

# Global Innovation Spillovers and Productivity: Evidence from 100 years of World Patent Data\*

Enrico Berkes

Ohio State

Kristina Manysheva

Northwestern

Martí Mestieri

FRB Chicago & Northwestern

This version: October 25, 2021

[\[Please click here for the latest version\]](#)

## Abstract

We use a panel of historical patent data covering a large range of countries over the past century to study the evolution of innovation across time and space and its effect on productivity. We document a substantial rise of international knowledge spillovers as measured by patent citations since the 1990s. This rise is mostly accounted for by an increase in citations to US and Japanese patents in fields of knowledge related to computation, information processing, and medicine. We estimate the causal effect of innovation induced by international spillovers on output per worker and TFP growth in a panel of countries-sectors from 2000 to 2014. To assess causality, we develop a shift-share instrument that leverages pre-existing citation linkages across countries and fields of knowledge, and heterogeneous countries' exposure to technology waves. On average, an increase of one standard deviation in log-patenting activity increases output per worker growth by 4.7%. We find an effect of similar magnitude when considering long-run income per capita growth for the post-war period.

*Keywords:* Innovation, Technology Diffusion, Patents.

*JEL Classification:* O10, O30, O33, O47.

---

\*We thank Joel David, Matthias Doepke, Ruben Gaetani, Giammario Impullitti, Ben Jones, Nan Li, Joel Mokyr, Thomas Sampson, Sergio Petralia and attendees at CREI, LSE, UCSD-UCLA-UCB trade conference, Midwest Macro, Nottingham, SED, UB, UCSD, NBER SI Growth Meeting and Northwestern for helpful comments and discussions. We are also grateful to Bart Hobijn for his discussion of our paper. Mestieri thanks CREI for their hospitality during part of this research. The views expressed are the authors' and do not necessarily reflect the views of the Federal Reserve Bank of Chicago or the Federal Reserve System. All errors are our own.

# 1 Introduction

Productivity is a key driver of economic growth within and across countries. [Clark and Feenstra \(2003\)](#) and [Klenow and Rodríguez-Clare \(1997\)](#) document that the majority of the divergence in income per capita over the twentieth century can be attributed to cross-country differences in total factor productivity (TFP) growth. The endogenous growth literature, starting with the seminal contributions of [Romer \(1990\)](#) and [Aghion and Howitt \(1992\)](#), has emphasized the role of innovation and idea generation as a central driver of technology and, ultimately, productivity growth. However, from an empirical point of view, direct measures of innovation that cover a large number of technologies, countries, and time periods are scant.<sup>1</sup>

In this paper, we use historical patent data spanning a vast range of countries over the last one hundred years to study the evolution of innovation across time and space. The use of patent data allows us to exploit a widely validated quantitative measure for the generation of new ideas and knowledge spillovers (i.e., how innovation builds on previous knowledge). We document a substantial rise of international knowledge spillovers since the 1990s mostly driven by the United States and Japan and the rise of innovation related to computation, information and communication technologies (ICTs), and medicine. We also leverage the rich structure of citation linkages across time, space, and fields of knowledge to propose an identification strategy to quantify the effect of innovation induced by knowledge spillovers on productivity and economic growth across countries and industries and which is novel to this literature.

We build our measure of innovation using patent data collected from the European Patent Office Worldwide Patent Statistical Database (PATSTAT). PATSTAT contains bibliographical and legal status information on more than 110 million patents from the main patent offices around the world, covering leading industrialized countries, as well as developing countries over the period 1782-2018. To avoid some of the arbitrariness of using broad patent technology classes ([Keller, 2002](#)), we classify patents into “fields of knowledge” that we obtain with a machine-learning approach. Based on the premise that knowledge is embedded in inventors, the algorithm bundles together patent classes based on the probability that the same inventor patents in these classes to infer the proximity of classes in the knowledge space.<sup>2</sup>

Armed with our newly defined technology classes, we show that their significance – as

---

<sup>1</sup>[Comin and Hobijn \(2004, 2010\)](#) and [Comin and Mestieri \(2018\)](#) have analyzed the diffusion of major technologies since the Industrial Revolution. [Comin and Mestieri \(2018\)](#) show that the productivity transitional dynamics implied by the observed diffusion patterns match well the evolution of the distribution of cross-country income per capita in the last two centuries. However, their analysis is circumscribed to 25 major technologies since 1780.

<sup>2</sup>As a robustness check, we also perform a clustering analysis where the strength of the linkages between different patent classes is based on citations and/or co-appearance of these classes on the same grant.

measured by the share of filed patents that goes to each field of knowledge – has importantly evolved over time. The data reveal substantial technological waves in the last one hundred years. Mechanical engineering accrued the largest share of innovations at the beginning of the twentieth century. Fields of knowledge related to chemistry and physics (e.g., macromolecular compounds) were the most prominent fields in the mid-century, while inventions related to medicine and the digital economy appear to be the most prevalent in the past decades. We also show that, while advanced economies account for the bulk of patenting activity, there is substantial variation in terms of countries’ specialization across fields of knowledge. Moreover, these patterns of specialization are heterogeneous over time.<sup>3</sup>

Next, we turn our attention to knowledge spillovers. We measure knowledge spillovers through citations across fields of knowledge and countries. For this exercise, we focus on the post 1970 sample for which we have data for virtually all countries in the world. We show that, for the average patent, citations tend to be biased towards domestic, as opposed to international, inventions and towards the same field of knowledge. We also document that, across all these categories, there is an upward trend over time in citations. That is, new patents tend to cite more other patents. A striking fact has emerged since the 1990s. Except for the US and Japan, international citations have grown *faster* than domestic citations. After the year 2000, excluding the US and Japan, international citations are more than twice more frequent than domestic citations. This finding suggests that the reliance on knowledge produced elsewhere – and particularly in the US and Japan – has increased over this period of time. Even for technology leaders like Germany or Great Britain, foreign citations now account for most of the citations. This fact may be interpreted as a decline in the prominence of European innovations relative to their US and Japanese counterparts. We also find that most of this increase is driven by a handful of fields of knowledge that are related to ICTs and medicine.

After having laid out these facts, we investigate the effect of innovation (as measured by patenting) on productivity and income. Our baseline exercise studies the effect of innovation induced by international spillovers on productivity in the latest part of the sample (2000-2014) for which we have high quality data on cross-country sectoral value added and TFP, as well as factors of production. We use patent data starting in 1970 to construct our instrument for this exercise. We then extend our analysis back in time and study directly the effect of innovation on long-run income growth (1980-2016 and 1960-2016), for which we use the full extent of our patent data to construct our instrument.

Simply correlating innovation and productivity (or income) is problematic due to mea-

---

<sup>3</sup>We also show that specialization in fields of knowledge tends to be clustered in space. Moreover, we document that inequality in patenting activity across countries has increased since the 2000s.

surement error (which would generate attenuation bias), potential reverse causality, and the presence of unobserved factors affecting simultaneously patenting and the dependent variables. Examples of such factors include financial or external shocks that affect both the output of a country and the amount of innovation produced. In this paper, we address these endogeneity concerns by constructing a shift-share instrument that leverages pre-existing knowledge linkages across countries and technologies and combines it with lagged foreign innovative output in other fields of knowledge and countries, in the spirit of [Acemoglu et al. \(2016\)](#) and [Berkes and Gaetani \(2018a\)](#). More precisely, we construct the instrument in two steps. First, we estimate the strength of the linkages across countries and fields of knowledge (measured by patent citations) in the pre-sample period. These linkages constitute our pre-determined *shares*. The *shifts* of our instrument for country and field of knowledge,  $c_o$  and  $k_o$ , are given by the patents filed in all other countries  $c_d \neq c_o$  and fields of knowledge  $k_d \neq k_o$  in the sample. We are thus implicitly assuming that the probability that patents in  $(c_d, k_d)$  generate a patent in  $(c_o, k_o)$  can be inferred from the network of patent citations.<sup>4</sup> Applying this procedure recursively, we obtain a predicted number of patents for each country and field of knowledge.

Our main variable of interest is value added per worker by country and sector (measured from the World Input Output Database) over the 2000-2014 period. The regression model includes controls that vary at the country-sector-time (e.g., sectoral capital and labor), as well as country-time and sector-time fixed effects to control for differential country and sectoral trends (e.g., overall patenting activity has been steadily growing over time). We find a robust effect of innovation on value added per employment growth. One standard deviation increase in patenting activity leads to a 0.047 standard deviation increase in output per worker growth, or equivalently, to an increase of 4.7% in output per worker growth. We then look on the effect of innovation on TFP growth, and find very similar results in magnitude to the one obtained using value added per employment as a dependent variable. We conduct a number of robustness checks to address concerns regarding the validity of the instrument such as the existence of demand-pull anticipatory effects that might be correlated with the contemporaneous state of the local economy. To do this, among the other things, we “reverse” the network of citations that we used to measure knowledge spillovers and calculate the amount of innovation we would have expected to observe *in the past* if the patenting activity was driven only by future demand. Reassuringly, we find no evidence supporting this alternative hypothesis.

We conclude our paper by extending our framework to study the effect of innovation on

---

<sup>4</sup>In fact, we refine this procedure and extend this logic to higher-order linkages to create our main instrument (see Section 5).

long-run growth. We reconstruct our shift-share instrument using patent data pre-1980 and estimate the effect of innovation on income per capita over the 1980-2016 period. We find a positive, significant coefficient that is very similar in magnitude to the elasticity of patents to value added per employment that we find for the period 2000-2014. In terms of magnitude, an increase in one standard deviation in patenting activity increases income per capita by 0.14 standard deviations.

**Related Literature** This paper relates to the vast and rich literature studying the link between innovation and productivity since the seminal work of [Griliches \(1979, 1986\)](#). Our paper focuses on knowledge spillovers and diffusion of technology. Knowledge spillovers have been extensively documented (e.g., [Jaffe et al., 1993](#) and [Murata et al., 2014](#)). However, most of this literature has focused on domestic spillovers, based on the premise that they are very localized. In this paper, we especially focus on international spillovers which have also been documented to be quantitatively important (e.g., [Eaton and Kortum, 1999](#); [Keller, 2002](#); [Keller and Yeaple, 2013](#); [Keller, 2004](#) provides an excellent survey). We contribute to this latter literature by documenting an increase of international spillovers since the 1990s and by leveraging international linkages to build our shift-share design and quantify the effect of innovation on productivity.

Our paper also relates to the recent strand of literature that have used historical patent data, e.g., [Nicholas \(2010\)](#), [Packalen and Bhattacharya \(2015\)](#), [Petrulia et al. \(2016\)](#) and [Akcigit et al. \(2017\)](#) to shed light on various linkages between innovation and long-run outcomes. One difference with most of this literature is that we extend our analysis beyond one country and aim to provide a global view. In this respect, our work is closest to [Bottazzi and Peri \(2003\)](#) who use R&D and patent data for European Regions in the 1977-1995 period to estimate research externalities.

Our shift-share instrumental approach is related to a number of papers that have used the network structure of citations to construct shift-share instruments. Our approach is most similar to [Berkes and Gaetani \(2018b\)](#), who construct a similar shift-share instrument across US cities and [Acemoglu et al. \(2016\)](#) who use a citation network to percolate innovations downstream and illustrate how technological progress builds upon itself. Both papers use only within country (US) variation.<sup>5</sup>

---

<sup>5</sup>A large number of papers have used more standard shift-share instruments in the innovation and productivity literature. For example, [Moretti et al. \(2019\)](#) estimate the effects of R&D subsidies and [Hornbeck and Moretti \(2019\)](#) estimate the effect of TFP growth in manufacturing across US cities.

## 2 Data

### 2.1 Data Sources

In this paper, we measure new ideas through patents data and productivity through value added and TFP data. Patent data are collected from the European Patent Office worldwide Patent Statistical Database (PATSTAT, Autumn 2018 version). PATSTAT contains bibliographical and legal status information on more than 110 million patents from the main patent offices around the world, covering leading industrialized countries, as well as developing countries over the period 1782-2018.<sup>6</sup> From PATSTAT, we collect information on patent filing years, inventor and assignee locations, citations, patent families, and technological classes. While PATSTAT provides the most comprehensive coverage of patenting activities worldwide, it also has some limitations (Kang and Tarasconi, 2016). The main limitation for our purposes is data availability in the earlier years. In fact, data along one or more dimensions are often missing for some countries in the years preceding 1970. We therefore split our sample into two groups of countries, that we use at different stages of our analysis. The first group is composed of six major technological leaders (the United States, Great Britain, France, Germany, the Soviet Union, and Switzerland)<sup>7</sup> for which all the patents' characteristics required by our analysis are available since 1920. The second group includes all the countries covered by PATSTAT and starts in 1970.<sup>8</sup> Appendix A provides more information about the composition of the samples and summary statistics.

