Federal Reserve Bank of Chicago

Global Innovation Spillovers and Productivity: Evidence from 100 Years of World Patent Data

Enrico Berkes, Kristina Manysheva, and Martí Mestieri

May 3, 2022

WP 2022-15

Global Innovation Spillovers and Productivity: Evidence from 100 Years of World Patent Data*

Enrico Berkes

Ohio State

Kristina Manysheva

Northwestern

Martí Mestieri

FRB of Chicago,

UPF, BSE and CREi

May 3, 2022

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 sectoral output per worker and total factor productivity (TFP) growth in a panel of country-sectors from 2000 to 2014, as well as on aggregate income per capita since 1960. To assess causality, we develop a shift-share instrument that leverages pre-existing citation linkages across countries and fields of knowledge, as well as heterogeneous countries' exposure to technology waves. On average, an increase of one standard deviation in log-patenting activity increases sectoral output per worker growth by 1.1 percentage points. We find results of similar magnitude for sectoral TFP growth and long-run aggregate income per capita growth.

Keywords: Innovation, Technology Diffusion, Patents.

JEL Classification: O10, O30, O33, O47.

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 20th 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.[1]

In this paper, we use historical patent data spanning a vast range of countries over the past 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 (through patent creation) and knowledge spillovers, i.e., how innovation builds on previous knowledge (through patent citations). We document a substantial rise of international knowledge spillovers since the 1990s mostly driven by the United States and Japan, as well as the rise of innovation related to computation, information and communication technologies (ICTs), and medicine. We leverage the rich structure of citation linkages across time, space, and fields of knowledge (FoK) to propose an identification strategy to quantify the effect of innovation induced by knowledge spillovers on productivity and economic growth across countries and industries. To the best of our knowledge, our identification strategy is novel to the endogenous growth 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. 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 first calculates the probability that the same inventor patents inventions in multiple technology classes. It then uses these probabilities to infer the proximity of technology classes in the knowledge space and to create knowledge clusters.[2]

---

[1]See Comin and Mestieri (2014) and the references therein documenting 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 past two centuries. Their analysis is circumscribed to 25 major technologies since 1780.

[2]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 patent grant.

Armed with our newly defined technology classes, we show that their significance—as measured by the share of patents across fields of knowledge—has importantly evolved over time. The data reveal substantial technological waves in the past one hundred years. For example, mechanical engineering accrued the largest share of innovations near the beginning of the 20th century. Fields of knowledge related to chemistry and physics (e.g., macromolecular compounds) were the most prominent fields around the mid-century mark, while inventions related to medicine and the digital economy appear to be the most prevalent at the end of the 20th century and over the most recent 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.

Next, we turn our attention to knowledge spillovers. We measure knowledge spillovers through patent 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 toward patents within the same field of knowledge. We also document that across all these categories, there is an upward trend over time in terms of citations. That is, new patents tend to cite other patents more.

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 patents are cited more than twice as much as domestic patents. This finding suggests that the reliance on knowledge produced elsewhere—and particularly in the US and Japan—has markedly increased over this period of time. Even for technology leaders such as Germany and the United Kingdom, foreign citations now account for most of the citations. The increase is mainly driven by a handful of fields of knowledge that are related to ICTs and medicine. This fact may be interpreted as a decline in the prominence of European inventions relative to their US and Japanese counterparts.

After having laid out these facts, we investigate the effect of innovation (as measured by patenting) on productivity and income. Our empirical specification is guided by a simple theoretical framework that incorporates patents and patent citations in a multi-sector growth model. Our baseline regression 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 then extend our analysis back in time and study the effect of innovation on long-run income growth (for the periods 1980-2016 up to 1960-2016).

