Spending on S&T intermediate inputs and equipment also grew rapidly. Spending on S&T intermediate inputs like materials, conference travel etc. increased from 2 million RMB in 1998 to 17 million RMB in 2012 and spending on S&T equipment rose from less than one million RMB in 1998 to over 6 million RMB in 2012. These increases represent an average growth rate of about 18% per year. We again see evidence of some very large research institutes. The standard deviation is 4 to 6 times the mean of spending and the maximum spending is over a billion RMB in 2012.

Finally, we turn our attention to the summary statistics on patenting. The average research institute applied for 0.5 patents (and was granted 0.35) in 1998. By 2012 the average research institute applied for 8.4 patents (and was granted 5). We also can see that among patent applications, most were for invention patents, a more novel idea compared to the more process-based innovations

captured by utility patents. The average research institute published 16 papers in 1998. By 2012 that number had risen to 38. The data also demonstrates a large amount of heterogeneity in patenting. The largest research institutes produce thousands of patents per year.

To visualize this heterogeneity, Figure 1 presents histograms for patent applications, patents granted, innovation patent applications and scientific papers. We use the sum of these values for the current year and the previous two years. For clarity we truncate the histograms at 50 for patent applications, patents granted, and scientific papers and at 30 for invention patent applications. We see that about 45% of research institutes failed to apply for a patent over the last three years. Additionally, half of research institutes were not granted a patent in the last three years. Eighty percent did not apply for an invention patent and 28% had no scientific publication in the last three years. However, of the research institutes that did patent and publish there are a substantial number with 10 or more patents or publications.

4.2 Geographic Distribution of Research Institutes

Next we explore the geographic distribution of the research institutes. Recall that we have detailed information on the location of the research institutes. Given the administrative region code which is in our data we have a precise measure of the geographic area where the institute is located. These codes vary depending on population density. In the most densely populated areas, the administrative code would correspond to a district. In less densely populated areas, the code corresponds to an entire city or a county in the least dense areas.

We use this geographic detail to calculate distance between research institutes for our regression. The last rows of table 1 presents overall summary statistics for distance. (We do not calculate it by year since distances do not change by year). The average research institute is 1,221 km from the other institutes in its industry. The most remote research institute is 3,521 km on average from the other research institutes in its industry and the standard deviation of this measure is 395 km. The average research institute is 145 km from the other research institutes in its province. The best-connected research institute is only 5.3 km on average from the other institutes in its province and the most remote is 1,086 km from the others in its province. The standard deviation of this

measure is 145 km. For more detail see appendix table 1 which breaks down distance by each industry and province.

To further analyze the distribution of research institutes, we present a cross-tabulation of research institutes at the aggregate province and sector level. This tabulation is contained in table 2. First, we aggregate our 46 industries into 10 over-arching sectors. These sectors are Agriculture, Mining, Low-tech, Med-Low-tech, Med-High-tech, High-tech, High-tech services, Sanitation, Electric and Gas, and Other, by grouping together similar industries. For example, industry 17, Manufacture of Textile, is classified as Low-tech, while, Manufacture of Electrical Machinery and Equipment is classified as Med-High-tech and Manufacture of Medicines is classified as High-Tech. Appendix A provides a complete list of all the classifications.

Next we assign numbers, for the 31 administrative level provincial units in China. The numbers begin at 11 and end at 65. Appendix B includes a map of China and the complete list of numbers and corresponding administrative level provincial units. Loosely speaking, the lower numbers correspond to the northeast beginning with Beijing, then the southeast beginning with Shanghai. The middle numbers correspond to the central regions of the country, and the highest numbers correspond to the western provinces.

Turning to table 2, the province with the largest number of research institutes is Beijing (#11) which contained 269 research institutes in 1998 and 293 by 2012. Shandong (#37) contained the second largest number of research institutes with 262 in 1998, though in contrast to the number in Beijing, the number of research institutes fell to 202 by 2012. On the other hand, the provinces with the fewest number of research institutes are Tibet (#54) which had 11 in 1998 and 16 in 2012. Qinghai (#63), one of the western provinces, and Ningxia (#64), a relatively small province in central China, also had few research institutes. These provinces had 33 and 44 research institutes respectively in 1998 and 35 research institutes each in 2012.

On the sector side, Agriculture was home to the most research institutes with 1,212 in 1998 and 1,066 in 2012. The second largest industry for the research institutes is High-tech services. High-tech services include industries like telecommunications, software and information technology (IT) services and research and experimental development. In contrast, there are relatively few research institutes in mining and electric and gas, about 40 for each industry.

An interesting question related to our current paper is if we observe clustering of industries or if they are evenly distributed across geographic locations. From table 2 we can observe some substantial clustering. For example, in 2012 while 9% of all the research institutes are located in Beijing, 13% of the high-tech services research institutes are. Similarly, Liaoning (#21) contains 7.5% of the Med-High-tech services like chemical or machinery manufacturing while containing only 4.5% of the research institutes overall. For more detail, see appendix tables 2 to 4 which are structured as table 2 is but with sector disaggregated by industry.

For a more graphical approach to demonstrating geographic clustering, view figure 2. For four sectors with the largest number of research institutes, we plot the percent of the total number of research institutes in that industry by province. All data is from 2012.

Since there are 31 provincial administrative areas, if research institutes were evenly distributed across the provinces, we would have 3.2% of the research institutes in each province. Instead we see a much less even distribution. For example, in Agriculture, Heilongjiang and Guangdong have 6% of the research institutes while Qinghai has less than 1%. For Low-tech industries, Beijing and Shandong have the highest percent of research institutes around 8% while the western provinces have none. We see a similar pattern for the Med-High-tech industries, but now Liaoning has the largest share, around 7.5%. Finally, Beijing is the leader in High-tech services with over 12% of the research institutes. However, Shandong and Guangdong have large shares as well. In contrast, Hainan, Chongqing, and Ningxia have very few research institutes.

4.3 Regression Results

We now turn our attention to the regression analysis. Our main regression equation is:

$$y_{i,t} = \eta_i + \lambda_t + \alpha l_{i,t} + \beta X_{i,t} + \gamma k^0_{i,t} + D_{i,t} * k^0_{i,t} + D_{i,t} + v_i,$$

Here $y_{i,t}$ is a measure of output like total revenue, $l_{i,t}$ is a measure of labor input, $X_{i,t}$ are additional input measures like equipment and intermediate inputs, η_i are sector and province fixed effects or institute fixed effects, λ_t are year fixed effects and $k^0_{i,t}$ is the spillover variable.

For our spillover variables we consider variables external to the system of research institute and variables internal to the system of research institutes. When we use a variable that is internal to

the system of research institutes, we interact it with the research institute's distance from other research institutes. Section 3 discusses how distance is calculated.

4.3.1 Patenting and Scientific Paper Results

Table 3 begins our analysis of the patenting and paper variables. Recall, that for these count variables we use a Poisson estimator. In this case, we are modeling the log of the Poisson parameter λ (the expected number of patents) as a linear function of our explanatory variables.

We consider three labor input measures: total personnel, S&T personnel and the fraction of S&T personnel which are R&D personnel. We have two additional input variables: log S&T equipment and log S&T intermediate inputs. We also include the fraction of S&T equipment which is R&D equipment. For our spillover variables we consider the external variables R&D expenditure within the province per worker and provincial patents per person. For our internal spillovers we consider the knowledge stock in the institute's industry and the knowledge stock in the institute's province and sector. We also include the number of other institutes in the same province and sector as the research institute. We generally enter the spillover variables individually to avoid collinearity. In each table and for each spillover variable, the first column's analysis is without institute fixed effects, the second adds institute fixed effects.

We see that total personnel increases patent applications as does the number of S&T employees. For both variables, a 10% increase in personnel leads to a 2% increase in patents on average based on the fixed effects specification. Having a larger fraction of R&D personnel increases the number of patents but this effect switches sign with institute fixed effects. This is most likely because R&D workers are focused on basic research and not patenting. We find support for this conjecture when we examine the impact on scientific papers where a large R&D share of personnel is associated with more publishing. S&T equipment and inputs both tend to increase patent applications. It also appears that when a larger share of S&T equipment is R&D equipment patenting will increase even more.

Next we examine the effect of our external spillover variables. We find that having more R&D spending per worker in the province spills over into more institute patenting. A 1% increase in

external R&D spending is associated with an expected increase in patenting of 0.33%, controlling for observable inputs and unobservable time invariant factors using institute fixed effects. We also find evidence that provincial patenting spills over to the research institutes. If the number of patents in the province per person rises so does research institute patenting, though the result is only statistically significant in our fixed effects specification.

For the internal spillovers we focus on a knowledge stock measure. This measure is the depreciated sum of past R&D expenditure. If there is a large knowledge stock in one's sector and province, patenting tends to increase. A 10% increase in the provincial-sector knowledge stocks leads to a 1.9% increase in the expected number of patent applications from the research institute. Interestingly, this effect is mitigated by distance. For every 100% increase in distance from the knowledge stock, the effect of the knowledge stock on the research institute's patents falls by 0.13 percentage points. We also find that an increase in the number of research institutes in your province and industry decreases patenting, suggesting the possibility of increasing returns to scale in patenting at the institute level as we would see more productivity if the existing institutes inputs were combined into few institutes. On the other hand, we see that having a large knowledge stock in one's industry does not affect patenting for the research institutes. Perhaps because while a larger knowledge stock may promote spillovers, many inventions have already been patented making it harder to patent more. In table 5, we find that there are more invention patents when industry knowledge stock is higher, consistent with this notion that institutes must be more innovative to patent in industries with much prior R&D expenditure.

It may seem surprising that the coefficient on distance for the provincial sector knowledge stock is positive. However, given the presence of the interaction term in the regression the marginal effect distance is given by the first distance coefficient plus the coefficient on the interaction term times the provincial sector knowledge stock. The (log) mean value of the provincial sector knowledge stock is 14.3. Evaluating the distance effect at this mean value we get a marginal effect of distance equal to -0.098 implying the intuitive result that 10% increase in distance from the knowledge stock is associated with a 1% reduction in patenting.

In addition to patent applications, we also have data on the number of patents granted. We present results for this measure in table 4. As with patent applications we find that increased S&T

personnel, inputs and equipment all increase patenting. There are some differences though. We find, with fixed effects, that total personnel and the fraction of R&D inputs negatively impacts patents granted. We take these results as evidence that expanding research institutes are focusing on basic research which does not necessarily lead to more patenting. We show in table 7 this increased spending does lead to more scientific publications though.