We assign each patent to a geographical unit according to the country of residence of its inventor(s). If this information is not available, then the country of the assignee(s) or publication authority is used, instead. When a given patent is associated with multiple inventors or applicants from different countries or territories, we assign weights to these patents. The weights are computed assuming that each inventor or applicant contributed

---

<sup>6</sup>PATSTAT is increasingly popular in economics as it provides rich information on patents. Most of its use has focused on particular sectors, countries or time periods. See, among others, Coelli et al. (2016); Aghion et al. (2016); Akcigit et al. (2018); Philippe Aghion and Melitz (2018); Bloom et al. (2020); Dechezleprêtre et al. (2020).

<sup>7</sup>Note that to compare consistent geographical units over time, when appropriate, we aggregate the patents filed in the German Democratic Republic and the Federal Republic of Germany. Similarly, for the Soviet Union, we consider all the patents produced by Armenia, Azerbaijan, Belarus, Estonia, Georgia, Kazakhstan, Kyrgyzstan, Latvia, Lithuania, Moldova, Russia, Tajikistan, Turkmenistan, Ukraine, and Uzbekistan.

<sup>8</sup>For our empirical analysis, we exclude China from our sample due to a substantial rise in the number of Chinese patents since the 3rd revision of Patent law in China in 2008. While we see a sharp increase in total number of Chinese patents after the implementation of the new law, the same pattern is not observed in the number of Triadic patents which include patents filed jointly in the largest patent offices (i.e., the United States Patent and Trademark Office (USPTO), the European Patent Office (EPO), and the Japan Patent Office (JPO)). For more details see Appendix A.1.

equally to the development of the invention.<sup>9</sup> To avoid double-counting patents that are filed in more than one patent office, we restrict most of our analysis to patents that are the first in their (DOCDB) family. We further collect the full distribution of technology classes associated with each patent based on the International Patent Classification (IPC). For our analysis, we first consider all the fields at the subclass level (e.g., A01B) – for a total of 650 classes – and we then cluster them into consistent groups following the procedure outlined in Section 2.2. Finally, to capture when an idea was completed and abstract from potential bureaucratic delays that are orthogonal to innovative activities, in our analysis we use the patent filing years instead of the years in which patents were granted.<sup>10</sup>

We supplement the patent data with the World Input Output Database (WIOD). This database provides data on prices and quantities of inputs, outputs, and trade flows covering 43 countries and the Rest of the World for the period 2000-2014. The data are classified according to the International Standard Classification revision 4 (ISIC) for a total of 56 sectors. Using the World Input-Output Tables (WIOT) for each set of countries, sectors, and years, we construct trade flows, gross output, intermediate purchases, and value added expressed in US dollars. Additionally, from the Socio-Economic Accounts (SEA) in the WIOD, we collect industry-level data on employment, capital stocks, gross output, and value added at current and constant prices. These data allow us to compute country-sector TFP paths and also to compute trade in intermediate and final goods across country-sector pairs.<sup>11</sup>

## 2.2 Construction of Fields of Knowledge

Innovation is the process of creating new knowledge building on existing knowledge across different fields. To operationalize our goal of measuring innovation waves across time and space, we build on the vast existing literature that measures innovative activities through patent data. We propose grouping finely-defined patent classes into broader “fields of knowledge,” which taken together constitute what we refer to as the *technology space* of the world. This conceptualization also provides a mapping between our patent data and the analytical framework developed in Section 4.<sup>12</sup> We employ a novel approach to grouping patent technology classes based on inventors’ information. Our procedure is based on the likelihood that the same inventor produces inventions associated with different patent subclasses. The idea

---

<sup>9</sup>For example, if a given patent has four inventors, one from the US and three from the UK, then the patent will be split between the US and the UK with weights of 0.25 and 0.75, respectively.

<sup>10</sup>We discuss in more detail our data construction procedure in the Appendix A.1

<sup>11</sup>See details in the Appendix A.2. In the Appendix we also discuss the additional database we use (UNIDO INDSTAT2) for historical data on sectoral manufacturing output by country and the Penn World Data Tables.

<sup>12</sup>See Kay et al. (2014), Leydesdorff et al. (2014) and Nakamura et al. (2015) for alternative definitions of technology space based on patent technology classes.



is that, since knowledge is embedded in people, it is possible to cluster fields of knowledge based on the IPC subclasses in which the same inventors tend to patent.<sup>13</sup> More precisely, we build a probability matrix  $T_{642 \times 642}$ ,<sup>14</sup> where each element  $(i, j)$  is the probability that an inventor patents in IPC subclass  $i$  conditional on also having a patent assigned to subclass  $j$ .<sup>15</sup> For example, a mechanical engineer specialized in brakes will most likely patent in IPCs B60T "Vehicle Brakes or parts thereof" and F16D "Clutches, Brakes", which our algorithm correctly bundles together.<sup>16</sup>

To obtain a symmetric matrix for the cluster analysis, we apply the following transformation:

$$D_{ij} = 1 - (T_{ij} + T_{ji}) = D_{ji}$$

where each element in the dissimilarity matrix  $D$  is interpreted as a measure of distance between subclass  $i$  and subclass  $j$ . We use this matrix together with a *k-medoids* clustering algorithm to group the IPC subclasses into clusters. A *k-medoids* algorithm minimizes the distance within clusters by comparing all possible permutations of subclasses, conditional on a specific number of clusters,  $k$ . Each resulting cluster represents a separate field of knowledge. To determine the optimal number of clusters, we first compute the optimal clustering for each possible  $k$  and we then "score" (using the silhouette coefficient) each result. The score takes into consideration the distance between elements within a cluster as well as the distance across clusters, while also penalizing the existence of singletons.<sup>17</sup> The optimal number of clusters implied by the silhouette coefficient is  $k = 164$ . Table E in the Appendix reports the subclasses assigned to each cluster.<sup>18</sup>

---

<sup>13</sup>Note that we do not distinguish whether IPC subclasses were assigned to different patents or to the same patent conditional on being from the same inventor.

<sup>14</sup>Eight IPC subclasses whose second level is 99 (i.e., "Subject Matter not otherwise Provided for in this Section"), were excluded from the analysis since they are assigned to patents with no clear identified technology.

<sup>15</sup>The diagonal elements of the matrix,  $i = j$ , are set to be equal to one. Note that the so-obtained matrix does not need to be symmetric. For example, according to the matrix manufacture of dairy products (A01J) is closest to dairy product treatment (A23C), while dairy product treatment is closest to foods, foodstuffs, or non-alcoholic beverages (A23L)

<sup>16</sup>Other procedures for bundling patent classes have been proposed in the literature. One strand of the measures uses patent citation information (e.g., Zitt et al., 2000; von Wartburg et al., 2005; Leydesdorff and Vaughan, 2006; Leydesdorff et al., 2014). We also conduct such grouping as a robustness check and find substantial overlap. Another strand of measures uses the "co-classification" information of patents (Jaffe, 1986; Engelsman and van Raan, 1994; Breschi et al., 2003; Leydesdorff, 2008; Kogler et al., 2013; Altuntas et al., 2015). Others used likelihood of diversification as measures of distance (Hidalgo et al., 2007) and analysis of patent texts (Fu et al., 2012; Nakamura et al., 2015)

<sup>17</sup>More details on the procedure used to construct fields of knowledge can be found in Appendix A.4.

<sup>18</sup>As a robustness check, we also construct the proximity matrix based on the citation linkages, instead, and apply the same procedure. The results are similar to the ones obtained with our proximity matrix: (i) the percentage of pairwise IPC subclasses that are in the same cluster is 50.6 (excluding singleton clusters, which accounts for 22.6 percent of all clusters); (ii) the percentage of pairwise IPC subclasses that are in the



### 3 Some Stylized Facts on World Innovation

We start our empirical analysis by presenting some stylized facts about the evolution of innovation and knowledge spillovers across time and space. Throughout the rest of the paper, we will use the fields of knowledge created in Section 2.2 as our main unit of analysis.

#### 3.1 Evolution of Fields of Knowledge across Space and Time

We first document the evolution of the major fields of knowledge for the last hundred years and highlight how different countries contributed to their growth at different points of time. To measure the importance of each field of knowledge at any point in time, we compute the share of patents belonging to that field of knowledge. Each patent is weighted by the total number of forward citations.<sup>19</sup> We split our dataset into nineteen 5-years periods from 1920 to 2015, plus a period prior to 1920 where we lump together all the patents filed before that year. For each time period, we rank every field of knowledge based on its relative contribution to the overall patent activity.

Figure 1 shows the evolution of the fields that were ever present in the top five fields at any point in time according to our measure. Two trends are readily noticeable. First, we observe a substantial increase in the concentration of innovation around the 1990s – approximately 10% of the fields of knowledge account for 60% percent of all patent activity in the 2000s compared to 30% in the first half of the 20th century. Second, there is substantial heterogeneity in the evolution of the fields of knowledge over time. At the beginning of the twentieth century, fields of knowledge belonging to Mechanical Engineering and Transportation (packaging & containers; geothermal systems) are the most prominent fields. Starting in the 1950s, we observe a shift towards chemistry and physics (e.g., macromolecular compounds). Around the 1980s there was substantial increase in medical and veterinary science (e.g., diagnosis and surgery; medical preparation). Finally, and as expected, around the mid 1990s the fields of knowledge related to computing and communication techniques started playing the leading role in the innovation landscape.

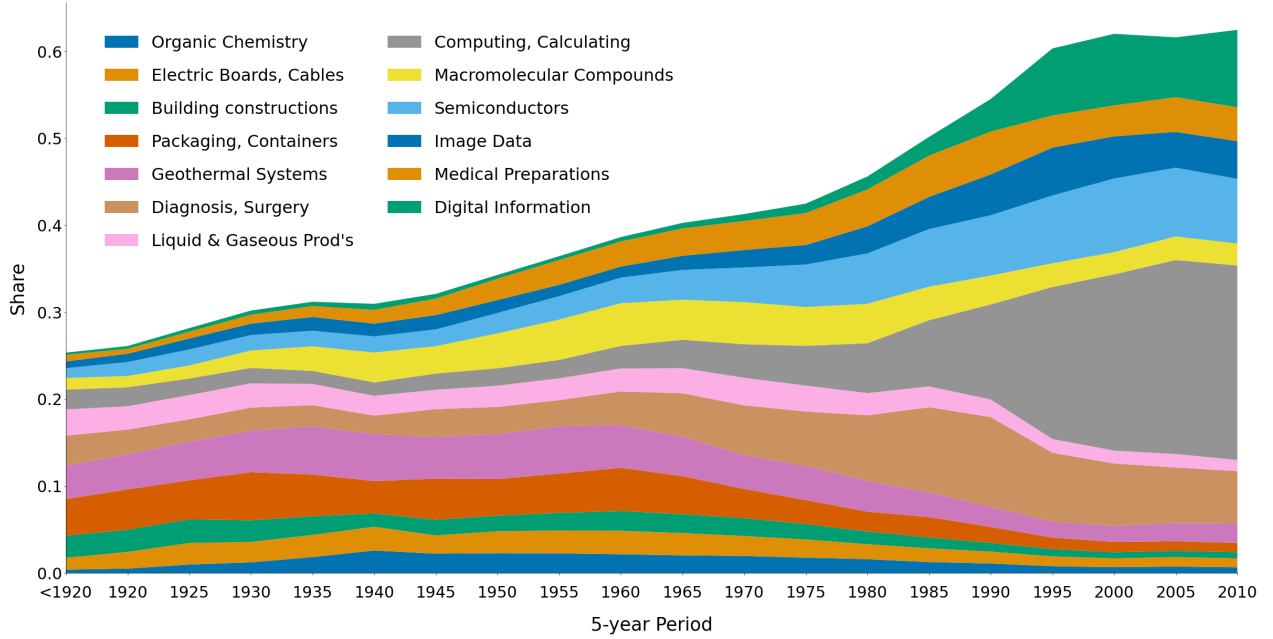
We also perform the same exercise using alternative measures of importance that address possible concerns related to, for example, heterogeneous patenting practices across countries or strategic patenting behavior that gained more prominence in the past few decades. To

---

same cluster weighted by importance, measured by the number of patents in the respective subclass relative to all patents, in the sample is 51.9 (excluding singletons); (iii) the percentage of clusters' centers that are the same is 67.1.