Again, for our spillover variables, more R&D per worker in the province and more patents per person in the province lead to more patenting. Turning to our internal spillover variables we find that the knowledge stock within one's province and sector promotes patenting and that this effect diminished with distance. We also find that fewer patents are granted when the research institute's national industry knowledge stock is large.

There are two main types of patents: invention patents and utility patents. Invention patents are more innovative compared to utility patents which are like adaptations of current innovations. For invention patents we see that increased personnel (fixed effects), S&T personnel, equipment and inputs all tend to promote invention patenting. Additionally, the use of R&D equipment and inputs help increase patenting while a larger share of R&D personnel is associated with less invention patents.

In terms of spillovers we find that provincial R&D spending and patenting spills over into increased invention patenting in the research institutes while the number of similar research institutes in your province is unimportant. For our internal spillovers we find that increases in the industry knowledge stock promotes patenting. A 10% increase in the industry's knowledge stock would lead to 4% increase in invention patents per research institute. Similarly, a 10% increase in the research institutes knowledge stock in the province and sector would lead to a 2% increase in patents per research institute. This effect falls with distance. A 100% increase in distance from the knowledge stock, reduces this effect by 0.13 percentage points (for a 10% increase in the knowledge stock).

For utility patents we find that having more personnel, S&T personnel, equipment and inputs all increases the number of utility patent applications. Moreover, having the inputs and equipment be a larger fraction of R&D raises utility patenting. However, institutes with a large fraction of R&D

personnel tended to produce fewer utility patent applications because, as we show in table 7, they are focusing on scientific papers.

We find positive spillovers from provincial R&D spending and provincial patenting and negative spillovers from the number of other research institutes in your province and industry. We find that the national knowledge stock reduces utility patenting, but that the provincial sector knowledge stock increases patenting. In the latter case a 10% increase in the provincial sector knowledge stock should lead to 1.6% more utility patent applications per research institute. Again, being farther from the knowledge stock will reduce this effect. A 100% increase in distance from the knowledge stock reduces patenting by 0.1 percentage points.

Finally, we turn our attention to scientific paper production. In general, all the inputs increase scientific paper production: total personnel, S&T personnel, more R&D personnel, S&T equipment, R&D equipment, S&T inputs and R&D inputs. We also see positive spillovers in the sense that more R&D spending in the province and more patenting in the province is associated with a significant increase in scientific paper production. On the other hand, an increase in the number of research institutes in your province and industry leads to a fall in paper production.

Looking for spillovers within the research institute system, we find that a larger national industry knowledge stock leads to more paper production. But the result is not significant with institute fixed effects. Similarly, a large provincial-sector knowledge stock leads to more scientific paper production. For example, a 10% increase in the provincial sector knowledge stock results in a 0.14 percent increase in the number of papers per institute. And this effect falls by 0.02 percentage points for every 100% increase in distance for the research institutes.

4.3.2 Revenue Results

Table 8 contains regression results with log total revenue as the dependent variable. We consider three labor input measures: total personnel, fraction of total personnel which are S&T personnel and fraction P&M personnel. We have two additional input variables: log S&T equipment and log S&T intermediate inputs. For our spillover variables we consider the external variables R&D

expenditure within the province and total industry sales within the province as a fraction of provincial GDP. For our internal spillovers we consider total revenue within the research institute's industry and total revenue for all other research institutes in the same province and sector as the institute. As with patents, we enter the spillover variables individually. The first column for each spillover variable is without institute fixed effects, the second adds institute fixed effects.

We find that our input measures generally have the correct sign and are of a reasonable magnitude. For example, a 1% increase in total personnel raises total revenue by 0.75%. This number falls to 0.55% with institute fixed effect, presumably because total personnel is correlated with differences in institute productivity. We find that without institute fixed effects the fraction of P&M personnel positively predicts the institute's total revenue. In this case, a 10-percentage point increase in the fraction of P&M personnel would lead to a 1.5% increase in total revenue. However, once we allow for institute fixed effects the P&M personnel coefficient is no longer significant. With institute fixed effects, the S&T personnel coefficient becomes marginally significant. In this case a 10-percentage point increase in the fraction of S&T personnel raises revenue by 1%. One possible explanation for this difference across specifications is that institutes with large amounts of revenue employ more P&M staff to manage that revenue but increases in revenue are coming from additional innovation accomplished by S&T employees.

Additionally, increases in inputs like S&T equipment and S&T intermediate inputs lead to an increase in total revenue. A 10% increase in S&T equipment leads to a 1% increase in total revenue and a 10% increase in S&T inputs leads to a 2.5% increase in total revenue. These estimates are smaller with fixed effects 0.6% and 1% respectively.

Now turning our attention to the external spillover variables we find no evidence that provincial R&D expenditure increases institute productivity. The coefficient is small and insignificant in both specifications. This result contrasts with the positive effect this variable had on patenting. Apparently, the patents do not immediately generate revenue for the research institutes. On the other hand, there is some evidence that institute revenue rises when sales in the province rise relative to provincial GDP. In our fixed effects specification, a 10-percentage point increase in

the provincial GDP share of the industry that corresponds to the research institute's industry raises total revenue by about 1%.

We now examine our internal spillover variables. We examine the effect of increases in total revenue within the research institute's sector and province. A 10% increase in this revenue leads to 1% increase in total revenue. Notably, this effect falls by 0.08 percentage points given a 100% increase in the research institute's average distance. We find a similar impact of total revenue of all the other research institutes in the institute's industry. We find that an increase in national industry revenue is associated with an increase in the research institute's total revenue. A 10% increase in national industry research institute revenue translates to a 1.3% increase in the institute's total revenue. Most important, this effect diminishes as the research institute's average distance from the other institutes in its industry increases. For a 100% increase in distance from the other research institutes in its industry, this effect falls by 0.2 percentage points. However, this latter effect is not robust to including institute fixed effects.

Now we consider the various components of revenue: S&T revenue and its subset government grant revenue and (P&M) production and management revenue. We first find that the effect of increasing personnel on S&T revenue is like the effect on total revenue. We find that a larger fraction of P&M personnel is associated with lower S&T revenue. This effect suggests that those institutes with a large fraction of P&M revenue focus more on business activities and less on basic science and technology research. We find, with fixed effects, that the fraction of S&T personnel has no effect on S&T revenue. This is a bit paradoxical but recall that S&T personnel are often engaged in basic research that is funded by the government and leads to fewer commercially viable products. S&T equipment and S&T intermediate inputs both increase S&T revenue.

For our external spillover variables, when provincial industry sales go up so does research institute S&T revenue. This result is consistent with the idea that stronger economic activity in the institute's industry increases its revenue from research related to the industry. It is also consistent with the idea that the government channels more funding to research institutes that are in booming industries. Surprisingly, we find that increased R&D expenditure in the province predicts lower S&T revenue. Possibly, R&D intensive provinces focus more on basic research and less on generating commercial revenue.

With our internal spillover variables, we see that when revenue for research institutes in the same province and sector as the research institute increases, it tends to generate more S&T revenue. This effect falls as the institute is further from other related research institutes. For example, a 1% increase in the revenue of related research institute leads to a 0.04% increase in S&T revenue, however this effect falls by 0.01 percentage points for a 100% increase in the research institute distance from other research institutes. When we add fixed effects the signs of the coefficients are consistent but the distance effect is no longer significant. We find inconsistent effects of increased national industry revenue on S&T revenue. National industry revenue is associated with more S&T revenue and this effect diminishes the farther the research institute is located from its related institutes. However, the sign of this effect changes once institute fixed effects are added.

Next we examine the impact on P&M revenue in table 10. Not surprisingly, more workers and more P&M personnel lead to higher P&M revenue. On the other hand, a larger fraction of S&T personnel lowers P&M revenue though the effect is not significant with institute fixed effects. More S&T equipment and more S&T intermediate inputs increase P&M revenue but the later effect is not significant with the institute fixed effects.

For the spillover variables, we find no evidence that our external spillover variables impact P&M revenue. However, we find that when national industry revenue rises by 1%, P&M revenue increases by 1.5% (fixed effects). However, for a 100% increase in distance away from national industry revenue, this effect falls by 0.18 percentage points. Similarly, a 1% increase in revenue among the research institutes in your sector and province, increases P&M revenue by 0.18% and this effect falls by 0.03 percentage points for a 100% increase in the research institute distance from other related research institutes.

Table 11 considers the results for Government Grant revenue. Similar to the effect on S&T revenue, more personnel, more S&T equipment, and more S&T inputs raise government grant revenue. On the other hand, an increased fraction of P&M personnel lowers government grant revenue.

We find no evidence that provincial R&D expenditures spill over into increased government grants, however a 10% increase in the research institutes industry's share of provincial GDP is

associated with 1-2% increase in government grant revenue. This result suggests the government may target basic research related to industries which are seen as important for a given geographical area. For the internal spillovers, we find that a 10% increase in national industry revenue leads to a 6% increase in government grant revenue. However, this effect falls by 0.6 percentage points for a 100% increase in the distance of the research institutes from other research institutes in its industry. Similarly, a 10% increase in revenue in the research institutes province and sector leads to a 0.2% increase in government grant revenue which falls by half when distance is doubled. However, the distance effect is not robust to including institute fixed effects.

4.4 Interpretation and Implications for Policy

In our analysis we have controlled for a substantial number of institute inputs, all time-invariant unobservable variables using institute fixed effects, and relied on spillover measures which are determined outside of the research institute's direct control. However, since we do not have pure exogenous variation in our spillover measures, we are unable to rule out the possibility that some time varying factor, correlated with our spillover variables are causing the positive productivity effects we document. For example, provincial R&D expenditure may go up at times when particularly promising technologies arrive, and the technology arrival not necessarily the external spending may cause the improvement in institute patenting and productivity. Similarly, distance from other institutes may be correlated with distance from economic activity and therefore we may overstate the role distance plays in reducing spillovers.

However, there are additional aspects of our empirical design which mitigate these concerns. First, many of our spillover variables are based on past decisions. For example, the knowledge stock is based on past R&D expenditure and the location of institutes is determined by government fiat often decades in the past. Additionally, even if some unobservable factor is correlated with our spillover variables it is still an important finding that the interaction of the spillover with distance variables is negative. Even though we may not exactly identify the factor that increases productivity, we have still demonstrated that distance attenuates its impact.

Our results have important implications for research policy in China. One stated Chinese policy goal is to increase development away from the coast especially in the western provinces. In light of our findings, this goal may prove challenging as research institutes located farther away from similar research institutes were less productive. However, the results also suggest ways in which China can mitigate this difficulty. Specifically, creating linkages between similar research institutes will promote knowledge transmission. Additionally, promoting clusters of similar research institutes will allow them to reap the benefits of increased spillovers.