<sup>19</sup>As a reminder, we are using only the first patent of the family. If a patent belongs to multiple fields, we add a fraction of the patent to each field proportional to the number of IPC subclasses reported on the patents.

Figure 1: Evolution of Top Fields of Knowledge



*Notes:* This figure represents the share of each field of knowledge, measured by the number of first in the family patents weighted by received citations, in total patent activity across all fields in a given period of time. The width of the line reflects the share of knowledge field. Exact values for shares can be found in Table B.1.

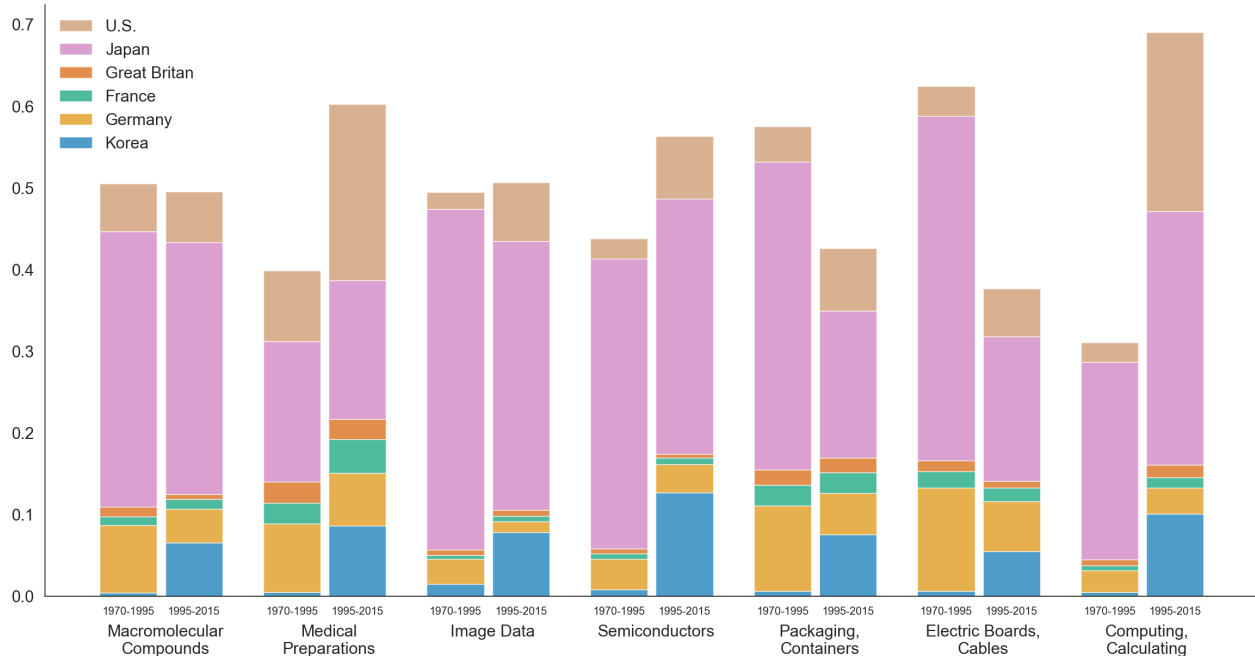
do this, we build importance measures that take into consideration country fixed effects, or measures that are only based on patents that were cited at least once. Table B.2 shows that these measures are highly correlated to our baseline.

Next, we turn to the spatial heterogeneity of innovation activities by studying the contribution of different countries to the growth of top fields of knowledge. We divide the sample into four periods: 1920-1945, 1945-1970, 1970-1995, and 1995-2015. We take seven fields of knowledge that took the leading role based on the number of patents throughout the entire period of study. Similarly to what we did in Figure 1, we assess the contribution of each country by computing its share in a certain field of knowledge.<sup>20</sup>

For the period 1920-1970, our sample is limited to six countries: the US, Great Britain, Germany, Switzerland, France, and the USSR. Figure B.1 shows that during this time period, the leading innovating role in major fields of knowledge was split between the US and Germany, followed by the UK and France. In fact, Germany overtook the US in every leading field in the period between the end of WWII and 1970.

<sup>20</sup>In this part of the analysis, we use the total number of patents without weighting by the number of citations for better comparability across countries. Different countries may use different procedures to assign citations, which would bias our results.

Figure 2: Countries Shares in Top Fields, 1970-2015



In Figure 2, we consider the whole sample in the years after 1970. Between 1970 and 1995, there are three clear technological leaders: Japan, the US, and Germany. The preponderant role played by Japan in the major fields of knowledge is remarkable. After 1995 other Asian countries, such as Korea, start rising to the forefront of the technological frontier. In this period, France experience a decrease of importance in the innovation landscape. Asian countries dominate in the fields related to computing, engineering, and digital information, while their role in chemistry and medicine is less pronounced.

We can also extend our analysis beyond the chosen fields of knowledge and compute an overall ranking by averaging the country ranking across all fields of knowledge. This exercise paints a picture similar to the one in Figure 2. Japan and the US are the technological leaders from 1970 until 1995, with Japan falling behind after the 2000s. The Soviet Union has an average ranking very similar to the US in 1970 but it falls behind subsequently, while Asian countries such as Taiwan gain prominence after the 2000s.<sup>21</sup>

<sup>21</sup>See Section B in the Appendix for further discussion. In the Appendix, we report two additional results that shed more light on the spatial heterogeneity of innovative activities over time. First, we decompose inequality in innovation within and between countries, and find that the inequality in patenting across countries has increased since the 2000s, while the within component has remained mostly stable. Second, we use a gravity-type regression to estimate the relationship between GDP per capita, geographical distance, and production of technologies. We find that changes in patenting shares across fields of knowledge are correlated across countries that are geographically and linguistically close to each other.

### 3.2 Using Citations to Measure Spillovers across Time and Space

So far, we have shown that there is substantial time variation in terms of composition of the technological output and in terms of geographical contribution to worldwide innovation. We now turn our attention to knowledge spillovers. We measure spillovers through patent citations across fields of knowledge and countries. There is an abundant literature studying within country spillovers using patent citations (e.g., [Jaffe et al., 1993](#), [Murata et al., 2014](#) for the United States), but the evidence on cross-country knowledge spillovers is more scarce. Despite being an imperfect measure of knowledge spillovers, patent citations provide a useful quantifiable benchmark that can be easily measured and used in our empirical analysis.

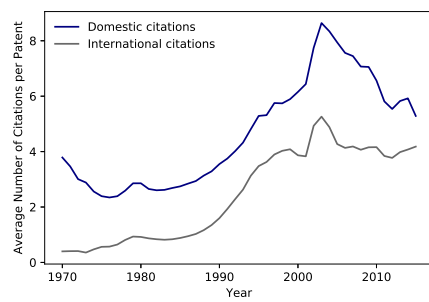
To illustrate this fact, we focus on the post 1970 sample, for which we have data for virtually all countries in the world. We compute backward citations given to patents filed after 1900. Panel (a) in [Figure 3](#) shows the evolution of the average number of citations given by patents filed after 1970. The average number of citations experiences an important increase starting around the 1980s. Domestic citations keep increasing up until 2002 and they then show a marked decline, whereas international citations plateau at about 4 international citations per patent in the late 1990s. A closer look at panel (a) further reveals that domestic citations tend to be more prominent than citations given to international patents: domestic patents are cited at a rate that is roughly double the one for international patents. Panel (b) breaks down these trend by additionally looking at citations within and outside the field of knowledge (FoK) of the citing patent.<sup>22</sup> The plot shows that citations are concentrated not only geographically but also technologically. Moreover, this gap has widened over the past decades.

As discussed in the previous section, an important pattern is that most knowledge (as measured by patent filings) is produced by a handful of countries, what we refer to as the “technological leaders.” Specifically, for the period 1970-2015 two countries – Japan and the United States – are responsible for the largest share of patents produced worldwide. Panels (c) and (d) of [Figure 3](#) separately depicts citation dynamics for Japan and the US and the rest of the world. While we observe an increase in the average number of citations per patent, there are two important differences between the two panels. First, the United States and Japan, on average, give more citations per patent than the rest of the world. Second, most of the citations in the US and Japan are given to domestic patents, while the rest of world mostly relies on knowledge produced in other countries.<sup>23</sup>

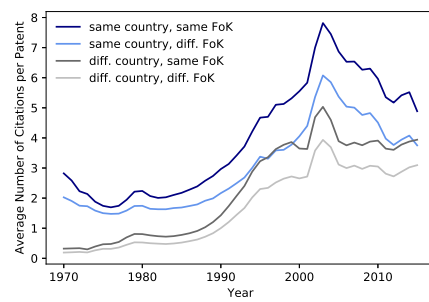
---

<sup>22</sup>The sum of the four lines in panel (b) is not equal to the total number of backward citations since there is some double-counting due to the fact that cited patents belong to multiple fields of knowledge and (more rarely) to multiple countries.

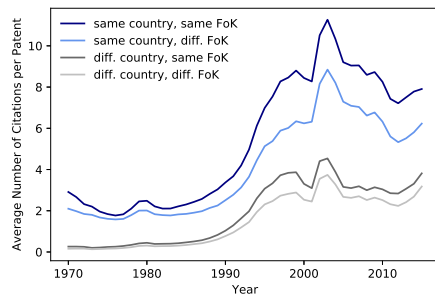
<sup>23</sup>Decomposition of citations for other countries, namely Germany, France and Great Britain, are reported in [Figure B.2](#).



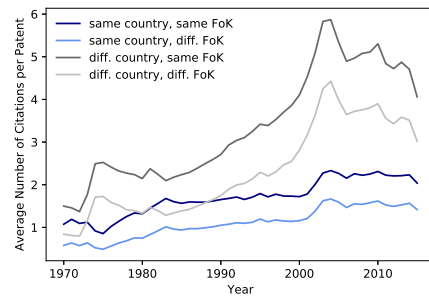
(a) Domestic vs International



(b) Citation Decomposition



(c) US and Japan



(d) Non-US and Non-Japan

Figure 3: Citation Dynamics, 1970-2015

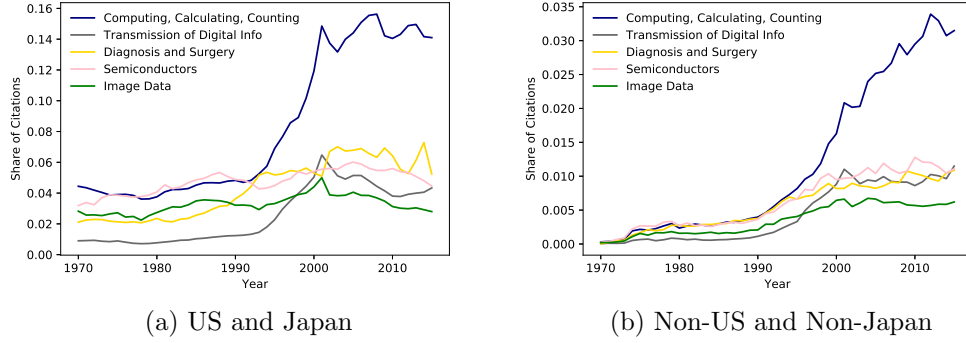


Figure 4: Share of citations going to the US and Japanese patents by FoK, 1970-2015. Each line in the plots represents the share of citations going to the US and Japan patents that belong to a given field of knowledge. Panel (a) depicts domestic citations given by American and Japanese patents, and panel (b) depicts international citation to the patents filed in the US and Japan given by other countries.

Figure 3 depicts a rapid increase in the overall average number of citations per patent. To better understand what lies behind this increase, we concentrate our attention to the citations received by five fields of knowledge that have become the leading technology fields over the past decades. Figure 4 shows that the substantial increase in the number of citations observed in Figure 3 is mainly driven by two fields of knowledge: “Computing, Calculating, Counting” and “Transmission of Digital Information.” What is perhaps even more striking is the fact that most citations to this field of knowledge are given to US and Japanese patents.

Taken together, the evidence presented in this section shows that knowledge spillovers are an important component of the innovation process. Although spillovers that originate from the same country and field of knowledge are still the most relevant, international knowledge spillovers have been steadily gaining importance over the last decades. This increase is visible when considering spillovers coming both from the same field of knowledge and from other fields of knowledge, and is mainly driven by a dramatic increase in the citations received by American and Japanese patents, especially in the fields of knowledge related to computing, information processing and medicine.