The results also point to inefficiencies in the research institute sector that can be resolved with better policy. For example, the number of research institutes in a province/sector was negatively correlated with productivity. This result suggests that merging institutes will lead to productivity gains. Additionally, despite the positive productivity effects of geographical proximity, government grant revenue did not fall with increased distance from similar research institutes. Therefore, a channel exists to allocate government funding more productively. Finally, as productivity increased as the research institute's industry became a more important part of regional GDP it suggests that the government should promote research institutes that are a good match for the local economy.

5. Conclusion

Research and development and the resulting knowledge and technology is seen as a key driver of economic growth and living standards. This is especially true in China, where the economy is growing rapidly, research and development levels are high, and R&D is promoted by the central government as a way to maintain the country's impressive economic growth.

The study of knowledge production though is beset by a set of thorny issues. Not the least of which are the inability to accurately measure research output, the inability to connect research institutes in the same area, and the ability to observe a large number of actors engaged in research and development.

We address these issues with a novel, administrative data set provided by the Chinese government's Ministry of Science and Technology. We observe the location of China's roughly 4,000 research institutes and detailed information on their inputs, outputs and industry of specialization.

We found that increasing inputs led to higher knowledge production measured both by institute revenue and patent and paper production. Additionally, activity outside the research institutes contributed to their success. When provincial R&D spending was high, research institutes patented more. Similarly, when the province had a high stock of patents per person, institutes were more likely to patent. The number of institutes in a similar area to the research institute lowered patenting and publication. The knowledge stock within the institute's province and sector positively contributed to innovative output. Revenue output of the research institutes increased when the institute's industry became a larger fraction of provincial GDP. The revenue output of institutes also responded to the revenue of other research institutes. Institutes had higher revenue when other institutes in their industry and province had higher revenue.

Importantly, in the most novel feature of our paper, we found that spillovers within the system of research institutes were diminished when research institutes were located farther from related research institutes. For example, the effect of the knowledge stock on productivity fell when distance from the knowledge stock increased. Similarly, the relationship between institute revenue and the revenue of other research institutes fell when distance from the research institutes increased.

Going forward, these results have important implications for research policy in China. Specifically, as China tries to increase development in the central and western provinces it should pay attention to creating linkages between similar research institutes as a way to promote knowledge transmission. Additionally, the spillover results reveal significant avenue for the government to improve the productivity of the research institute sector.

References

- Anselin, Luc, Attila Varga & Zoltan Acs, 1997. "Local Geographic Spillovers between University Research and High Technology Innovations," *Journal of Urban Economics*, Elsevier, vol. 42(3), pages 422-448, November.
- Audretsch, David and Maryann P Feldman, (1996), "R&D Spillovers and the Geography of Innovation and Production", *American Economic Review*, 86, issue 3, p. 630-40.
- Bernstein, Jeffrey I. & M. Ishaq Nadiri, 1989. "Research and Development and Intra-industry Spillovers: An Empirical Application of Dynamic Duality," *Review of Economic Studies*, Oxford University Press, vol. 56(2), pages 249-267.
- Bloom, Nick, Brian Lucking & John Van Reenen, 2018. "Have R&D Spillovers Changed?" CEP Discussion Papers dp1548, Centre for Economic Performance, LSE.
- Branstetter, Lee, (2006), "Is foreign direct investment a channel of knowledge spillovers? Evidence from Japan's FDI in the United States", *Journal of International Economics*, 68, issue 2, p. 325-344.
- Buzard, Kristy & Gerald Carlino, 2013. "The geography of research and development activity in the US," Chapters, in: *Handbook of Industry Studies and Economic Geography*, chapter 16, pages 389-410 Edward Elgar Publishing.
- Cao, Cong, Ning Li, Xia Li, & Li Liu, 2013, "Reforming China's S&T system." *Science*. 341. 460-462.
- Carlino, Gerald & William R. Kerr, 2015. "Agglomeration and Innovation," *Handbook of Regional and Urban Economics*, Elsevier.
- Chen, Cheung, Kui-yin & Lin, Ping, (2004), "Spillover effects of FDI on innovation in China: Evidence from the provincial data", *China Economic Review*, 15, issue 1, p. 25-44.
- Chen, Kaihua , Yi Zhang, Guilong Zhu & Rongping Mu, "Do research institutes benefit from their network positions in research collaboration networks with industries or/and universities?" *Technovation*, 2017.
- China Statistics Press, 1998-2012, *China Statistical Yearbook on Science and Technology*, Beijing, China.
- China Statistics Press, 1998-2012, *China Statistical Yearbook*, Beijing China.
- Conroy, Richard, 1982 "China's Local Scientific Research Sector: Its Role, Impact and Future Prospects, *The Australian Journal of Chinese Affairs*, Vol. 7, pp.85-118

- Ellison, Glenn & Edward L. Glaeser & William R. Kerr, 2010. "What Causes Industry Agglomeration? Evidence from Coagglomeration Patterns," *American Economic Review*, American Economic Association, vol. 100(3), pages 1195-1213, June.
- Griliches, Zvi, 1992 " The Search for R&D Spillovers," *Scandinavian Journal of Economics*, Wiley Blackwell, vol. 94(0), pages 29-47.
- Hall, Bronwyn, Jacques Mairesse, & Pierre Mohnen, 2010, "Measuring the Returns to R&D", Chapter 24, p. 1033-1082, Elsevier, in *Handbook of the Economics of Innovation*, Bronwyn H. Hall and Nathan Rosenberg editors.
- Hu, Albert G. Z., Gary H. Jefferson & Qian Jinchang, 2005."R&D and Technology Transfer: Firm-Level Evidence from Chinese Industry," *The Review of Economics and Statistics*, MIT Press, vol. 87(4), pages 780-786, November.
- Hu, Mei-Chih & John Mathews, 2008, "China's National Innovative Capacity," *Research Policy*. 37. 1465-1479.
- Jaffe, Adam B, 1986. "Technological Opportunity and Spillovers of R&D: Evidence from Firms' Patents, Profits, and Market Value," *American Economic Review*, American Economic Association, vol. 76(5), pages 984-1001, December.
- Jaffe, Adam B, 1989. "Real Effects of Academic Research," *American Economic Review*, American Economic Association, vol. 79(5), pages 957-970, December.
- Jaffe, Adam B, Manuel Trajtenberg & Rebecca Henderson, 1993. "Geographic Localization of Knowledge Spillovers as Evidenced by Patent Citations," *The Quarterly Journal of Economics*, Oxford University Press, vol. 108(3), pages 577-598.
- Jiang, Renai, Daniel L. Tortorice & Gary H. Jefferson, 2016. "Restructuring China's research institutes," *The Economics of Transition*, The European Bank for Reconstruction and Development, vol. 24(1), pages 163-208, January.
- Kantor, Shawyn & Alexander Whalley, 2014. "Knowledge Spillovers from Research Universities: Evidence from Endowment Value Shocks," *The Review of Economics and Statistics*, MIT Press, vol. 96(1), pages 171-188, March.
- Kuo, Chun-Chien & Yang, Chih-Hai, 2008. "Knowledge capital and spillover on regional economic growth: Evidence from China," *China Economic Review*, Elsevier, vol. 19(4), pages 594-604, December.
- Liu, Feng-chao & Simon, Denis Fred & Sun, Yu-tao & Cao, Cong, 2011. "China's innovation policies: Evolution, institutional structure, and trajectory," *Research Policy*, Elsevier, vol. 40(7), pages 917-931, September.
- Motohashi, Kazuyuki & Xiao Yun, 2007. "China's innovation system reform and growing industry and science linkages," *Research Policy*, Elsevier, vol. 36(8), pages 1251-1260, October.

Neves, Pedro Cunha, and Tiago Neves Sequeira, 2018, "Spillovers in the production of knowledge: A meta-regression analysis." *Research Policy*, 47.4, p.750-767.

Qiu, S., Liu, X., & Gao, T. (2017). Do emerging countries prefer local knowledge or distant knowledge? Spillover effect of university collaborations on local firms. *Research Policy*, 46(7), 1299-1311.

Rosenthal, Stuart & William Strange, 2001, The Determinants of Agglomeration, *Journal of Urban Economics*, 50, issue 2, p. 191-229.

Sinnott, R.W. "Virtues of the Haversine", *Sky and Telescope*, vol. 68, no. 2, 1984, p. 159

Thompson, Peter & Melanie Fox-Kean, 2005. "Patent Citations and the Geography of Knowledge Spillovers: A Reassessment," *American Economic Review*, American Economic Association, vol. 95(1), pages 450-460, March.

Thompson, Peter, 2006. "Patent Citations and the Geography of Knowledge Spillovers: Evidence from Inventor- and Examiner-added Citations," *The Review of Economics and Statistics*, MIT Press, vol. 88(2), pages 383-388, May.

Ugur, Mehmet, Sefa Awaworyi Churchill and Hoang M. Luong, 2020. "What do we know about R&D spillovers and productivity? Meta-analysis evidence on heterogeneity and statistical power." *Research Policy*, 49.1.

Wieser, Robert, 2005, "Research and Development Productivity And Spillovers: Empirical Evidence At The Firm Level," *Journal of Economic Surveys*, Wiley Blackwell, vol. 19(4), pages 587-621, September.

Xie, Qingnan & Richard B. Freeman, 2018, "Bigger Than You Thought: China's Contribution to Scientific Publications," NBER Working Papers 24829, National Bureau of Economic Research, Inc.

Xiwei, Zhong & Yang Xiangdong, 2007, "Science and technology policy reform and its impact on China's national innovation system, *Technology in Society*, Elsevier, vol. 29, pp 317-325.

Zhang, Lin, 2017. "The knowledge spillover effects of FDI on the productivity and efficiency of research activities in China," *China Economic Review*, Elsevier, vol. 42(C), pages 1-14.

Figure 1: Histograms of Patent and Paper Output

Patent Applications

Invention Patent Applications

Patents Granted

Scientific Papers

Note: This figure presents histograms for patent applications, invention patent applications, patents granted and scientific papers. These variables are calculated by summing across the current year and previous two years. Histograms are truncated at 50 in all cases except for invention patent applications where the histogram is truncated at 30.