## 4 Conceptual Framework

In this section, we present a framework that incorporates the elements of our data analysis in the previous sections and that serves as a guide for empirical exercises. Our framework builds on the canonical growth literature and it applies it to a multi-sector growth model. The fundamental element of our analysis is the production function of ideas and its link to patenting activity. We choose our formulation of the idea production function to remain

relatively parsimonious so that it encompasses alternative formulations of endogenous growth theory (see, e.g., [Jones, 1999](#) for a discussion).<sup>24</sup> Our framework only specifies the production-side of the economy, and it does not assume the existence of a constant growth path of output or productivity at the sectoral level of analysis. This is indeed empirically the case for the United States and other advanced economies for which we have long, sectoral-level time series of output and productivity ([Comin et al., 2019](#)).

Consider a world economy with  $C$  countries,  $S$  sectors and  $K$  fields of knowledge, where we index countries by  $c$ , sectors by  $s$ , fields of knowledge by  $k$ , and time by  $t$ . There is a representative firm in each country-sector that produces sectoral output combining physical inputs (labor, capital, etc.) according to the best production methods used in that country-sector at time  $t$ , which are summarized by sectoral TFP,  $TFP_{sct}$ . Following the endogenous growth literature, we refer to these best production methods as best ideas—thus assuming that the role of ideas is to increase firms’ productivity by developing and improving methods of production (see, e.g., [Acemoglu, 2009a](#)).

We denote by  $N_{cskt}$  the stock of ideas available in country  $c$ , sector  $s$ , field of knowledge  $k$ , and time  $t$ . The state of world ideas at time  $t$  is thus summarized by the vector  $\mathbf{N}_t \equiv (N_{111t}, \dots, N_{cskt}, \dots, N_{CSKt})$ . There is a production function for new ideas,  $I(\cdot)$ , that establishes the relationship between the flow of new ideas in a given field of knowledge and production sector,  $\Delta N_{cskt}$ , the current stock of knowledge,  $\mathbf{N}_t$ , and inputs devoted to generate new ideas,  $R_{cskt}$ ,

$$\Delta N_{cskt} = I(S_{csk}(\mathbf{N}_t), R_{cskt}), \quad (1)$$

where  $\Delta$  denotes the time difference operator between  $t + 1$  and  $t$ . The spillover function  $S_{csk}(\mathbf{N}_t)$  captures how the current world stock of knowledge  $\mathbf{N}_t$  helps generate new ideas in country  $c$  in field of knowledge  $k$  for sector  $s$ . We take the spillover function to be

$$S_{csk}(\mathbf{N}_t) = \sum_{c' \in C} \sum_{s' \in S} \sum_{k' \in K} \alpha_{c's'k't} N_{c's'k't}, \quad (2)$$

where  $\alpha_{c's'k't}$  captures the reliance of the production function of ideas in  $csk$  on ideas from  $c's'k'$  at time  $t$ . Note that we purposefully state Equation (1) generically so that it subsumes the first generation of endogenous growth models as in [Romer \(1990\)](#) or [Aghion and Howitt \(1992\)](#), semi-endogenous growth as in [Jones \(1995\)](#), [Kortum \(1997\)](#) or [Segerstrom \(1998\)](#), or second generation as [Aghion and Howitt \(1998\)](#), [Young \(1998\)](#) or [Peretto \(1998\)](#).<sup>25</sup>

<sup>24</sup>Our formulation builds on previous studies that have been applied to the study of the patent network of citations ([Acemoglu et al., 2016](#)). Relative to [Acemoglu et al. \(2016\)](#), we present additional model elements to relate our results to TFP and output per capita and also extend the model to a multi-country setting.

<sup>25</sup>For example, one specification extensively used in the literature (e.g., [Romer, 1990](#); [Jones, 1995](#)) ignores cross-country spillovers, and corresponds to having  $S = K = 1$  and  $S_c(\mathbf{N}_t) = N_{ct}$  and postulates a log-linear



Since ideas are to a large extent non-rival (Romer, 1990), the vast majority of these theories resort to intellectual protection in the form of patents to ensure that investments in new ideas can be recovered with future profits.<sup>26</sup> This observation motivates our empirical strategy to proxy the generation of new ideas through patent filings. Patents provide a quantifiable measure over time and space that is arguably very hard (or impossible!) to replicate with other measures of ideas or innovation. Moreover, through citations, patents also provide an empirical measure of reliance on existing ideas across space and fields of knowledge. We rely on these spillover measures in our empirical analysis and, in particular, on our instrumental variables strategy. In practice, however, not all ideas are patented, and not all ideas a patent builds on are cited. We thus think of patents as a *proxy* for new ideas,  $\Delta N_{cskt}$  and citations as a *proxy* for spillovers. We discuss in the next section how our empirical specification addresses these potential discrepancies between idea generation and patenting.

Regardless of their vintage, endogenous growth theories argue that there is a positive, monotonic relationship between the ideas produced and sectoral TFP growth. However, they differ on the implied effect of the current stock of ideas on the generation of new ideas: first-generation theories emphasize the standing on the shoulders of giants effect, while semi-endogenous theories allow for fishing-out effects. To build a connection with our empirical specification, we assume a flexible, iso-elastic relationship between ideas and TFP growth

$$\log TFP_{cst+1} = \beta_0 + \beta_A \log TFP_{cst} + \beta_N \log(1 + \Delta N_{cst}), \quad (3)$$

with  $\beta_0, \beta_A, \beta_N \geq 0$  and  $\Delta N_{cst} = \sum_{k=1}^K \Delta N_{cskt}$  denoting the total number of ideas generated in country  $c$  sector  $s$  at time  $t$  across all fields of knowledge.

Equation (3) nests a number of cases often considered in the literature and it is at the core of our empirical specifications in the next section. For example  $\beta_0 = 0$  and  $\beta_A = \beta_N = 1$  generates building-on-the-shoulders-of-giants dynamics, whereby the growth rate of  $TFP_{cst}$  is directly controlled by the number of ideas produced at time  $t$ . In this case, if no ideas are produced at time  $t$ ,  $\Delta N_{cst} = 0$ , there is no TFP growth. Letting  $\beta_A < 1$  introduces the fishing-out-of-the-same-pond effect in the sense that more ideas become necessary over time to sustain constant TFP growth.

Finally, we are ultimately interested on the effect of innovation on output per worker—which proxies for the level of development, besides of being an indirect proxy for productivity. We thus extend our framework to study output per worker. Suppose that sectoral value added per worker,  $y_{cst}$ , is given by a Cobb-Douglas production function,  $\log y_{cst} = \log TFP_{cst} +$

---

relationship,  $I = N_{ct}^\phi R_{ct}$  with  $\phi \leq 1$ .

<sup>26</sup>See, among others, Aghion and Howitt (1998), Acemoglu (2009b) and the references therein.

$\alpha \log k_{cst}$ , where  $k_{cst}$  denotes capital per worker and  $0 < \alpha < 1$ . We begin making the following observation. Under the assumption of competitive markets, firm optimization implies that the ratio of sectoral output per worker between two sectors,  $s$  and  $s'$ , is proportional to their TFPs,

$$\log y_{sct} - \log y_{s'ct} = \log TFP_{sct} - \log TFP_{s'ct}. \quad (4)$$

Equation (4) implies that the differential growth rate in output per worker across sectors coincides with the differential growth rate in sectoral TFPs. If we allow  $\alpha$  to be sector specific, we have that the difference in output per worker growth rates has an additional term that depends on factor prices weighted by factor share differences which can be absorbed using a country-time fixed effect.<sup>27</sup> This result allows us to build a connection between empirical specifications in which the dependent variables are value added per worker and TFP, and explain why we find very similar estimated effects.

Our baseline empirical specification corresponds to a standard log-linear growth regression specification to which we explicitly incorporate the production function of new ideas,<sup>28</sup>

$$\begin{aligned} \log y_{cst+1} &= \log y_{cst} + \Delta \log TFP_{cst} + \beta (\log y_{ct} - \log TFP_{cst}) + \theta \log(1 + \Delta N_{cst}) + \delta_{cs} \\ &= \phi_N \log(1 + \Delta N_{cst}) + \phi_Y \log y_{cst} + \phi_K \log k_{cst} + \delta_{cs}, \end{aligned} \quad (5)$$

where  $\delta_{cs}$  is a country-sector specific intercept that absorbs the normalized asymptotic output per worker of the sector. We have used Equation (3) to go from the first to the second line. The noteworthy feature of Equation (5) relative to Equation (3) is that the level of output per worker also appears on the right-hand-side. This term controls for convergence effects and its analysis has been the focus of empirical growth theories in the last decades. By contrast, however, the focus of our analysis will be on the elasticity of patenting on output growth,  $\phi_N$ , rather than the convergence term  $\phi_Y$ .

## 5 Empirical Analysis

This section presents the main empirical exercises of the paper to study the effect of innovation on productivity. We begin analyzing the effect of innovation on sectoral output per worker and TFP using cross-country panel data. We present our identification strategy in Section 5.1 and report our baseline results in Section 5.2. We finalize the section presenting

---

<sup>27</sup>Letting  $R_{ct}$  denote the price of capital, we would have the term  $(\alpha_s - \alpha_{s'}) \log R_{ct}$  appearing in addition to  $\log TFP_{sct} - \log TFP_{s'ct}$  in Equation (4).

<sup>28</sup>This equation is obtained by obtained by log-linearizing a Solow model. See Barro (1991); Barro and Sala-i Martin (1992); Barro et al. (2004); Acemoglu (2009a); Durlauf et al. (2005) for a detailed derivation and further discussion.

two extensions. In Section 5.3, we extend our baseline estimation to longer time horizons where the dependent variable is output per capita starting in 1980 (thus loosing sectoral variation).

## 5.1 Estimating Equations and Identification Strategy

Our empirical model is based on Equation (5) from our analytical framework. The specification of our baseline regression model is the following:

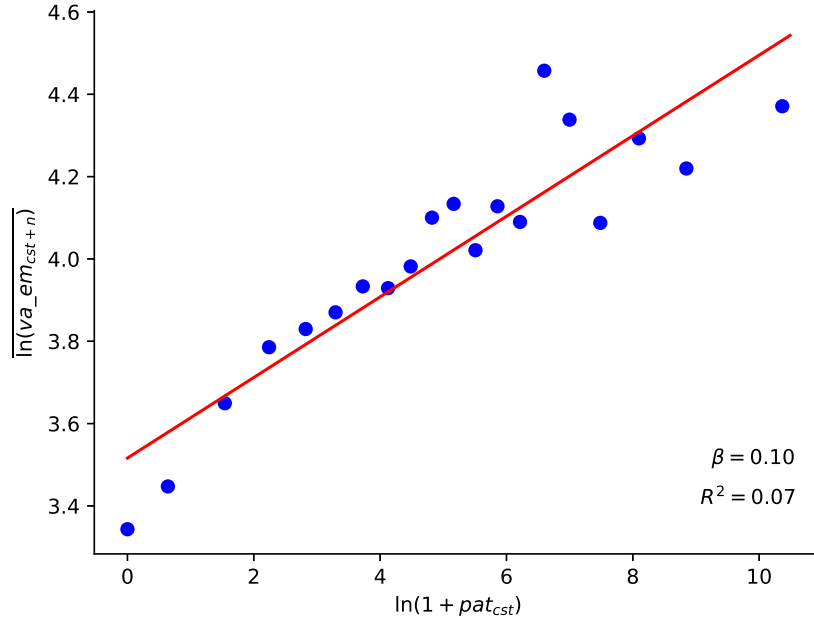
$$\ln(y_{cst+n}) = \phi_A \ln(y_{cst}) + \phi_N \ln(1 + pat_{cst}) + \phi_0 X_{cst} + \delta_{ct} + \delta_{st} + \epsilon_{cst}, \quad (6)$$

where  $\ln(y_{cst+n})$  is future annual output per worker in the period  $t + n$ ,  $X_{cst}$  denotes a set of controls for country  $c$  and sector  $s$ ,  $\delta_{ct}$  and  $\delta_{st}$  denote country-time and sector-time fixed effects and  $\epsilon_{cst}$  denotes the error term. Thus, relative to the model presented in the analytical framework, there are two departures. First, rather than looking one period ahead, we look at output per worker  $n$  years ahead of period  $t$  (we take average annual output per worker for  $n \in \{1, \dots, 3\}$  as our baseline, and show that the results are robust to  $n \in \{1, \dots, 3\}$ ), as it is common in the growth literature (e.g., Arcand et al., 2015), to smooth out short-term fluctuations in the variable of interest and concentrate on long-run trends. Second, we use country $\times$ time and sector $\times$ time fixed effects to allow flexible differential trends across countries and sectors.

Our main results use the output per worker and TFP measures derived from the World Input Output Database. The data used in our baseline analysis span from year 2000 through 2014, covers 36 countries and considers 20 sectors. As we discuss below in more detail, we use the 1970-2000 patent data to construct our instrument. Figure 5 shows the binscatter plot of the raw correlation between patent activity  $\ln(1 + pat_{cst})$  and productivity, measured by value added per employment,  $\overline{\ln(va\_em_{cst+n})}$  over our sample period. In the cross-section of countries and sectors, a one percent increase in the number of patents is associated with a 0.10 percent increase in future output per worker averaged over the next three years. The coefficient is statistically significant (we cluster the standard errors two-way at a country and sector level).

Equation (6) is our baseline model to study the effects of innovation on productivity. The coefficient of interest is  $\phi_N$ . It relates changes in number of patents at the country-sector level in a given year to changes in output per worker in the following years. The inclusion of sector-year dummies accounts for the fact that different industries rely differently on innovations, and that this relationship can vary over time. In addition, sector-year dummies allow us to control for the presence of technological waves and other sectoral shocks that are common

Figure 5: Unconditional Correlation between Value-added per worker and Number of Patents



across all countries. The inclusion of country-year fixed effects accounts, first, for the fact that different countries have different propensities to innovate, and, second, for any business cycles fluctuations at a country level, e.g., a financial crisis.

To evaluate the strength of the causal relationship between innovation and productivity, we need to identify variation in patent activity that is orthogonal to unobserved factors that might affect both innovation activity and productivity at the same time. There is a wide range of such possible factors and the direction of the bias is ex-ante ambiguous. An example of such factors is technological obsolescence of some industries. Reverse causality is also a concern, with higher productivity being the cause, rather than consequence, of higher innovation activity in a given sector. Finally, estimates might be suffering from attenuation bias, due to presence of measurement error since patents are an imperfect measure of ideas and innovation.

### 5.1.1 Instrument Construction

To deal with these threats to identification, we build an instrumental variable for the number of patents. Our instrument is based on a shift-share design that leverages pre-existing cross-country, cross-sector variation to predict the current level of patenting. More specifically, we exploit the pre-determined network of patent citations during the period 1970-90 to identify knowledge links to construct the “shares” part of our instrument. We then construct the

“shifts” for the period 1990-2014 using a mix of observed and predicted number of patents in other countries and sectors starting from the year 1980 on a rolling basis.<sup>29</sup> Interacting the shares with the shifts and adding those up, we obtain the “predicted” number of patents in the period 2000-2014 as our shift-share instrument. Importantly, we only use “predicted” patents as shifts to generate the instrument for our baseline sample. Thus, our instrument predicts patenting activity in the current period based on the knowledge spillovers from other countries and sectors. In this sense, our shift-share design can be interpreted as a particular application of the linear knowledge spillover function presented in Equation (2) in Section 4.

Before delving into the details of the instrument, it is worth emphasizing that our proposed shift-share design differs from a more standard “Bartik” design. The reason is that we exploit the directed network of citations to construct linkages across country-sector pairs and then use shift terms that also vary at the country-sector level. In contrast, a standard “Bartik” only uses as sources of variation the own country-sector exposure (shares) and the world patenting activity in a sector (shift). For our purposes, the standard Bartik design is unappealing since it may confound innovation shocks with world industry or technological trends that also affect productivity.<sup>30</sup>

To compute the “share” terms of our instrument, we gather patent information on the country of origin, technological field, backward and forward citations, and the sequence of the patent within its family (as described in Section 2) for all patents filed from  $T_0^{share} = 1970$  to  $T_1^{share} = 1990$ . We use a correspondence from technological fields to industry codes to assign each patent to one or multiple sectors, with their respective weights in the latter case.<sup>31</sup> The underlying idea is to measure knowledge flows across countries and sectors through the share of citations that each patent produced in the country and sector of origin  $o$  gives to patents in the destination country and sector  $d$ . In particular, for each patent of sector  $s_o$  belonging to country  $c_o$  at time  $t$ , we calculate the share of citations given to patents produced in sector  $s_d$ , country  $c_d$  at time  $t - \Delta$  for some citation lag  $\Delta > 0$ . We repeat this procedure for each time period  $t$  between  $T_0^{share}$  and  $T_1^{share}$  and sum these shares to obtain the total number of citations over the  $T_1^{share}$  to  $T_0^{share}$  period. Importantly, to control for size effects due to the fact that some locations and/or sectors tend to patent more for idiosyncratic reasons, we normalize this measure by the total number of patents produced in the country-sector of the

---

<sup>29</sup>We start by computing “predicted” number of patents in year 1990 by using actual number of patents filed during the period 1980-1989 as shifts. For the year 1991, we use the actual number of patents filed during the period 1981-1989 and the predicted number of patents in 1990 computed in the previous step as shifts to generate the predicted number of patents. Starting from the year 2000, only the predicted number of patents are used as shifts to generate the instrument.

<sup>30</sup>Consider, for example, a world where a few countries leaders determine in which sectors most of innovation activity is going to happen. In this case, the shift components that we would use in the construction of the instrument would not be orthogonal neither to patent activity nor to productivity.

<sup>31</sup>We use Eurostat correspondence tables, Van Looy et al. (2014).

destination country,  $d$ .

Formally, the adjacency matrix of the knowledge network for a citation lag  $\Delta$  is given by:

$$m_{c_o, c_d, s_o, s_d, \Delta} = \frac{\sum_{t=T_0^{share}}^{T_1^{share}} \sum_{p \in \mathcal{P}(c_o, s_o, t)} s_{p \rightarrow (c_d, s_d, t-\Delta)}}{\sum_{t=T_0^{share}}^{T_1^{share}} |\mathcal{P}(c_d, s_d, t-\Delta)|}, \quad (7)$$

where  $s_{p \rightarrow (c_d, s_d, t-\Delta)}$  denotes the share of citations that patent  $p$  gives to patents of sector  $s_d$  produced in country  $c_d$  filed at time  $t-\Delta$ ,  $\mathcal{P}(s_o, c_o, t)$  denotes the set of patents in  $s_o, c_o$  at time  $t$ , and  $|\mathcal{P}(\cdot)|$  denotes the total number of patents in the set (i.e., set cardinality). As the numerator shows, we add the citations of all patents originating in country-sector  $(c_o, s_o)$  at time  $t$  over the time period from  $T_0^{share}$  through  $T_1^{share}$  going to patents filed in country-sector  $(c_d, s_d)$  at time  $t-\Delta$ , and normalize by the patent count in the destination country-sector at time  $t-\Delta$ . As we explain below, we use resulting object  $m_{c_o, c_d, s_o, s_d, \Delta}$  to construct the “shares” in our shift-share instrument.<sup>32</sup> Note that the “share” terms  $m_{c_o, c_d, s_o, s_d, \Delta}$  do not add up to one, since their levels capture the number of citations that are typically received by patents filed in  $(c_d, s_d)$  from  $(c_o, s_o)$  with a lag  $\Delta$ .

Our network analysis also takes into account that the speed at which ideas diffuse might differ across locations and sectors. We formally capture this effect by allowing the weights in our network to be time specific. We compute the citation shares at different time horizons, with citations lags  $\Delta \in \{1, \dots, 10\}$ . In other words, we allow for the strength of the links to depend on how many years have passed between the time cited and citing patents were filed. Thus, our share terms are allowed to vary by country-sector citing-cited pairs, and by time lag between cited and citing patents.

Finally, we describe our “shift” terms and the construction of our instrument. Our shift-share design is based on the idea of predicting the number of patents in a country and sector of interest based on predicted knowledge spillovers, i.e. as if only the pre-existing knowledge network mattered for generating knowledge. Intuitively, this approach mirrors the one of an input-output model except that recognizes the non-rival nature of ideas (an idea in one country-sector can potentially spillover to multiply country-sector pairs). To this end, we

---

<sup>32</sup>Let us reiterate here that, as we have done in Section 2, we restrict our sample to patents that are the first in their family to avoid double-counting of the same idea and capture only knowledge creation originated in a particular country-sector. However, for cited patents, we count all cited patents irrespective of whether they are the first or not in their family to capture all innovations on which any given patent builds on. We also note that [Berkes and Gaetani \(2018b\)](#) show that the network of patents in the United States is stable in the time frame they consider, which roughly coincides with ours.

then use as shift terms patents filed  $\Delta$  years before the period of interest  $t$  in other countries and sectors (or predicted patents as we explain below), and use the strength of the linkages to predict the number of patents in the country-sector of interest. In particular, we assume that the strength of knowledge spillovers between country-sector dyads is mediated through how ideas in other country-sectors (as measured by our shift terms) diffuse through the knowledge network (as measured by the linkages  $m_{c_o, c_d, s_o, s_d, \Delta}$ ). By interacting the shift and share terms and summing across countries, sectors and diffusion lags, we then obtain a predicted number of patents  $\widehat{pat}_{c_o, s_o, t}$  in country  $c_o$ , sector  $s_o$  and time  $t$ .

Formally, our baseline shift-share design is constructed iteratively as follows. For 1990, we obtain predicted patents as

$$\widehat{pat}_{c_o, s_o, 1990} = a_{1990} \sum_{s_d \in \mathcal{S} \setminus s_o} \sum_{c_d \in \mathcal{N} \setminus c_o} \sum_{\Delta=1}^{10} m_{c_o, c_d, s_o, s_d, \Delta} \cdot pat_{c_d, s_d, 1990-\Delta},$$

where  $a_t$  is a re-scaling term that ensures that predicted number of patents is equal to the actual number of patents in period  $t$  worldwide and  $pat_{c_d, s_d, 1990-\Delta}$  is the actual number of patents filed in  $c_d, s_d, 1990-\Delta$ .<sup>33</sup> Between 1991 and 1999 we construct the predicted number of patents using previously computed predicted number of patents for years since 1990 and observed patenting activity prior to 1990. That is, for  $t \in (1990, 2000)$  we have that

$$\widehat{pat}_{c_o, s_o, t} = a_t \sum_{s_d \in \mathcal{S} \setminus s_o} \sum_{c_d \in \mathcal{N} \setminus c_o} \left( \sum_{\Delta=1}^{t-1990} m_{c_o, c_d, s_o, s_d, \Delta} \cdot \widehat{pat}_{c_d, s_d, t-\Delta} + \sum_{\Delta=t-1990}^{10} m_{c_o, c_d, s_o, s_d, \Delta} \cdot pat_{c_d, s_d, t-\Delta} \right),$$

where  $\widehat{pat}_{c_o, s_o, t}$  denotes predicted patenting. Finally, starting in year 2000 we construct predicted patenting off the predicted patenting computed in the decade of the 1990s:

$$\widehat{pat}_{c_o, s_o, t} = a_t \sum_{s_d \in \mathcal{S} \setminus s_o} \sum_{c_d \in \mathcal{N} \setminus c_o} \sum_{\Delta=1}^{10} m_{c_o, c_d, s_o, s_d, \Delta} \cdot \widehat{pat}_{c_d, s_d, t-\Delta}$$

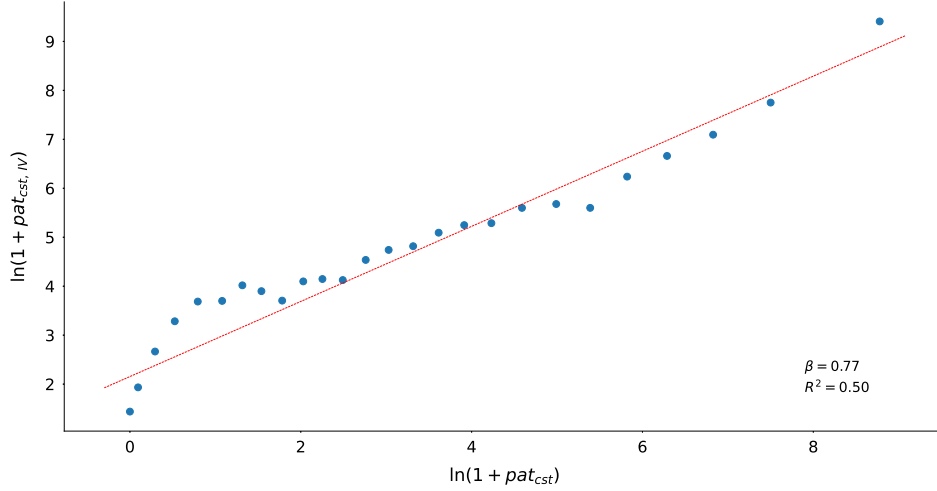
Note that, to mitigate endogeneity concerns, the proposed shift-share design avoids using contemporaneous shares and shifts. First, to construct the share terms, we use the pre-sample period 1970-1990 to construct the knowledge network off patent citations which. Second, when constructing the shift terms, we diffuse the observed patents filed pre-1990 over the period 1990-1999 to predict the patenting activity in the 1990s. We then use this predicted patenting activity to predict patenting activity over the sample period (2000-2014). Third, we discard citations coming from the same country and from the same sector when we construct

---

<sup>33</sup>Figure C.1 in the appendix represents a simple example of described procedure



Figure 6: Unconditional Correlation between Actual and Predicted Patents



predicted patents. This is as if when constructing the knowledge network  $m_{c_o, c_d, s_o, s_d, \Delta}$  in Equation (7), we eliminated the own-country and own-sector terms

$$m_{c_o, c_d, s_o, s_d, \Delta} = \begin{cases} 0 & c_o = c_d \\ 0 & s_o = s_d \end{cases}$$

We do this to avoid endogeneity concerns arising from the fact that edges that link the same country or sector might be correlated with future shocks (despite being at least 10 years apart).