Figure 2: Percent of Research Institutes in Each Province by Industry

Note: This figure plots the percent of research institutes in each province by industry

Table 1: Summary Statistics

		<u>1998</u>	<u>2003</u>	<u>2008</u>	<u>2012</u>			<u>1998</u>	<u>2003</u>	<u>2008</u>	<u>2012</u>
Total Revenue	mean	8,566	16,591	51,508	103,958	S&T Inputs	mean	2,211	2,120	8,896	16,811
	sd	30,447	62,274	462,032	1,616,176		sd	12,211	9,468	36,108	61,024
	min	0	0	0	7		min	0	0	0	0
	max	899,159	2,339,754	24,300,000	90,300,000		max	547,061	232,572	1,196,400	1,388,374
Fraction S&T Rev.	mean	0.869	0.864	0.851	0.842	S&T Equipment	mean	858	1,489	3,776	6,446
	sd	0.265	0.271	0.293	0.296		sd	3,891	6,596	26,320	38,434
	min	0	0	0	0		min	0	0	0	0
	max	1	1	1	1		max	127,000	153,421	974,862	1,300,048
Fraction P&M Rev	mean	0.15	0.141	0.149	0.157	Patent Applications	mean	0.46	1.38	3.16	8.39
	sd	0.284	0.277	0.293	0.296		sd	3.02	9.66	15.69	51.08
	min	0	0	0	0		min	0	0	0	0
	max	1	1	1	1		max	103	298	311	1852
Fraction Gov Grant Rev.	mean	0.545	0.607	0.642	0.662	Patents Granted	mean	0.25	0.74	1.62	5.03
	sd	0.356	0.367	0.373	0.366		sd	1.49	6.14	9.08	28.76
	min	0	0	0	0		min	0	0	0	0
	max	1	1	1	1		max	33	205	233	1217
Total Personnel	mean	138.7	108.7	128.5	153.3	Invention Patents	mean		0.97	2.31	6.01
	sd	261.9	202.3	474.5	758		sd		8.07	13.41	35.85
	min	1	0	1	0		min		0	0	0
	max	4737	5343	19430	33812		max		248	305	834
Fraction S&T Personnel	mean	0.695	0.7103	0.74	0.7642	Utility Patents	mean		0.41	0.84	2.38
	sd	0.2197	0.2299	0.237	0.2388		sd		2.64	4.43	22.19
	min	0	0	0	0		min		0	0	0
	max	1	1	1	1		max		91	97	1153
Fraction P&M Personnel	mean	0.185	0.166	0.138	0.119	Scientific Papers	mean	15.9	20.93	30.49	38.16
	sd	0.197	0.205	0.204	0.201		sd	38.78	57.33	80.9	99.54
	min	0	0	0	0		min	0	0	0	0
	max	1	1	1	1		max	574	860	1184	17.49
Fraction R&D Personnel	mean	0.167	0.211	0.297	0.413	Distance: Ind.	mean	1221	Prov	mean	145
	sd	0.259	0.312	0.456	1.466		sd	395		sd	126
	min	0	0	0	0		min	0		min	5.3
	max	1	1	1	1		max	3521		max	1086

Note: This table gives summary statistics for the key variables in the analysis. Monetary values are in 1,000 RMB. Other measures are single units. Distances are km

Table 2: Distribution of Research Institutes Across Sectors and Provinces

1998	sector/prov.	11	12	13	14	15	21	22	23	31	32	33	34	35	36	37	41	42	43	44	45	46	50	51	52	53	54	61	62	63	64	65	total
	Agriculture	25	11	38	41	41	51	30	66	12	47	31	42	33	30	58	50	67	58	87	44	19	17	52	42	57	6	39	36	6	21	55	1212
	Mining	4	1	3	3	1	2	1			4	2	1	1	1	4	1	3			1		1	1	1	1		3	2		1	1	44
	Low-tech	23	15	10	11	8	29	4	19	18	24	11	11	7	6	36	12	12	10	20	6	3	7	7	6	8		8	7	2		4	344
	Med-Low-tech	18	12	5	4	5	19	5	5	16	15	5	6	4	7	14	5	11	5	9	2		4	7	3	4		4	3		2	3	202
	Med-High-tech	33	27	16	14	12	51	16	25	43	45	21	18	9	6	37	17	27	17	28	16	1	10	20	6	9		18	10		1	2	555
	High-tech	31	11	11	7	1	20	10	12	21	35	11	18	4	4	21	4	17	11	18	7	1	10	9	2	4		11	8		1	3	323
	High-tech Services	101	27	24	28	18	56	46	39	40	37	44	50	21	29	67	36	47	43	56	21	11	8	49	24	24	4	30	27	22	16	24	1069
	Sanitation	29	3	7	22	8	6	5	12	12	17	2	9	13	4	25	10	4	3	9	11	4	2	11	3	5		7	3	2	2	3	253
	Electric & Gas	5	2		1	4	3	1		1	7	2		2	1		1	1	1	1					1		1	1	2	1		1	40
	Other				1		1				1	1	2	1					1	1	1			1		1		1					13
	Total	269	109	114	132	98	238	118	178	163	232	130	157	95	88	262	136	189	149	229	109	39	59	157	88	113	11	122	98	33	44	96	4055
2005	sector/prov.	11	12	13	14	15	21	22	23	31	32	33	34	35	36	37	41	42	43	44	45	46	50	51	52	53	54	61	62	63	64	65	total
	Agriculture	23	12	33	39	41	46	30	67	13	43	26	37	35	30	51	45	59	58	84	41	20	11	46	37	43	5	22	35	6	23	51	1112
	Mining	6	1	2	2	1	2	1			4	1	1	1	1	5	1				1		1		1	1		3	1		1	1	39
	Low-tech	16	13	5	10	4	19	3	16	17	18	9	3	5	6	23	11	10	9	13	6	2	4	6	4	7		6	5			5	255
	Med-Low-tech	16	11	2	4	4	13	2	4	11	10	4	3	3	7	14	6	8	4	9	2		1	7	3	3		3	2		1	3	160
	Med-High-tech	24	25	11	13	11	38	12	20	35	34	20	13	4	6	32	16	17	13	22	9	1	7	17	3	9		15	8		1		436
	High-tech	22	5	5	6	2	14	5	11	15	21	8	8	4	4	20	3	9	6	16	5	1	6	9	1	4		7	4		2	3	226
	High-tech Services	133	32	26	26	23	45	47	45	56	38	44	40	18	31	67	33	46	36	58	18	6	10	49	21	33	8	34	23	22	12	26	1106
	Sanitation	29	3	5	16	9	4	4	11	10	14	2	7	11	3	22	9	4	2	7	11	1	2	9	3	5		6	3	2	2	3	219
	Electric & Gas	6	2		1	3	2	1	1	1	6	3		1	1		1		2	1					1		2	1	2	1		1	40
	Other				1																			1									2
	Total	275	104	89	118	98	183	105	175	158	188	117	112	82	89	234	125	153	130	210	93	31	42	144	74	105	15	97	83	31	42	93	3595
2012	sector/prov.	11	12	13	14	15	21	22	23	31	32	33	34	35	36	37	41	42	43	44	45	46	50	51	52	53	54	61	62	63	64	65	total
	Agriculture	29	14	32	39	33	44	25	68	12	36	26	30	35	31	49	38	56	58	78	43	19	8	42	36	40	8	19	42	6	19	51	1066
	Mining	9	1	2	3	1	2	1			3	1	1	1	1	4	1				1		1		1	1		3	1		1	1	41
	Low-tech	16	10	3	8	4	12	3	10	13	10	9	2	5	6	17	6	9	7	9	7	2	4	6	4	7		4	5			5	203
	Med-Low-tech	15	9	1	3	4	7	2	4	5	7	4	1	3	7	9	6	5	4	8	1		1	5	4	3		3	2		1	3	127
	Med-High-tech	23	23	5	11	10	28	10	18	25	28	19	11	4	5	25	14	15	12	18	7	1	6	16	3	6		11	8		1		363
	High-tech	20	7	5	4	2	11	4	9	8	20	8	8	4	4	17	3	7	5	14	4	1	6	7	1	4		8	5		1	3	200
	High-tech Services	144	33	27	27	22	41	41	45	56	39	43	33	19	32	64	31	38	37	58	20	6	7	48	22	22	6	31	24	26	10	28	1080
	Sanitation	33	3	5	15	10	4	4	9	10	14	2	6	9	3	17	9	3	2	7	11		2	7	3	7		4	4	2	2	3	210
	Electric & Gas	4	2		1	3	2	1	2		7	3		1	1		1		2	1					1		2	1	2	1		1	39
	Other				1						1													1									3
	Total	293	102	80	112	89	151	91	165	129	165	115	92	81	90	202	109	133	127	193	94	29	35	132	75	90	16	84	93	35	35	95	3332

Table 3: Patent Applications

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
	Pat App	Pat App	Pat App	Pat App	Pat App	Pat App	Pat App	Pat App
Total Personnel	0.113 (0.070)	0.213*** (0.026)	0.115 (0.070)	0.227*** (0.026)	0.055 (0.069)	0.265*** (0.024)	0.040 (0.072)	0.270*** (0.025)
S&T Personnel	0.452*** (0.083)	0.193*** (0.027)	0.450*** (0.084)	0.181*** (0.027)	0.486*** (0.077)	0.156*** (0.025)	0.508*** (0.080)	0.190*** (0.025)
Fraction R&D Personnel	0.187*** (0.015)	-0.426*** (0.017)	0.187*** (0.015)	-0.465*** (0.017)	0.182*** (0.015)	-0.478*** (0.016)	0.186*** (0.015)	-0.499*** (0.016)
S&T Equipment	0.356*** (0.018)	0.096*** (0.005)	0.355*** (0.018)	0.097*** (0.005)	0.358*** (0.018)	0.077*** (0.005)	0.351*** (0.018)	0.085*** (0.005)
Fraction R&D Equipment	1.141*** (0.061)	0.334*** (0.019)	1.139*** (0.062)	0.351*** (0.019)	1.034*** (0.059)	0.330*** (0.018)	1.077*** (0.059)	0.276*** (0.018)
S&T Intermediate Inputs	0.189*** (0.024)	0.101*** (0.007)	0.190*** (0.024)	0.108*** (0.007)	0.201*** (0.023)	0.109*** (0.006)	0.212*** (0.023)	0.100*** (0.006)
Fraction R&D Inputs	0.027** (0.014)	0.023*** (0.006)	0.026* (0.014)	0.024*** (0.006)	0.030** (0.012)	0.025*** (0.005)	0.031** (0.012)	0.019*** (0.005)
Prov. R&D per worker	0.433*** (0.145)	0.334*** (0.025)						
Num. of Inst. in same Prov. and Ind.	-0.056** (0.027)	-0.050** (0.023)	-0.056** (0.027)	-0.031 (0.023)				
Prov. Patents Per Person			0.074 (0.079)	0.047*** (0.013)				
Natl Industry Knowledge Stock (KS)					-0.135 (0.258)	0.057 (0.089)		
Distance from Natl KS					-0.345 (0.556)	-0.421** (0.185)		
Distance*Natl Ind. KS					0.030 (0.038)	0.022* (0.012)		
Prov. Sector KS							0.012 (0.027)	0.185*** (0.011)
Distance from Prov. Sector KS							0.053 (0.047)	0.088*** (0.016)
Distance*Prov. Sector KS							-0.005 (0.005)	-0.013*** (0.002)
Constant	-14.177*** (2.382)		-7.201*** (0.214)		-5.808 (3.788)		-7.421*** (0.405)	
Observations	32,526	17,157	32,526	17,157	34,756	18,620	34,296	18,245
Year Fixed Effect	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Area and Sector Fixed Effects	Yes	No	Yes	No	Yes	No	Yes	No
Institute Fixed Effects	No	Yes	No	Yes	No	Yes	No	Yes

Note: All variables in logarithms. All input variables are three year backward moving averages lagged one year. Robust standard errors in parentheses. *** p<0.01, ** p<0.05, * p<0.1.