Figure 6 visually compares the actual and predicted number of patents by providing a binscatter plot. The two variables are strongly but not perfectly correlated: the coefficient of the regression is 0.77 and  $R^2 = 0.50$ . The Cragg-Donald Wald F statistics in the benchmark regression is 2,070, which rules out weak instrument concerns.

Our proposed instrument belongs to the family of shift-share instruments: weighted averages of a common set of shocks, with weights reflecting heterogeneous shock exposure. The key difference of our shift-share design relative to standard “Bartik” type designs is that our shares leverage the entire information of the citation network structure rather than only using information on the country-sector of interest. Despite this, the analysis of the validity of our instrument still falls within the shift-share instrumental variable framework and it must rely on some assumptions about the exogeneity of the shift terms, exposure shares, or both (see [Borusyak et al. \(2018\)](#) and [Goldsmith-Pinkham et al. \(2020\)](#) for a technical discussion of those assumptions).

To provide evidence in support of our instrument, we test for a number of assumptions underlying the identification restrictions of shift-share designs, along the lines of [Tabellini](#)

(2020). First, the validity of the shift-share instrument rests on the assumption that countries and sectors giving more citations (to other sectors and countries) in the period between 1970 and 1990 are not on different trajectories for the evolution of output per worker in the period of analysis (2000-2014). We test this assumptions in two ways: i) regressing productivity in 1990 against average patent activity in the period of 2000-14 predicted by the instrument, ii) we check that results are unchanged when controlling for an average level of patent activity in the period 1970-90.<sup>34</sup>

Second, we rule out the possibility that the links of knowledge diffusion used to construct the instrument capture a demand pull factors from the destination country and sector, rather than a supply push from the origin country and sector. We do so by directly controlling by shift-share variable constructed analogously to our instrument but with the timing reversed, so that it predicts the number of patents that should have been produced in the past in other countries and sectors to generate the current level of patenting given lagged citation patterns. More precisely, we start by constructing the pre-determined network of citations, but now using forward citations instead of backward. Then, using the patenting activity across country-sector pairs during our sample period (2000-2014) and the forward citation network generated in the previous step, we infer the number of patents in the period 1970-1990 that would be necessary to rationalize the 2000-2014 period. Then, we include this predicted number of patents in our baseline regression as an additional control. In other words, these predicted patents are patents that should have been filed in the period of 1970-1990 to generate patent activity in the period 2000-2014 that we observe in the data.

## 5.2 Innovation and Productivity

Our identification strategy relies on the pre-determined network knowledge linkages that allows us to capture country and sector specific shocks to innovation activity, measured by a number of patents, due to knowledge created in other geographical and sectoral areas. In this section, we explore the effects of these shocks on productivity.

Table 1 shows our benchmark estimates of the relationship between value added per employment and innovation instrumented with predicted innovation. As we have discussed, we use a three year average of output per worker to remove any short-term business cycle fluctuations.<sup>35</sup> Also, our benchmark regression uses data from years 1970-90 to compute pre-determined network linkages, and the period of our analysis is 2000-2014. The first

<sup>34</sup>We use value added per employment obtained from UNIDO data as a measure of productivity.

<sup>35</sup>To retain zero-valued observations, we use in our baseline specification  $\ln(1 + pat)$ , but the results are robust to the inverse hyperbolic sine transformation used instead. Results for alternative log transformation of patents and forward lags for dependent variable are reported in Table C.1 in the Appendix.

Table 1: 2SLS Estimates: 2000-2014

	$\overline{\ln(\text{va\_em}_{c,s,t+n})}$		$n \in \{1, 2, 3\}$	
	(1)	(2)	(3)	(4)
$\ln(\text{patent}_{c,s,t})$	0.004 (0.003)	0.017 (0.008)	0.005 (0.003)	0.019 (0.007)
$\ln(\text{va\_em}_{c,s,t})$	0.942 (0.016)	0.937 (0.016)	0.934 (0.016)	0.928 (0.015)
$\ln(\text{capital}_{c,s,t})$	-0.016 (0.008)	-0.014 (0.008)	-0.015 (0.008)	-0.014 (0.008)
$\ln(\text{employment}_{c,s,t})$	0.020 (0.010)	0.015 (0.010)	0.012 (0.011)	0.006 (0.010)
$\ln(\text{imports}_{c,s,t})$			0.009 (0.009)	0.010 (0.009)
Country-Year FE	Y	Y	Y	Y
Sector-Year FE	Y	Y	Y	Y
# obs.	8,357	8,357	8,357	8,357
# countries	36	36	36	36
First-stage estimates				
Predicted		0.461		0.461
$\ln(1 + \text{patent}_t)$		(0.079)		(0.079)
CD Wald F		2,085		2,069

*Notes:* Period of the analysis is 2000-14 using pre-determined matrix based on the data from 1970-90. First-stage estimates include all the controls. Standard errors are two-way clustered at a country and sector levels in parentheses. Columns (1) and (3) report the results using OLS, and Columns (2) and (4) report the results obtained with 2SLS.

two columns report the estimated results when we only include capital and employment as controls, as well as country-year and sector-year fixed. In the third and fourth columns, we exploit the trade linkages given by the world input-output structure and add as controls the value of intermediates imported by country-sector pairs to explore the role of trade linkages.

The coefficient on innovation activity is positive, and statistically significant across the board. The coefficient obtained using instrumental variable approach is substantially larger than the one using OLS. The magnitude of the two-stage least squares regressions is also stable across specifications and it suggests that a 1% increase in patenting between 2000 and 2014 leads to 0.017% increase in output per worker. This estimated elasticity implies that 1 standard deviation increase in log patents generates an increase in log value added per employment of 4.7% (i.e., 4.7% growth of output per worker, or a 0.047 standard deviation

increase in logarithm of output per worker). One standard deviation increase corresponds to an increase in innovation activity in the pharmaceutical sector from the level of innovation observed in Canada to the level observed in the US in 2000. This also approximately corresponds to an increase in innovation activity in computer and electronic products sector from the level of innovation observed in Australia or France to the level observed in the US in 2000. Similarly, an interquartile shift in log of number of patents implies 0.055 interquartile shifts in logarithm value added per employment. Looking at countries at the bottom quartile of the patenting distribution in our sample, our estimated elasticity implies that, *ceteris paribus*, if Mexico in 2000 innovated in computer and electronic products and pharmaceuticals at the level of the US, output per worker in these sectors would have been higher by 12.0% and 11.3%, respectively.

Given the presence of fixed effects in our regression, it is important to interpret the coefficient as a change in output per worker caused by the growth rate of innovation activity that is beyond the average growth rate of innovations across the world in a given sector and beyond the average growth rate of innovation across all sectors in given country in a given period of time.<sup>36</sup>

The estimated 2SLS coefficients are larger than the ones obtained in the OLS regression. This increase is consistent with the likely scenario in which our OLS estimates suffer from attenuation bias because patents are an imperfect measure of innovation activity. Another possible explanation for the bias could be an increase in market concentration – a trend observed in most advanced countries since 2000. Higher market concentration leads to slow-down in productivity, while stimulates innovation activity due to the fact that leader(s) don't want to give up their leading role (Akcigit and Ates, 2021). Finally, we look whether our estimates are consistent if we use other measures of productivity, and more specifically if we use TFP growth instead of output per worker.<sup>37</sup> Table 2 shows our estimates for two measures of TFP growth, as well as value added per employment growth. The coefficient on innovation activity is positive, and statistically significant across different measures, and are consistent with baseline results.<sup>38</sup>

---

<sup>36</sup>Bringing those numbers to actual data means that 1 standard deviation in increase in log annual number of patents after partialing out all controls in column (2), i.e, country-year and sector-year fixed effects, value added per employment, capital and employment, leads to 0.076 standard deviations increase in log output per worker after partialling out the same controls which, in turn, imply an increase in log value added per employment (growth rate) of 1.1%. The implied magnitude is thus similar to the magnitudes without partialling out the controls.

<sup>37</sup>We obtain measures of TFP growth at a country-sector level at a given period of time using "dual" and "primal" approaches as in Hsieh (1999) and Hsieh (2002)

<sup>38</sup>Results for TFP using alternative log transformation of patents and forward lags are reported in Tables C.2 and C.3

Table 2: 2SLS Estimates: 2000-2014 TFP and VA/EMP growth

	$\Delta \ln(y_{t+n}) \quad n \in \{1, 2, 3\}$								
	VA/EMP			Dual TFP			Primal TFP		
	IV	IV	IV	IV	IV	IV	IV	IV	IV
$\ln(\text{patent}_t)$	0.011 (0.004)	0.009 (0.004)	0.009 (0.003)	0.004 (0.004)	0.008 (0.003)	0.008 (0.003)	0.007 (0.004)	0.010 (0.006)	0.010 (0.005)
$\ln(y_t)$	-0.044 (0.009)	-0.031 (0.007)	-0.033 (0.007)	-0.018 (0.009)	-0.009 (0.005)	-0.009 (0.005)	-0.017 (0.010)	-0.010 (0.009)	-0.011 (0.008)
$\ln(\text{capital}_t)$		-0.005 (0.003)	-0.005 (0.003)		-0.031 (0.003)	-0.031 (0.004)		-0.023 (0.003)	-0.022 (0.003)
$\ln(\text{employ}_t)$		0.005 (0.004)	0.002 (0.005)		0.026 (0.004)	0.023 (0.003)		0.021 (0.004)	0.022 (0.004)
$\ln(\text{int\_import}_t)$			0.003 (0.004)			0.001 (0.003)			-0.002 (0.005)
Country-Year FE	Y	Y	Y	Y	Y	Y	Y	Y	Y
Sector-Year FE	Y	Y	Y	Y	Y	Y	Y	Y	Y
# obs.	8,834	8,357	8,357	7,931	7,931	7,931	8,554	8,336	8,336
# countries	36	36	36	36	36	36	36	36	36
First-stage estimates									
Predicted	0.468	0.461	0.461	0.498	0.470	0.472	0.498	0.472	0.473
$\ln(\text{patent}_t)$	(0.085)	(0.079)	(0.079)	(0.081)	(0.080)	(0.080)	(0.085)	(0.083)	(0.083)
CD Wald F	2,214	2,085	2,069	2,297	2,153	2,165	2,406	2,204	2,212

*Notes:* Period of the analysis is 2000-14 using pre-determined matrix based on the data from 1970-90. First-stage estimates include all the controls. Standard errors are two-way clustered at a country and sector levels in parentheses.

**First-Stage Estimates and Knowledge Spillovers** Before turning to the analysis of our robustness checks, we discuss the first-stage results reported in Table 1. We find positive and significant coefficients across the board of predicted patents constructed using our shift-share design on actual patenting. These estimates inform us directly on the average knowledge spillovers from other country-sector pairs on a given country-sector pair. The estimated coefficient implies an elasticity of 0.46 between the predicted patents from our shift-share design and the actual patenting activity. In terms of magnitude, a 1 standard deviation increase in predicted patents outside country-sector  $(c, s)$  implies a 0.43 increase in actual patenting in country sector  $(c, s)$ . in a sample period.<sup>3940</sup>

<sup>39</sup> An analogous exercise when we use residualized variables with all controls implies 0.46 standard deviation increase.