Table 4: Patents Granted

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
	Pat. Grant	Pat. Grant	Pat. Grant	Pat. Grant	Pat. Grant	Pat. Grant	Pat. Grant	Pat. Grant
Total Personnel	0.221*** (0.075)	-0.088*** (0.032)	0.222*** (0.076)	-0.063** (0.032)	0.202*** (0.068)	0.122*** (0.031)	0.174** (0.072)	0.075** (0.031)
S&T Personnel	0.330*** (0.086)	0.383*** (0.034)	0.329*** (0.087)	0.371*** (0.033)	0.348*** (0.078)	0.300*** (0.032)	0.379*** (0.082)	0.346*** (0.032)
Fraction R&D Personnel	0.183*** (0.018)	-0.308*** (0.024)	0.184*** (0.018)	-0.353*** (0.023)	0.180*** (0.017)	-0.296*** (0.022)	0.183*** (0.017)	-0.308*** (0.022)
S&T Equipment	0.316*** (0.019)	0.124*** (0.007)	0.316*** (0.019)	0.124*** (0.007)	0.304*** (0.018)	0.067*** (0.006)	0.300*** (0.019)	0.084*** (0.007)
Fraction R&D Equipment	1.001*** (0.057)	0.285*** (0.025)	1.000*** (0.058)	0.319*** (0.025)	0.925*** (0.055)	0.316*** (0.024)	0.960*** (0.055)	0.280*** (0.024)
S&T Intermediate Inputs	0.211*** (0.023)	0.035*** (0.009)	0.212*** (0.023)	0.044*** (0.009)	0.224*** (0.022)	0.042*** (0.008)	0.235*** (0.023)	0.055*** (0.008)
Fraction R&D Inputs	0.023* (0.013)	-0.048*** (0.011)	0.022* (0.013)	-0.043*** (0.011)	0.023* (0.012)	-0.025*** (0.009)	0.026** (0.012)	-0.022** (0.009)
Prov. R&D per worker	0.358** (0.146)	0.551*** (0.033)						
Num. of Inst. in same Prov. and Ind.	-0.049* (0.028)	0.164*** (0.032)	-0.048* (0.028)	0.201*** (0.032)				
Prov. Patents Per Person			0.108 (0.080)	0.151*** (0.018)				
Natl Ind. Knowledge Stock (KS)					-0.343 (0.221)	-0.309** (0.122)		
Distance from Natl KS					-0.921* (0.492)	-1.905*** (0.254)		
Distance*Natl Ind. KS					0.060* (0.032)	0.117*** (0.017)		
Prov. Sector KS							0.063** (0.026)	0.186*** (0.016)
Distance from Prov. Sector KS							0.157*** (0.058)	0.128*** (0.026)
Distance*Prov. Sector KS							-0.015*** (0.005)	-0.015*** (0.002)
Constant	-13.464*** (2.414)		-7.775*** (0.219)		-2.343 (3.365)		-8.601*** (0.412)	
Observations	32,526	15,653	32,526	15,653	34,756	17,064	34,296	16,721
Year Fixed Effect	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Area and Sector Fixed Effects	Yes	No	Yes	No	Yes	No	Yes	No
Institute Fixed Effects	No	Yes	No	Yes	No	Yes	No	Yes

Note: All variables in logarithms. Robust standard errors in parentheses. All input variables are three year backward moving averages lagged one year. *** p<0.01, ** p<0.05, * p<0.1.

Table 5: Invention Patents

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
	Inv. Pat.	Inv. Pat.	Inv. Pat.	Inv. Pat.	Inv. Pat.	Inv. Pat.	Inv. Pat.	Inv. Pat.
Total Personnel	-0.108 (0.087)	0.123*** (0.033)	-0.106 (0.088)	0.132*** (0.033)	-0.171** (0.084)	0.198*** (0.032)	-0.181** (0.087)	0.186*** (0.031)
S&T Personnel	0.625*** (0.105)	0.254*** (0.034)	0.621*** (0.106)	0.243*** (0.034)	0.654*** (0.097)	0.184*** (0.032)	0.670*** (0.099)	0.229*** (0.032)
Fraction R&D Personnel	0.203*** (0.017)	-0.448*** (0.020)	0.205*** (0.017)	-0.475*** (0.020)	0.198*** (0.016)	-0.514*** (0.019)	0.202*** (0.016)	-0.526*** (0.019)
S&T Equipment	0.430*** (0.022)	0.084*** (0.007)	0.429*** (0.023)	0.086*** (0.007)	0.439*** (0.022)	0.075*** (0.006)	0.436*** (0.022)	0.084*** (0.006)
Fraction R&D Equipment	1.398*** (0.071)	0.378*** (0.024)	1.395*** (0.070)	0.387*** (0.024)	1.296*** (0.072)	0.394*** (0.023)	1.364*** (0.068)	0.344*** (0.023)
S&T Intermediate Inputs	0.219*** (0.030)	0.152*** (0.009)	0.220*** (0.030)	0.159*** (0.009)	0.223*** (0.029)	0.149*** (0.008)	0.232*** (0.028)	0.138*** (0.009)
Fraction R&D Inputs	0.021 (0.023)	0.042*** (0.007)	0.016 (0.027)	0.042*** (0.007)	0.021 (0.021)	0.037*** (0.006)	0.019 (0.024)	0.029*** (0.006)
Prov. R&D per worker	0.597*** (0.168)	0.299*** (0.030)						
Num. of Inst. in same Prov. and Ind.	-0.042 (0.034)	-0.023 (0.027)	-0.042 (0.034)	-0.008 (0.027)				
Prov. Patents Per Person			0.132 (0.092)	0.073*** (0.017)				
Natl Ind. Knowledge Stock (KS)					-0.058 (0.296)	0.400*** (0.112)		
Distance from Natl KS					-0.356 (0.633)	0.015 (0.228)		
Distance*Natl Ind. KS					0.025 (0.044)	-0.019 (0.016)		
Prov. Sector KS							0.031 (0.035)	0.196*** (0.013)
Distance from Prov. Sector KS							0.047 (0.058)	0.093*** (0.022)
Distance*Prov. Sector KS							-0.006 (0.006)	-0.013*** (0.002)
Constant	-17.954*** (2.759)		-8.384*** (0.251)		-7.343* (4.324)		-8.714*** (0.511)	
Observations	32,526	14,932	32,526	14,932	34,756	16,251	34,296	15,905
Year Fixed Effect	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Area and Sector Fixed Effects	Yes	No	Yes	No	Yes	No	Yes	No
Institute Fixed Effects	No	Yes	No	Yes	No	Yes	No	Yes

Note: All variables in logarithms. All input variables are three year backward moving averages lagged one year. Robust standard errors in parentheses. *** p<0.01, ** p<0.05, * p<0.1.

Table 6: Utility Patents

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
	Util Pat	Util Pat	Util Pat	Util Pat	Util Pat	Util Pat	Util Pat	Util Pat
Total Personnel	0.511*** (0.070)	0.332*** (0.042)	0.509*** (0.070)	0.358*** (0.042)	0.478*** (0.068)	0.397*** (0.040)	0.467*** (0.071)	0.411*** (0.040)
S&T Personnel	0.138* (0.075)	0.101** (0.045)	0.141* (0.075)	0.098** (0.045)	0.154** (0.071)	0.094** (0.041)	0.171** (0.074)	0.132*** (0.041)
Fraction R&D Personnel	-0.030 (0.114)	-0.758*** (0.040)	-0.035 (0.114)	-0.796*** (0.040)	0.028 (0.101)	-0.677*** (0.039)	0.022 (0.099)	-0.684*** (0.039)
S&T Equipment	0.196*** (0.021)	0.113*** (0.008)	0.195*** (0.021)	0.112*** (0.008)	0.192*** (0.021)	0.070*** (0.008)	0.171*** (0.021)	0.079*** (0.008)
Fraction R&D Equipment	0.555*** (0.088)	0.308*** (0.031)	0.559*** (0.088)	0.343*** (0.031)	0.448*** (0.093)	0.271*** (0.029)	0.413*** (0.096)	0.198*** (0.030)
S&T Intermediate Inputs	0.173*** (0.024)	0.040*** (0.010)	0.174*** (0.024)	0.045*** (0.010)	0.200*** (0.025)	0.058*** (0.009)	0.214*** (0.026)	0.052*** (0.010)
Fraction R&D Inputs	0.028** (0.011)	0.015* (0.009)	0.028*** (0.011)	0.016* (0.009)	0.034*** (0.010)	0.019*** (0.007)	0.035*** (0.011)	0.017*** (0.007)
Prov. R&D per worker	0.321* (0.171)	0.557*** (0.045)						
Num. of Inst. in same Prov. and Ind.	-0.106*** (0.028)	-0.095** (0.048)	-0.106*** (0.028)	-0.066 (0.048)				
Prov. Patents Per Person			0.021 (0.093)	0.068*** (0.023)				
Natl Ind. Knowledge Stock (KS)					-0.459* (0.237)	-0.386** (0.152)		
Distance from Natl KS					-0.689 (0.500)	-1.048*** (0.332)		
Distance*Natl Ind. KS					0.058* (0.035)	0.086*** (0.022)		
Prov. Sector KS							-0.046** (0.023)	0.164*** (0.019)
Distance from Prov. Sector KS							0.047 (0.038)	0.060** (0.024)
Distance*Prov. Sector KS							-0.003 (0.004)	-0.010*** (0.002)
Constant	-12.357*** (2.821)		-7.157*** (0.244)		-1.942 (3.391)		-6.825*** (0.380)	
Observations	32,526	13,707	32,526	13,707	34,756	14,916	34,296	14,592
Year Fixed Effect	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Area and Sector Fixed Effects	Yes	No	Yes	No	Yes	No	Yes	No
Institute Fixed Effects	No	Yes	No	Yes	No	Yes	No	Yes

Note: All variables in logarithms. All input variables are three year backward moving averages lagged one year. Robust standard errors in parentheses. *** p<0.01, ** p<0.05, * p<0.1.