<sup>40</sup> It is also possible to further investigate knowledge spillovers across countries and sectors by relaxing the restriction we impose in our baseline exercise of only including country sector pairs outside from the country-sector pair of interest. Of course, this is at the expense of endogeneity concerns. However, since we include country-time and sector-time fixed effects, a large array of potential concerns is taken care of by these. We find that if we include the own sector (but exclude own country) in the shift-share design to create predicted patents, the estimated coefficient is 0.45 and 1 std. dev. increase in predicted number of patents implies 0.44 std. dev. increase in actual patents (0.48 resid). Conversely, if we include the own country sectors but still exclude the own sector we find first stage coefficient to be equal to 0.35 and 1 std. dev. increase in predicted number of patents implies 0.34 std. dev. increase in actual patents (0.4 resid). Finally, including both own

Table 3: Checking for Pre-trends

	$\ln(\overline{\text{va.emp}}_t)$			
	(1)	(2)	(3)	(4)
$\ln(\overline{\text{patent}}_{2000-14})$	0.080 (0.036)	0.102 (0.053)	0.032 (0.056)	0.014 (0.047)
Controls	✓	✓	✓	✓
Country FE	Y	Y	Y	Y
Sector FE	Y	Y	Y	Y
# obs.	641	433	433	424
CD Wald F	211.6	159.4	130.0	118.9

*Notes:* Columns (1) and (2) use average value added per employment in the period 2000-14 as a dependent variable computed with WIOD and UNIDO data, respectively. The latter one is included for better compatibility with results in columns (3) and (4), where dependent variable is average value added per employment computed with UNIDO data for the periods 1981-90 and 1971-90, respectively. All regressions include average (log) values for capital, employment and intermediate imports in period 2000-14. Standard errors are clustered at a country level in parentheses.

### 5.2.1 Robustness Checks

The key identifying assumption behind the instrument can be violated if the characteristics of countries and sectors that give more citations to particular sectors and countries in the period 1970-90 had persistent effects on patent activity as well as on changes in the outcomes of interest (beyond our regression controls). We test this assumption in a variety of ways. First, we test for pre-trends, by showing that the pre-period productivity is uncorrelated with subsequent patent activity predicted by the instrument. Table 3 presents the results of regressing average value of productivity during the pre-sample period against average annual number of patent in period 2000-14.<sup>41</sup> The coefficients of this regression, reported in Columns (3) and (4), are not statistically significant. Importantly, they are quantitatively different from the estimates obtained for the period used in main exercises, reported in Columns (1) and (2).

Second, in Column (2) of Table 4, we check that results do hold when we also control for an average level of patent activity in the period 1970-90. The results remain unchanged, however, the coefficient of interest becomes larger in magnitude (in absolute value). Yet,

---

country and sector yields to the coefficient of 0.3, and 1 std. dev. increase in predicted number of patents implies 0.3 std. dev. increase in actual patents (0.39 resid).

<sup>41</sup>As a measure of productivity we use value added per employment data from UNIDO database as data for historical periods is not available in WIOD. We also averaged all the variables in order to suppress the time dimension as the left-hand side and right-hand side of our regression belong to different time periods.

statistically we can not distinguish it from the baseline level.

Next, we rule out the possibility that the links of knowledge diffusion used to construct the instrument capture a demand pull factors from the destination country and sector, rather than a supply push from the origin. To do that, we include in our baseline regression number of patents that should have been filed in pre-sample period to explain actual number of patents observed in the sample in the period of study given citations linkages in pre-sample.<sup>42</sup> Results presented in Column (3) of Table 4 are very stable and the coefficient remains statistically significant and quantitatively close to the baseline. Column (4) includes both historical patent activity and demand-driven number of patents in past simultaneously in a baseline regression as controls.

Table 4: 2SLS Estimates: Robustness

	$\overline{\ln(\text{va\_em}_{c,s,t+n})} \quad n \in \{1, 2, 3\}$			
	(1)	(2)	(3)	(4)
$\ln(\text{patent}_t)$	0.017 (0.008)	0.029 (0.010)	0.025 (0.010)	0.030 (0.010)
$\ln(\overline{\text{patent}_{1970-90}})$		-0.009 (0.005)		-0.009 (0.005)
$\ln(\widehat{\text{patent}}_{t-30})$			-0.006 (0.007)	-0.001 (0.006)
Controls	✓	✓	✓	✓
Country-Year FE	Y	Y	Y	Y
Sector-Year FE	Y	Y	Y	Y
# obs.	8,357	8,357	8,357	8,357
First-stage estimates				
Predicted	0.461	0.264	0.388	0.305
$\ln(\text{patent}_t)$	(0.079)	(0.058)	(0.065)	(0.056)
CD Wald F	2,085	499	815	536

*Notes:* Column (1) shows the results of our baseline regression, Column (2) and (3) include separately to baseline regression historical levels of average patent activity and predicted number of patents driven by demand pull factors, respectively. And Column (4) includes them together. All regressions include (log) values for value added per employment, capital, and employment as controls. Standard errors are two-way clustered at a country and sector levels in parentheses.

<sup>42</sup>We describe procedure used to compute predicted number of patents in pre-sample period driven by demand pull factors in previous section. To deal with time dimension of data, we include in the regression predicted number of patents that should have been filed 30 years in past. The results hold for other choices of lag.



Finally, to check for outliers driving our results, we show that our results remain unchanged if we exclude one country or sector at a time.<sup>43</sup>

### 5.3 Innovation and Long-term Development

We extend now our analysis to longer-time periods. One challenge of looking at long-term outcomes is that high quality value added per employment or TFP panel data spanning a large number of countries and sectors is not available. To circumvent this problem, we adapt our empirical strategy to study the relationship between innovation activity and real GDP per capita at the aggregate country level since 1980.<sup>44</sup> That is, we depart in two dimensions relative to our baseline exercise. First, we abstract from sectoral variation both when we construct instrument, and when we conduct regression analysis. Second, we use real GDP per capita rather than output per worker as our outcome variable.

The choice of the time period for our analysis is the result of a balancing act. On the one hand, since we are interested in long-run growth, we would like to study a long time period. On the other hand, given that comprehensive data for the period prior 1970 are available mostly for advanced countries, and that for the most developing countries we observe almost no innovation activity measured by patents, our shift-share design may miss a part of the variation we are interested to capture. For these reasons, we choose as our baseline time period of analysis of growth rates the 1980-2016, while we use the pre-1980 data to construct our instrument (so that we include the 1970s which have a substantial number of patenting by middle-income economies). Finally, we choose as our baseline set of countries High and Upper Middle Income countries based on the World Bank classification,<sup>45</sup> for which we have substantial variation in patenting activity.<sup>46</sup>

To obtain our shift-share instrument in this cross-country setup, we use only country-time variation in citations to generate the pre-determined matrix of linkages. Each element of the

---

<sup>43</sup>The largest change in magnitude that we obtain in  $\phi_N$  is when we exclude the sector "Manufacture of wood and of products of wood and cork, except furniture; manufacture of articles of straw and plaiting materials." In this case, it increases from 0.017 to 0.023.

<sup>44</sup>Data for real GDP per capita is from Maddison Project Database (Inklaar et al., 2018)

<sup>45</sup><https://datahelpdesk.worldbank.org/knowledgebase/articles/906519-world-bank-country-and-lending-groups>

<sup>46</sup>We have extensively check the robustness of our findings. For example, we find similar results if we construct our shift-share instrument with data pre-1970 or pre-1960 or include all countries in our sample.

matrix is computed as

$$m_{c_o, c_d, \Delta} = \frac{\sum_{t=T_0^{share}}^{T_1^{share}} \sum_{p \in \mathcal{P}(c_o, t)} s_{p \rightarrow (c_d, t - \Delta)}}{\sum_{t=T_0^{share}}^{T_1^{share}} |\mathcal{P}(c_d, t - \Delta)|},$$

and abstract from sectoral variation.<sup>47</sup> We use the citation data observed in the period prior 1980 to construct the pre-existing linkages across countries, and countries' patenting activity during the period starting in 1970 as shifts to construct our instrument for the period 1980-2016.<sup>48</sup>

The empirical specification we run corresponds to Equation (5) in our motivating framework. As a reminder, it is obtained from a combination of a log-linearization of output dynamics around the steady state (as in the standard growth regressions) and our law of motion for TFP. The following specification is used in the analysis

$$\overline{\ln(gdp\_cap_{ct+n})} = \phi_A \ln(gdp\_cap_{ct}) + \phi_N \ln(1 + total\_pat_{ct}) + \delta_t + \delta_c + \varepsilon_{ct}$$

where on the left-hand side we use the average level of GDP per capita over  $n = 3$  years after  $t$  to smooth out variation driven by business cycles and other idiosyncratic shocks.

Table 5 shows the results estimated using citation and patent data for the period prior 1980 to generate the knowledge network, and the period of the analysis for the regression analysis is 1980-2016. As in the previous section, the 2SLS estimates reported in columns (2) and (4) imply a higher elasticity of patenting on income than the OLS estimates in columns (1) and (3). In our preferred specification, which includes country and year fixed effects, we find a positive, significant coefficient that is similar in magnitude to the elasticity of patents to output per worker that we find for the period 2000-2014. The elasticity of patenting to income per capita is 0.034. Quantitatively, this elasticity implies that one standard deviation

<sup>47</sup>As a robustness check, we have also computed our shift-share instrument using cross country and sector variation and then aggregating up the sectoral variation. That is, we compute the linkages at the country-sector level as in our baseline regression and then create our shift-share instrument at the country-sector level first. Then, we aggregate the predicted number of patents across sectors within a country (and year) to construct the instrument. We find very similar results with this alternative procedure.

<sup>48</sup>Similar to our baseline instrument, we use a mix of actual patents filed and predicted patents as shifts. We also do not take into account domestic spillovers when construct the instrument, i.e.  $m_{c_o, c_d} = 0$ , when  $o = d$ . However, we no longer have the intermediate 10 years period between the pre-determined matrix and instrument in our baseline to ensure sufficient sample size of growth rates and inclusion of the 1970s to construct our shift-share. We also perform robustness check analysis where we use all citations available prior 1960/70 to construct the pre-determined matrix of citation linkages, and 1970/80-2016 as a period for the regression analysis. Our main results are robust to alternative samples and can be found in the Table C.5 in the Appendix.

Table 5: 2SLS Estimates: Innovation and Long-term Development: 1980-2016

	Dependent Variable is: $\ln(\overline{gdp\_cap}_{c,t+n})$			
	(1)	(2)	(3)	(4)
$\ln(patent_{c,t})$	0.013 (0.005)	0.086 (0.020)	0.005 (0.003)	0.034 (0.012)
$\ln(gdp\_cap_{c,t})$	0.906 (0.026)	0.735 (0.052)	0.852 (0.025)	0.804 (0.028)
Country FE	Y	Y	Y	Y
Year FE	N	N	Y	Y
# obs.	1,951	1,951	1,951	1,951
# countries	59	59	59	59
First-stage estimates				
Predicted		0.787		1.915
$\ln(patent)$		(0.200)		(0.695)
CD Wald F		203.6		130.0

*Notes:* Period of the analysis is 1980-2016 using pre-determined matrix based on the data for the period pre 1980. Standard errors are clustered at a country level in parentheses. Columns (1) and (3) present the results for OLS, and columns (2) and (4) presents the results obtained with 2SLS. In regressions (1) and (2) only country fixed effects are used. To account for a trend in a number of patents regressions in columns (3) and (4) also include year fixed effects.

increase in logarithm of annual number of patents leads to 0.16 standard deviations increase in logarithm of annual GDP per capita.

## 6 Conclusion

This paper uses a panel of historical patent data spanning the last hundred years and a large range of countries to study the evolution of innovation across time and space and its effect on productivity. First, we have proposed a clustering algorithm to classify finely-defined patent classes based on inventors' patent activity to distill different fields of knowledge. Second, we have documented broad technological waves over the twentieth century and heterogeneous contribution of countries to these. Third, we have documented a substantial rise of international knowledge spillovers as measured by patent citations since the 1990s. This rise is mostly accounted for rising citations to the US and Japanese patents in fields of knowledge related to computation, information processing, and medicine.

After having documenting these facts, we propose a shift-share identification that leverages the knowledge spillovers across fields of knowledge and countries (to construct a the

shift) and the heterogeneity in exposure of countries to technological waves (to construct the share). We then estimate the effect of innovation on output per worker and TFP growth in a panel of countries-sectors for the period 2000-2014 using historical patent data spanning 1970 through 2000. On average, an increase of one standard deviation in patents imply around 4.7% of output per worker growth. We also estimate the effect of innovation on income per capita since 1980 and find similar in magnitude results.

## References

- ACEMOGLU, D. (2009a): *Introduction to Modern Economic Growth*, Princeton University Press.
- (2009b): *Introduction to modern economic growth*, Princeton, NJ [u.a.]: Princeton Univ. Press.
- ACEMOGLU, D., U. AKCIGIT, AND W. R. KERR (2016): “Innovation network,” *Proceedings of the National Academy of Sciences*, 113, 11483–11488.
- AGHION, P., A. DECHEZLEPRETRE, D. HEMOUS, R. MARTIN, AND J. VAN REENEN (2016): “Carbon Taxes, Path Dependency, and Directed Technical Change: Evidence from the Auto Industry,” *Journal of Political Economy*, 124, 1 – 51.
- AGHION, P. AND P. HOWITT (1992): “A Model of Growth through Creative Destruction,” *Econometrica*, 60, 323–51.
- (1998): *Endogenous Growth Theory*, Cambridge, MA: MIT Press.
- AKCIGIT, U. AND S. T. ATEŞ (2021): “Ten Facts on Declining Business Dynamism and Lessons from Endogenous Growth Theory,” *American Economic Journal: Macroeconomics*, 13, 257–98.
- AKCIGIT, U., S. BASLANDZE, AND F. LOTTI (2018): “Connecting to Power: Political Connections, Innovation, and Firm Dynamics,” Working Paper 25136, National Bureau of Economic Research.
- AKCIGIT, U., J. GRIGSBY, AND T. NICHOLAS (2017): “The Rise of American Ingenuity: Innovation and Inventors of the Golden Age,” CEPR Discussion Papers 11755, C.E.P.R. Discussion Papers.
- ALTUNTAS, S., T. DERELI, AND A. KUSIAK (2015): “Analysis of patent documents with weighted association rules,” *Technological Forecasting and Social Change*, 92, 249–262.
- ARCAND, J.-L., E. BERKES, AND U. PANIZZA (2015): “Too much finance?” *Journal of Economic Growth*, 20, 105–148.
- BARRO, R. (1991): “Economic Growth in a Cross Section of Countries,” *The Quarterly Journal of Economics*, 106, 407–443.
- BARRO, R., R. BARRO, X. SALA-I MARTIN, X. SALA-I MARTIN, AND M. I. OF TECHNOLOGY (2004): *Economic Growth*, The MIT Press, MIT Press.

- BARRO, R. J. AND X. SALA-I MARTIN (1992): “Convergence,” *Journal of Political Economy*, 100, 223–251.
- BERKES, E. AND R. GAETANI (2018a): “Income Segregation and Rise of the Knowledge Economy,” 2018 Meeting Papers 213, Society for Economic Dynamics.
- (2018b): “Income Segregation and the Rise of the Knowledge Economy,” 2018 Meeting Papers 213, Society for Economic Dynamics.
- BLOOM, N., C. I. JONES, J. VAN REENEN, AND M. WEBB (2020): “Are Ideas Getting Harder to Find?” *American Economic Review*, 110, 1104–44.
- BORUSYAK, K., P. HULL, AND X. JARAVEL (2018): “Quasi-Experimental Shift-Share Research Designs,” NBER Working Papers 24997, National Bureau of Economic Research, Inc.
- BOTTAZZI, L. AND G. PERI (2003): “Innovation and spillovers in regions: Evidence from European patent data,” *European Economic Review*, 47, 687–710.
- BRESCHI, S., F. LISSONI, AND F. MALERBA (2003): “Knowledge-relatedness in firm technological diversification,” *Research Policy*, 32, 69–87.
- CLARK, G. AND R. C. FEENSTRA (2003): “Technology in the Great Divergence,” in *Globalization in Historical Perspective*, National Bureau of Economic Research, Inc, NBER Chapters, 277–322.
- COELLI, F., A. MOXNES, AND K. H. ULLTVEIT-MOE (2016): “Better, Faster, Stronger: Global Innovation and Trade Liberalization,” NBER Working Papers 22647, National Bureau of Economic Research, Inc.
- COMIN, D. AND B. HOBIJN (2004): “Cross-country technology adoption: making the theories face the facts,” *Journal of Monetary Economics*, 51, 39–83.
- (2010): “An Exploration of Technology Diffusion,” *American Economic Review*, 100, 2031–2059.
- COMIN, D., D. LASHKARI, AND M. MESTIERI (2019): “Structural Transformation of Innovation,” 2019 Meeting Papers 1394, Society for Economic Dynamics.
- COMIN, D. AND M. MESTIERI (2018): “If Technology Has Arrived Everywhere, Why Has Income Diverged?” *American Economic Journal: Macroeconomics*, 10, 137–178.

- DECHEZLEPRÊTRE, A., D. HÉMOUS, M. OLSEN, AND C. ZANELLA (2020): “Automating Labor: Evidence from Firm-Level Patent Data,” CEP Discussion Papers dp1679.pdf, Centre for Economic Performance, LSE.
- DURLAUF, S. N., P. A. JOHNSON, AND J. R. TEMPLE (2005): “Growth Econometrics,” in *Handbook of Economic Growth*, ed. by P. Aghion and S. Durlauf, Elsevier, vol. 1 of *Handbook of Economic Growth*, chap. 8, 555–677.
- EATON, J. AND S. KORTUM (1999): “International Technology Diffusion: Theory and Measurement,” *International Economic Review*, 40, 537–570.
- ENGELSMAN, E. AND A. VAN RAAN (1994): “A patent-based cartography of technology,” *Research Policy*, 23, 1–26.
- FU, K., J. CHAN, J. CAGAN, K. KOTOVSKY, C. SCHUNN, K. WOOD, AND P. PILLAR (2012): “The Meaning of “Near” and “Far”: The Impact of Structuring Design Databases and the Effect of Distance of Analogy on Design Output,” vol. 135.
- GOLDSMITH-PINKHAM, P., I. SORKIN, AND H. SWIFT (2020): “Bartik Instruments: What, When, Why, and How,” *American Economic Review*, 110, 2586–2624.
- GRILICHES, Z. (1979): “Issues in Assessing the Contribution of Research and Development to Productivity Growth,” *Bell Journal of Economics*, 10, 92–116.
- (1986): “Productivity, R&D, and the Basic Research at the Firm Level in the 1970’s,” *American Economic Review*, 76, 141–154.
- HIDALGO, C., B. KLINGER, A.-L. BARABASI, AND R. HAUSMANN (2007): “The Product Space Conditions the Development of Nations,” *Science (New York, N.Y.)*, 317, 482–7.
- HORNBECK, R. AND E. MORETTI (2019): “Estimating Who Benefits from Productivity Growth: Direct and Indirect Effects of City Manufacturing TFP Growth on Wages, Rents, and Inequality,” IZA Discussion Papers 12277, Institute of Labor Economics (IZA).
- HSIEH, C.-T. (1999): “Productivity Growth and Factor Prices in East Asia,” *American Economic Review*, 89, 133–138.
- (2002): “What Explains the Industrial Revolution in East Asia? Evidence From the Factor Markets,” *American Economic Review*, 92, 502–526.
- INKLAAR, R., H. DE JONG, J. BOLT, AND J. VAN ZANDEN (2018): “Rebasing ‘Maddison’: new income comparisons and the shape of long-run economic development,” GGDC



- Research Memorandum GD-174, Groningen Growth and Development Centre, University of Groningen.
- JAFFE, A. (1986): “Technological Opportunity and Spillovers of R&D: Evidence from Firms’ Patents, Profits, and Market Value,” *American Economic Review*, 76, 984–1001.
- JAFFE, A., M. TRAJTENBERG, AND R. HENDERSON (1993): “Geographic Localization of Knowledge Spillovers as Evidenced by Patent Citations,” *The Quarterly Journal of Economics*, 108, 577–598.
- JONES, C. I. (1995): “R&D-Based Models of Economic Growth,” *Journal of Political Economy*, 103, 759–784.
- (1999): “Growth: With or Without Scale Effects?” *American Economic Review*, 89, 139–144.
- KANG, B. AND G. TARASCONI (2016): “PATSTAT revisited: Suggestions for better usage,” *World Patent Information*, 46, 56–63.
- KAY, L., N. NEWMAN, J. YOUTIE, A. L. PORTER, AND I. RAFOLS (2014): “Patent overlay mapping: Visualizing technological distance,” *Journal of the Association for Information Science and Technology*, 65, 2432–2443.
- KELLER, W. (2002): “Geographic Localization of International Technology Diffusion,” *American Economic Review*, 92, 120–142.
- (2004): “International Technology Diffusion,” *Journal of Economic Literature*, 42, 752–782.
- KELLER, W. AND S. R. YEAPLE (2013): “The Gravity of Knowledge,” *American Economic Review*, 103, 1414–1444.
- KLENOW, P. AND A. RODRÍGUEZ-CLARE (1997): “The Neoclassical Revival in Growth Economics: Has It Gone Too Far?” in *NBER Macroeconomics Annual 1997, Volume 12*, National Bureau of Economic Research, Inc, NBER Chapters, 73–114.
- KOGLER, D., D. RIGBY, AND I. TUCKER (2013): “Mapping Knowledge Space and Technological Relatedness in US Cities,” *European Planning Studies*, 21, 1374.
- KORTUM, S. S. (1997): “Research, Patenting, and Technological Change,” *Econometrica*, 65, 1389–1420.

- LEYDESDORFF, L. (2008): “Patent classifications as indicators of intellectual organization,” *Journal of the American Society for Information Science and Technology*, 59, 1582–1597.
- LEYDESDORFF, L., D. KUSHNIR, AND I. RAFOLS (2014): “Interactive overlay maps for US patent (USPTO) data based on International Patent Classification (IPC),” *Scientometrics*, 98, 1583–1599.
- LEYDESDORFF, L. AND L. VAUGHAN (2006): “Co-occurrence matrices and their applications in information science: Extending ACA to the Web environment,” *Journal of the American Society for Information Science and Technology*, 57, 1616–1628.
- MORETTI, E., C. STEINWENDER, AND J. V. REENEN (2019): “The Intellectual Spoils of War? Defense R&D, Productivity and International Spillovers,” NBER Working Papers 26483, National Bureau of Economic Research, Inc.
- MURATA, Y., R. NAKAJIMA, R. OKAMOTO, AND R. TAMURA (2014): “Localized Knowledge Spillovers and Patent Citations: A Distance-Based Approach,” *The Review of Economics and Statistics*, 96, 967–985.
- NAKAMURA, H., S. SUZUKI, I. SAKATA, AND Y. KAJIKAWA (2015): “Knowledge combination modeling: The measurement of knowledge similarity between different technological domains,” *Technological Forecasting and Social Change*, 94, 187–201.
- NICHOLAS, T. (2010): “The Role of Independent Invention in U.S. Technological Development, 1880–1930,” *The Journal of Economic History*, 70, 57–82.
- PACKALEN, M. AND J. BHATTACHARYA (2015): “Cities and Ideas,” Working Paper 20921, National Bureau of Economic Research.
- PERETTO, P. (1998): “Technological Change and Population Growth,” *Journal of Economic Growth*, 3, 283–311.
- PETRALIA, S., P.-A. BALLAND, AND D. L. RIGBY (2016): “Unveiling the geography of historical patents in the United States from 1836 to 1975,” *Scientific Data*, 3, 160074.
- PHILIPPE AGHION, ANTONIN BERGEAUD, M. L. AND M. J. MELITZ (2018): “The Impact of Exports on Innovation: Theory and Evidence,” Working papers 678, Banque de France.
- ROMER, P. (1990): “Endogenous Technological Change,” *Journal of Political Economy*, 98, S71–102.

- SEGERSTROM, P. S. (1998): “Endogenous Growth without Scale Effects,” *American Economic Review*, 88, 1290–1310.
- TABELLINI, M. (2020): “Gifts of the Immigrants, Woes of the Natives: Lessons from the Age of Mass Migration,” *Review of Economic Studies*, 87, 454–486.
- VAN LOOY, B., C. VEREYEN, AND U. SCHMOCH (2014): “Patent Statistics: Concordance IPC V8 – NACE Rev.2,” *EUROSTAT*.
- VON WARTBURG, I., T. TEICHERT, AND K. ROST (2005): “Inventive progress measured by multi-stage patent citation analysis,” *Research Policy*, 34, 1591–1607.
- YOUNG, A. (1998): “Growth without Scale Effects,” *Journal of Political Economy*, 106, 41–63.
- ZITT, M., E. BASSECOULARD, AND Y. OKUBO (2000): “Shadows of the Past in International Cooperation: Collaboration Profiles of the Top Five Producers of Science,” *Scientometrics*, 47, 627–657.

**[Link to Appendix \(Click Here\)](#)**