Table 7: Scientific Papers

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
	Papers	Papers	Papers	Papers	Papers	Papers	Papers	Papers
Total Personnel	0.035 (0.031)	0.101*** (0.008)	0.035 (0.032)	0.100*** (0.008)	0.034 (0.030)	0.140*** (0.008)	0.041 (0.031)	0.121*** (0.008)
S&T Personnel	0.705*** (0.035)	0.256*** (0.008)	0.705*** (0.035)	0.257*** (0.008)	0.676*** (0.033)	0.246*** (0.008)	0.704*** (0.034)	0.267*** (0.008)
Fraction R&D Personnel	0.163*** (0.016)	0.069*** (0.004)	0.163*** (0.016)	0.068*** (0.004)	0.152*** (0.015)	0.079*** (0.004)	0.166*** (0.016)	0.064*** (0.004)
S&T Equipment	0.021*** (0.008)	0.037*** (0.001)	0.021*** (0.008)	0.037*** (0.001)	0.029*** (0.007)	0.032*** (0.001)	0.021*** (0.008)	0.036*** (0.001)
Fraction R&D Equipment	0.642*** (0.026)	0.092*** (0.005)	0.642*** (0.026)	0.093*** (0.005)	0.443*** (0.025)	0.090*** (0.005)	0.645*** (0.025)	0.083*** (0.005)
S&T Intermediate Inputs	0.132*** (0.009)	0.018*** (0.002)	0.132*** (0.009)	0.019*** (0.002)	0.125*** (0.008)	0.014*** (0.002)	0.128*** (0.008)	0.019*** (0.002)
Fraction R&D Inputs	0.037*** (0.005)	0.011*** (0.001)	0.037*** (0.005)	0.011*** (0.001)	0.031*** (0.005)	0.010*** (0.001)	0.036*** (0.005)	0.011*** (0.001)
Prov. R&D per worker	0.020 (0.057)	0.019*** (0.007)						
Num. of Inst. in same Prov. and Ind.	0.015 (0.011)	-0.164*** (0.009)	0.015 (0.011)	-0.163*** (0.009)				
Prov. Patents Per Person			0.049 (0.035)	0.010*** (0.004)				
Natl Ind. Knowledge Stock (KS)					0.232** (0.095)	-0.023 (0.026)		
Distance from Natl KS					0.171 (0.207)	-0.481*** (0.056)		
Distance*Natl Ind. KS					0.005 (0.014)	0.043*** (0.004)		
Prov. Sector KS							0.038*** (0.013)	0.014*** (0.003)
Distance from Prov. Sector KS							0.095*** (0.021)	0.005 (0.004)
Distance*Prov. Sector KS							-0.012*** (0.002)	-0.002*** (0.000)
Constant	-1.541* (0.930)		-1.285*** (0.083)		-5.711*** (1.455)		-1.527*** (0.197)	
Observations	32,526	28,165	32,526	28,165	34,756	30,093	34,296	29,653
Year Fixed Effect	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Area and Sector Fixed Effects	Yes	No	Yes	No	Yes	No	Yes	No
Institute Fixed Effects	No	Yes	No	Yes	No	Yes	No	Yes

Note: All variables in logarithms. All input variables are three year backward moving averages lagged one year. Robust standard errors in parentheses. *** p<0.01, ** p<0.05, * p<0.1.

Table 8: Spillovers to Total Revenue

	(1) Total Rev.	(2) Total Rev.	(3) Total Rev.	(4) Total Rev.	(5) Total Rev.	(6) Total Rev.	(7) Total Rev.	(8) Total Rev.
Total Personnel	0.747*** (0.006)	0.549*** (0.027)	0.747*** (0.006)	0.549*** (0.027)	0.747*** (0.006)	0.544*** (0.027)	0.747*** (0.006)	0.544*** (0.028)
Fraction S&T Personnel	-0.056 (0.036)	0.109 (0.067)	-0.055 (0.036)	0.111* (0.067)	-0.060* (0.036)	0.113* (0.067)	-0.058 (0.035)	0.113* (0.065)
Fraction P&M Personnel	0.154*** (0.041)	0.064 (0.076)	0.154*** (0.041)	0.065 (0.076)	0.150*** (0.041)	0.066 (0.076)	0.165*** (0.041)	0.051 (0.075)
S&T Equipment	0.140*** (0.003)	0.057*** (0.004)	0.140*** (0.003)	0.057*** (0.004)	0.140*** (0.003)	0.057*** (0.004)	0.138*** (0.003)	0.056*** (0.004)
S&T Intermediate Inputs	0.247*** (0.004)	0.093*** (0.005)	0.247*** (0.004)	0.092*** (0.005)	0.247*** (0.004)	0.093*** (0.005)	0.243*** (0.004)	0.093*** (0.005)
Natl Ind. R.I. Rev.					0.135*** (0.044)	0.047 (0.121)		
Distance from Natl. Ind. R.I. Rev.					0.338*** (0.096)	-0.096 (0.267)		
Distance*Natl Ind. Rev.					-0.019*** (0.006)	0.007 (0.017)		
Prov. R&D Expenditure	-0.008 (0.016)	-0.008 (0.021)						
Prov. Sales to RGDP			0.019 (0.030)	0.086*** (0.033)				
Prov. Sector R.I. Rev.							0.047*** (0.006)	0.101*** (0.018)
Distance from Prov. Sector R.I. Rev.							0.092*** (0.009)	0.083*** (0.030)
Distance*Prov.Sector R.I. Rev.							-0.010*** (0.001)	-0.008*** (0.003)
Constant	2.883*** (0.253)	4.826*** (0.326)	2.785*** (0.063)	4.808*** (0.135)	0.394 (0.677)	4.064** (1.878)	2.257*** (0.086)	3.595*** (0.244)
Observations	40,944	40,944	40,944	40,944	40,939	40,939	40,478	40,478
R-squared	0.873	0.595	0.873	0.595	0.873	0.596	0.874	0.598
Year Fixed Effect	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Area and Sector Fixed Effects	Yes	No	Yes	No	Yes	No	Yes	No
Institute Fixed Effects	No	Yes	No	Yes	No	Yes	No	Yes

Note: All variables in logarithms except fractions. All input variables are three year backward moving averages lagged one year. Robust standard errors in parentheses. *** p<0.01, ** p<0.05, * p<0.1.

Table 9: Spillovers to S&T Revenue

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
	S&T Rev.	S&T Rev.	S&T Rev.	S&T Rev.	S&T Rev.	S&T Rev.	S&T Rev.	S&T Rev.
Total Personnel	0.662*** (0.007)	0.479*** (0.033)	0.662*** (0.007)	0.479*** (0.033)	0.660*** (0.007)	0.477*** (0.033)	0.665*** (0.007)	0.479*** (0.034)
Fraction S&T Personnel	-0.136*** (0.041)	0.067 (0.090)	-0.134*** (0.041)	0.071 (0.090)	-0.141*** (0.041)	0.068 (0.090)	-0.132*** (0.041)	0.071 (0.089)
Fraction P&M Personnel	-1.186*** (0.049)	-0.663*** (0.105)	-1.185*** (0.050)	-0.663*** (0.104)	-1.186*** (0.050)	-0.653*** (0.105)	-1.180*** (0.050)	-0.667*** (0.104)
S&T Equipment	0.120*** (0.003)	0.053*** (0.004)	0.120*** (0.003)	0.053*** (0.004)	0.120*** (0.003)	0.052*** (0.004)	0.119*** (0.003)	0.052*** (0.004)
S&T Intermediate Inputs	0.300*** (0.004)	0.143*** (0.007)	0.299*** (0.004)	0.143*** (0.007)	0.299*** (0.004)	0.142*** (0.007)	0.295*** (0.004)	0.141*** (0.007)
Natl Ind. R.I. Rev.					0.178*** (0.054)	-0.463*** (0.143)		
Distance from Natl. Ind. R.I. Rev.					0.331*** (0.120)	-1.075*** (0.310)		
Distance*Natl Ind. Rev.					-0.023*** (0.008)	0.072*** (0.021)		
Prov. R&D Expenditure	-0.051*** (0.017)	-0.051** (0.024)						
Prov. Sales to RGDP			0.032 (0.033)	0.100*** (0.038)				
Prov. Sector R.I. Rev.							0.036*** (0.006)	0.095*** (0.018)
Distance from Prov. Sector R.I. Rev.							0.081*** (0.010)	0.027 (0.028)
Distance*Prov.Sector R.I. Rev.							-0.009*** (0.001)	-0.003 (0.003)
Constant	3.985*** (0.273)	5.490*** (0.372)	3.235*** (0.071)	4.871*** (0.168)	0.611 (0.845)	11.745*** (2.160)	2.822*** (0.096)	3.645*** (0.255)
Observations	40,508	40,508	40,508	40,508	40,503	40,503	40,048	40,048
R-squared	0.832	0.440	0.832	0.440	0.832	0.440	0.833	0.443
Year Fixed Effect	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Area and Sector Fixed Effects	Yes	No	Yes	No	Yes	No	Yes	No
Institute Fixed Effects	No	Yes	No	Yes	No	Yes	No	Yes

Note: All variables in logarithms. All input variables are three year backward moving averages lagged one year. Robust standard errors in parentheses. *** p<0.01, ** p<0.05, * p<0.1.

Table 10: Spillovers to P&M Revenue

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
	P&M Rev.	P&M Rev.	P&M Rev.	P&M Rev.	P&M Rev.	P&M Rev.	P&M Rev.	P&M Rev.
Total Personnel	0.775*** (0.022)	0.555*** (0.072)	0.776*** (0.022)	0.552*** (0.072)	0.790*** (0.022)	0.544*** (0.072)	0.767*** (0.023)	0.540*** (0.072)
Fraction S&T Personnel	-0.378*** (0.143)	-0.007 (0.254)	-0.372*** (0.143)	-0.001 (0.255)	-0.323** (0.143)	-0.028 (0.254)	-0.418*** (0.144)	-0.049 (0.253)
Fraction P&M Personnel	2.129*** (0.145)	0.654** (0.269)	2.129*** (0.145)	0.647** (0.270)	2.152*** (0.145)	0.576** (0.269)	2.163*** (0.145)	0.563** (0.267)
S&T Equipment	0.141*** (0.010)	0.081*** (0.013)	0.140*** (0.010)	0.080*** (0.013)	0.141*** (0.010)	0.081*** (0.013)	0.136*** (0.010)	0.081*** (0.013)
S&T Intermediate Inputs	0.059*** (0.012)	-0.020 (0.017)	0.059*** (0.012)	-0.020 (0.017)	0.059*** (0.012)	-0.016 (0.017)	0.061*** (0.012)	-0.020 (0.017)
Natl Ind. R.I. Rev.					0.018 (0.197)	1.532*** (0.510)		
Distance from Natl. Ind. R.I. Rev.					0.309 (0.421)	2.401** (1.098)		
Distance*Natl Ind. Rev.					-0.023 (0.028)	-0.182** (0.074)		
Prov. R&D Expenditure	-0.096 (0.071)	-0.111 (0.091)						
Prov. Sales to RGDP			0.189 (0.131)	0.074 (0.138)				
Prov. Sector R.I. Rev.							-0.036* (0.020)	0.184*** (0.056)
Distance from Prov. Sector R.I. Rev.							0.025 (0.034)	0.310*** (0.099)
Distance*Prov.Sector R.I. Rev.							-0.005 (0.003)	-0.030*** (0.009)
Constant	2.151* (1.134)	5.247*** (1.341)	0.926*** (0.254)	3.745*** (0.446)	0.304 (2.930)	-16.705** (7.666)	1.369*** (0.310)	1.741** (0.743)
Observations	16,266	16,266	16,266	16,266	16,261	16,261	16,027	16,027
R-squared	0.534	0.203	0.534	0.202	0.537	0.209	0.535	0.205
Year Fixed Effect	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Area and Sector Fixed Effects	Yes	No	Yes	No	Yes	No	Yes	No
Institute Fixed Effects	No	Yes	No	Yes	No	Yes	No	Yes

Note: All variables in logarithms. All input variables are three year backward moving averages lagged one year. Robust standard errors in parentheses. *** p<0.01, ** p<0.05, * p<0.1.

Table 11: Government Grants

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
	Gov. Grant	Gov. Grant	Gov. Grant	Gov. Grant	Gov. Grant	Gov. Grant	Gov. Grant	Gov. Grant
Total Personnel	0.625*** (0.008)	0.382*** (0.030)	0.625*** (0.008)	0.382*** (0.030)	0.616*** (0.008)	0.381*** (0.030)	0.629*** (0.008)	0.379*** (0.030)
Fraction S&T Personnel	-0.230*** (0.046)	0.236*** (0.080)	-0.229*** (0.046)	0.239*** (0.080)	-0.254*** (0.046)	0.235*** (0.080)	-0.216*** (0.046)	0.247*** (0.080)
Fraction P&M Personnel	-1.324*** (0.053)	-0.259*** (0.092)	-1.324*** (0.053)	-0.260*** (0.092)	-1.315*** (0.054)	-0.257*** (0.092)	-1.321*** (0.053)	-0.252*** (0.092)
S&T Equipment	0.095*** (0.003)	0.043*** (0.004)	0.095*** (0.003)	0.044*** (0.004)	0.095*** (0.003)	0.043*** (0.004)	0.094*** (0.003)	0.043*** (0.004)
S&T Intermediate Inputs	0.235*** (0.005)	0.099*** (0.006)	0.235*** (0.005)	0.098*** (0.006)	0.234*** (0.005)	0.099*** (0.006)	0.230*** (0.005)	0.100*** (0.006)
Natl Ind. R.I. Rev.					0.563*** (0.065)	-0.232* (0.138)		
Distance from Natl. Ind. R.I. Rev.					1.039*** (0.144)	-0.552* (0.310)		
Distance*Natl Ind. Rev.					-0.069*** (0.009)	0.034* (0.020)		
Prov. R&D Expenditure	-0.019 (0.021)	0.005 (0.024)						
Prov. Sales to RGDP			0.075* (0.039)	0.179*** (0.041)				
Prov. Sector R.I. Rev.							0.020** (0.008)	0.065*** (0.018)
Distance from Prov. Sector R.I. Rev.							0.081*** (0.012)	0.001 (0.024)
Distance*Prov.Sector R.I. Rev.							-0.010*** (0.001)	0.000 (0.002)
Constant	3.947*** (0.335)	4.640*** (0.377)	3.756*** (0.083)	4.898*** (0.157)	-4.694*** (1.020)	8.469*** (2.201)	3.537*** (0.116)	3.916*** (0.265)
Observations	38,415	38,415	38,415	38,415	38,410	38,410	38,054	38,054
R-squared	0.740	0.495	0.740	0.496	0.743	0.496	0.744	0.500
Year Fixed Effect	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Area and Sector Fixed Effects	Yes	No	Yes	No	Yes	No	Yes	No
Institute Fixed Effects	No	Yes	No	Yes	No	Yes	No	Yes

Note: All variables in logarithms. All input variables are three year backward moving averages lagged one year. Robust standard errors in parentheses. *** p<0.01, ** p<0.05, * p<0.1.

Appendix

A. Sectors and Industries

This section contains a numbered list of the sectors and the corresponding sub industries with industry numbers.

1. Agriculture

1. Farming
2. Forestry
3. Animal Husbandry
4. Fishery
5. Service Activities for Farming, Forestry, Animal Husbandry, and Fishery

2. Mining

6. Mining and Washing of Coal
7. Extraction of Petroleum and Natural Gas

3. Low-tech

13. Processing of Food from Agricultural Products
14. Manufacture of Foods
15. Manufacture of Liquor, Beverages and Refined Tea
16. Manufacture of Tobacco
17. Manufacture of Textile
18. Manufacture of Textile Apparel and Accessories
19. Manufacture of Leather, Fur, Feather and Related Products and Shoes
20. Processing of Timbers and Manufacture of Wood, Bamboo, Rattan, Palm and Straw
21. Manufacture of Furniture
22. Manufacture of Paper and Paper Products
23. Printing, Reproduction of Recording Media
24. Manufacture of Artworks and Articles for Culture, Education, Sports and Recreation

4. Med-Low Tech

25. Processing of Petroleum, Cooking and Processing of Nucleus Fuel
29. Manufacture of Rubber and Plastic
30. Manufacture of Non-metallic Mineral Products
31. Manufacture and Processing of Ferrous Metals
32. Manufacture and Processing of Non-ferrous Metals
33. Manufacture of Metal Products

5. Med-High Tech

26. Manufacture of Chemical Raw Material and Chemical Products
28. Manufacture of Chemical Fiber
34. Manufacture of General Purpose Machinery
35. Manufacture of Special Purpose Machinery
36. Manufacture of Motor Vehicles
38. Manufacture of Electrical Machinery and Equipment.

6. High Tech

- 27 Manufacture of Medicines
- 37 Manufacture of Railway Equipment, Aerospace Equipment and Other Transport
- 39 Manufacture of Computer, Communication and Other Electronic Equipment
- 40 Manufacture of Measuring Instrument and Meter

7. High-Tech Service

- 63 Telecommunications, Broadcasting, Television and Satellite Transmission Services
- 64 Internet and Related Services
- 65 Software and Information Technology Services
- 73 Research and Experimental Development
- 74 Professional Technique Services
- 75 Technique Generalization and Application Services

8. Sanitation

- 83 Sanitation

9. Electric and Gas

- 44. Production and Supply of Electric Power and Heat Power
- 45 Production and Distribution of Gas

10. Other

- 41 Other Manufacturing
- 42 Waster Recycling and Recovery
- 43 Repair Service of Metal Products, Machinery and Equipment

B. Provinces

Source: https://all-free-download.com/free-vector/download/china-political-map-vector_587825.html

11	Beijing	41	Henan
12	Tianjin	42	Hubei
13	Hebei	43	Hunan
14	Shanxi	44	Guangdong
15	Inner Mongolia	45	Guangxi
21	Liaoning	46	Hainan
22	Jilin	50	Chongqing
23	Heilongjiang	51	Sichuan
31	Shanghai	52	Guizhou
32	Jiangsu	53	Yunnan
33	Zhejiang	54	Tibet
34	Anhui	61	Shaanxi
35	Fujian	62	Gansu
36	Jiangxi	63	Qinghai
37	Shandong	64	Ningxia
		65	Xinjiang

C. Supplementary Tables

Table A1: Summary Statistics of Distance by Industry and Province

<u>Industry</u>	<u>Mean</u>	<u>Std.</u> <u>Dev.</u>	<u>Industry</u>	<u>Mean</u>	<u>Std.</u> <u>Dev.</u>	<u>Province</u>	<u>Mean</u>	<u>Std.</u> <u>Dev.</u>
Farming	1,425	435	Manufacture of:			Beijing	10	5
Forestry	1,460	457	Mineral Products	1,077	312	Tianjin	9	10
Animal Husbandry	1,591	409	Ferrous Metals	1,089	422	Hebei	213	68
Fishery	1,097	381	Non-ferrous Metals	1,162	329	Shanxi	151	78
Services for Agriculture	1,337	362	Metal Products	735	186	Inner Mongolia	497	253
Coal	1,085	360	General Purpose Machinery	1,035	311	Liaoning	159	57
Petroleum and Natural Gas	602	256	Special Purpose Machinery	1,150	293	Jilin	112	84
Food Processing	1,227	312	Motor Vehicles	1,069	310	Heilongjiang	207	115
Manufacture of:			Transport Equipment	891	247	Shanghai	8	4
Foods	1,177	326	Electrical Machinery	937	249	Jiangsu	155	58
Liquor	1,087	120	Electronic Equipment	966	229	Zhejiang	123	49
Tobacco	639	278	Measuring Instruments	963	264	Anhui	116	52
Textile	932	251	Other	839	300	Fujian	113	49
Apparel	1,255	492	Waste Recycling and Recovery	0	0	Jiangxi	154	64
Leather	994	263	Repair of Metal Products	192	298	Shandong	191	63
Wood	781	113	Production and Electricity and Heat	1,183	376	Henan	126	60
Furniture	611	251	Production and Distribution of Gas	396	106	Hubei	158	77
Paper	848	165	Telecommunications	960	242	Hunan	127	62
Artworks	1,068	276	Internet and Related Services	1,056	376	Guangdong	161	97
Chemicals	1,045	240	Software and ITS	1,284	353	Guangxi	175	61
Medicines	1,260	373	Research and Development	1,269	391	Hainan	83	31
Fibers	1,118	350	Professional Technique Services	1,221	352	Chongqing	41	40
Rubber and Plastic	1,022	250	Technique Generalization	1,138	289	Sichuan	177	97
Printing and Media	1,271	476	Sanitation	1,137	357	Guizhou	105	62
Processing of Petroleum	1,017	386				Yunnan	208	98
			Total	1,221	395	Tibet	113	114
						Shaanxi	128	82
						Gansu	193	141
						Qinghai	79	96
						Ningxia	35	48
						Xinjiang	328	233
						Total	145	126

Note: This table gives summary statistics for distance (km) by industry and province. Complete industry names are given in the appendix.

Table A2: Distribution of Research Institutes by Industry and Province 1998																																		
		11	12	13	14	15	21	22	23	31	32	33	34	35	36	37	41	42	43	44	45	46	50	51	52	53	54	61	62	63	64	65	total	
Agriculture																																		
	1	9	5	20	21	12	15	10	30	7	18	18	18	12	13	26	29	25	25	37	23	11	11	24	21	29	4	15	15	2	10	25	540	
	2	7		5	5	8	9	3	10		4	5	6	2	6	10	8	10	10	18	10	1	3	10	7	10	1	5	8		1	10	192	
	3	3	2	2	4	4	3	3	6	2	4	1	2		2		2	2	1	11	3	1	1	3	2	3		6	2		1	5	81	
	4			1		1	5	1	3	1	9	3	3	3	2	8	4	8	3	8	3	1	2		1	2					1	1	74	
	5	6	4	10	11	16	19	13	17	2	12	4	13	16	7	14	7	22	19	13	5	5		15	11	13	1	13	11	4	8	14	325	
Mining																																		
	6	2		2	3	1	2	1			3	1	1	1	1	2	1				1		1		1	1		2	1		1	1	30	
	7	2	1	1							1	1				2		3						1				1	1				14	
Low-tech																																		
	13	3	2	3	2	5	7	1	6	1	5	2	3	2	2	7	2	6	2	7	2	1		2	1	3		3	2			1	83	
	14	6	2	1	3	1	6	2	2	5	3		1	1	1	8	1	2	2	8	2	1	3	2	2	1		1	1	1			69	
	15			1	1									1								1			1			1					6	
	16																3								1								4	
	17	1	4	3	1		6		2	3	8	4	2	1	1	8	4	1	3	1			1	2		1		1	1	1			60	
	18	2	1	1		1	4		4	2	4	1	3		1	3		1		1	1						2	1			2		35	
	19	3	1	1		1	2	1		2						2	1		1	1				1				1					18	
	20	1				1			1		1																						3	
	21	1	1			1		1	1				1			2							1										9	
	22	1	2							1		1				2	1		1				1										10	
	23	2			1		2		1	2			1			1		1		1			1			1					1		15	
	24	3	2		3				2	1	4	3		2	1	3		1	1	1	1				2	1			1				32	
Med-Low-tech																																		
	25	3					1					1								1													6	
	29	4	2	1	1	1	4	3		2	4		2			2		2	2	2			1			1		1	1				36	
	30	5	3	3	2	1	10	2	4	4	8	3	2	2	4	8	2	4	1	5	2			4	1	1		2	1		1	1	86	
	31	3	2	1	1		2		1	3						2	1	4		1				1								1	23	
	32	1	1			3				3	2	1	1	1	2	1	1		2				1	2	2	2				1	1		28	
	33	2	4				2			4	1		1	1	1	1	1	1					2					1	1				23	
Med-High-tech																																		
	26	3	7	4	7	3	17	7	4	12	16	7	7	6	4	13	7	9	8	8	4		2	8	4	4		6	1				178	
	28		1		1		1	1	1		1					1				1	1												9	
	34	10	5	4		1	11		7	12	7	7	3	2		8	5	5		5	2		2	6	1	1		3	1			1	109	
	35	11	8	7	4	7	16	5	12	11	14	7	6	1	2	11	5	9	9	13	8		3	5	1	3		4	7		1	1	191	
	36	3	2		1		1	2		1	1					3		2			1	1	1										19	
	38	6	4	1	1	1	5	1	1	7	6		2			1		2		1			2	1		1		5	1				49	
High-tech																																		
	27	8	1	2	2		4	4	4	3	12	4	3	1	2	6	2	5	1	8	5	1	3	5	2	3		2	3		1	3	100	
	37	1	1				1	2	1	2		2			1	1		2	1				2										17	
	39	12	4	5	3		9	1	2	9	16	4	7	1	1	8	2	8	8	9	2		2	4				8	1				126	
	40	10	5	4	2	1	7	4	4	8	5	3	6	2		6		2	1	1			3			1		1	4				80	
High-Tech Service																																		
	63	7		2	1		2			1	2	2	1			1	1	1	1	2				1		1		1		1			28	
	64	2			1				1				1				1				1												7	
	65	6	1	1	2	2	7	1	1	2	1	5	1	1	1	2	4	4	3	5	5		1	4	2	3			1	1	1	3	71	
	73	41	6		2		9	18	13	20	7	10	13	1	11	19	3	7	13	13	4	8		8	7	9	2	8	9	3	3	9	276	
	74	35	17	17	9	4	29	16	17	12	24	13	24	12	10	33	12	28	13	29	6	3	6	25	8	7	2	11	14	15	8	10	469	
	75	10	3	4	13	12	9	11	7	5	3	14	10	7	7	12	15	7	13	7	5		1	11	7	4		10	3	2	4	2	218	
Sanitation																																		
	83	29	3	7	22	8	6	5	12	12	17	2	9	13	4	25	10	4	3	9	11	4	2	11	3	5		7	3	2	2	3	253	
Elec and Gas																																		
	44	4	2		1	4	2	1		1	7	2		2	1		1	1	1	1					1		1	1	2	1		1	38	
	45	1					1																										2	
Other																																		
	41				1		1				1	1	2						1	1				1		1		1					11	
	43													1								1											2	
	total	264	107	114	130	94	234	117	178	162	224	127	155	92	87	262	135	188	147	227	108	39	59	156	87	112	10	120	96	32	44	95	4055	

Sector ind./prov.		11	12	13	14	15	21	22	23	31	32	33	34	35	36	37	41	42	43	44	45	46	50	51	52	53	54	61	62	63	64	65	total
Agriculture																																	
	1	6	5	17	20	13	16	10	30	8	16	14	16	13	13	24	27	23	27	33	24	11	8	23	19	19	4	5	14	2	12	24	496
	2	5		5	5	8	9	3	9		4	5	5	2	6	7	8	11	9	19	10	1	1	9	6	9		2	8		1	13	180
	3	3	2	1	3	4	2	3	6	2	3		2		2		1	1	1	10	2	1		3	2	2		4	2			8	70
	4			1		1	3	1	3	1	10	2	3	3	2	6	3	8	3	8	3	1	1		1	2					1	1	68
	5	9	5	9	11	15	16	13	19	2	10	5	11	17	7	14	6	16	18	14	2	6	1	11	9	11	1	11	11	4	9	5	298
Mining																																	
	6	2		2	2	1	2	1			3	1	1	1	1	3	1				1		1		1	1		2	1		1	1	30
	7	4	1								1					2												1					9
Low-tech																																	
	13	3	2		1	3	6	1	3	1	3	2	1	2	2	5	2	4	3	6	2	1		2	1	3		3	1			1	64
	14	4	2	1	3	1	3	1	2	5					1	4	1	2		4	2			2	1			1				1	41
	15			1	1									1								1				1			1			6	
	16																2									1						3	
	17	1	4	2	1		4		2	3	6	2		1	1	4	4	1	3	1			1	2		1		1	1			46	
	18		1			3		4	2	3	1	2			1	2		1		1	1						1	1			2	26	
	19	1		1		1	1		1							1	1		1	1				1								10	
	20	1							1																							2	
	21		1						1	1						2																	5
	22	1	2							1		1				2	1		1				1										10
	23	2			1		2		1	2	1							1								1					1		12
	24	3	1		3				2	1	5	3		1	1	3		1	1		1				2	1			1				30
Med-Low-tech						</																											

Sector ind/prov	11	12	13	14	15	21	22	23	31	32	33	34	35	36	37	41	42	43	44	45	46	50	51	52	53	54	61	62	63	64	65	total		
Agriculture																																		
1	7	5	18	20	11	16	7	30	7	15	14	14	13	14	23	24	22	28	32	26	10	4	21	20	18	6	5	19	2	10	26	487		
2	6		3	6	8	8	3	8		4	5	4	2	6	6	6	10	9	15	10	1	1	8	6	9	1		6		1	12	164		
3	3	2	1	3	3	2	3	6	2	3		2		2	1	1	1	1	10	3	1		3	2	1	1	3	3		1	8	72		
4			1		1	4	1	3	1	8	2	3	1	5	2	8	3	6	3	1	1		1	2						1	1	61		
5	13	7	9	10	10	14	11	21	2	6	5	8	17	8	14	5	15	17	15	1	6	2	10	7	10		11	14	4	6	4	282		
Mining																																		
6	2		2	3	1	2	1			2	1	1	1	1	3	1				1		1		1	1		2	1		1	1	30		
7	7	1								1					1												1					11		
Low-tech																																		
13	4	2		1	3	6	1	2	1	2	2	1	2	2	4		4	1	5	3	1		2	1	4		2	1		1		58		
14	4	2	1	2	1		1	1	4				1	3		1	1		2	2		2	1				1			1		31		
15				1									1						2	2			1									5		
16																2					1				1							3		
17	1	3	2	1		1		1	3	2	2		1	1	3	2	1	3	1			1	2					1					32	
18		1				2		2	1	2	1	1		1			1		1	1							1	1			2		18	
19	1					1	1		1						1			1					1										7	
20	1							1																									2	
21		1							1						1																			7
22	1										1				2	1		1				1												3
23	2			1		2		1	2	1								1								1					1			12
24	2	1		2				2		3	3		1	1	3		1	1		1				2	1				1					25
Med-Low-tech																																		
25	2					1		1			1								1															6
29	3	2		1		2	1		1	2					1				1	1				1	1			1	1					19
30	5	3		2	1	2	1	2	1	5	2	1	2	4	5	3																		