Simply correlating innovation and productivity or output per worker is problematic be-

cause of measurement 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. We address these endogeneity concerns by constructing a shift-share instrument that exploits country and time variation in technological waves and the network structure of knowledge spillovers, in the spirit of Acemoglu et al. (2016) and Berkes and Gaetani (2022). In particular, our proposed shift-share design leverages pre-existing knowledge linkages across countries and technologies measured through patent citations to construct the share component of our instrument. The shift component is obtained using lagged foreign innovative output in other fields of knowledge and countries as measured by patent filings.[3]

In our baseline regression, the 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 level (e.g., sectoral capital and labor, along with differential country and sectoral trends). We use patent data starting in 1970 to construct our instrument for this exercise. We find a robust effect of innovation on value added per employment growth. One standard deviation increase in patenting activity leads to a 0.078 standard deviation increase in output per worker growth (after partialling out the regression controls), implying an increase in output per worker growth of 1.1 percentage points. When we estimate the effect of innovation on TFP growth, we find a very similar result in magnitude—a result that is implied by our theoretical framework.

We conduct a number of robustness checks to address concerns regarding the validity of the instrument, such as the existence of pre-trends or demand-pull anticipatory effects that might be correlated with the contemporaneous state of the local economy. To do this, among other things, we show that the pre-period productivity is uncorrelated with subsequent patent activity predicted by the instrument. We also "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.

---

[3]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, respectively) 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 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, and it is an increasing function of the strength of these links. Applying this procedure recursively, we obtain a predicted number of patents for each country and field of knowledge. In fact, we refine this procedure and extend this logic to higher-order linkages to create our main instrument (see Section 5).

We conclude our paper by extending our framework to study the effect of innovation on long-run income per capita growth. In our first exercise, we estimate the effect of innovation on income per capita over the 1980-2016 period. We reconstruct our shift-share instrument using pre-1980 patent data, thus exploiting the full extent of the time coverage of our dataset. Using pre-1980 data allows us to cover patenting activity of virtually all high-income and upper-middle-income countries (as defined by the World Bank). We perform a second exercise by estimating the effect on income per capita growth starting in 1960 and 1970. In this case, we construct our instrument using the pre-1960 and pre-1970 patent data, respectively. While covering a longer time span, we lose information on the patenting activity of many upper-middle-income countries. Despite this, we find a positive, comparable-in-magnitude, significant effect of patenting on income per capita growth across these different time periods. An increase of one standard deviation in log patenting implies an increase in the growth of income per capita between 1.6 and 2.8 percentage points. The implied changes in growth rates represent 24% and 41% of a standard deviation of income per capita growth, respectively.

Related Literature This paper relates to the vast and rich literature studying the link between innovation and productivity since, at least, the seminal work of Griliches (1979, 1986). Similar to Kogan et al. (2017), who find large positive effects of patented inventions on firm growth and productivity, we document positive effects of innovation on output and productivity growth at the country-sector level. Our instrumental variable approach leverages knowledge spillovers and the diffusion of technology as measured by patent citations. The existence of knowledge spillovers has 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; Buera and Oberfield, 2020; Keller, 2004; and Melitz and Redding, 2021 provide excellent surveys). We contribute to this strand of the literature by documenting an increase of international spillovers since the 1990s and by using international linkages to build our shift-share design and, ultimately, quantify the effect of innovation on productivity.

In addition, our paper contributes to a recent literature that uses historical patent data to shed light on various linkages between innovation and long-run outcomes, e.g., Nicholas (2010), Packalen and Bhattacharya (2015), Petralia et al. (2016), and Akcigit et al. (2017). One difference with most of this literature is that we extend our analysis beyond one country and aim to provide a global view. To the best of our knowledge, this is the first paper that uses the entire coverage of the PATSTAT database to study patenting activity. Regarding

the goal of providing a global view, our work is perhaps 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.

This paper is also related to the growing literature that incorporates networks in the analysis of different aspects of economic growth and trade (e.g., Acemoglu et al., 2015; Oberfield, 2018; Liu, 2019; Baqaee and Farhi, 2019; and Kleinman et al., 2021). In this regard, our work complements the recent work by Ayerst et al. (2020) and Liu and Ma (2021), who use international patent data to study the diffusion of knowledge embedded in trade patterns and the design of optimal R&D policies in the presence of international knowledge spillovers, respectively. Finally, our network-based shift-share instrumental approach is related to a number of papers that have used the network structure of patent citations to construct shift-share instruments. Our approach is most similar to Berkes and Gaetani (2022), who construct a shift-share instrument leveraging the network of citations across US cities, and Acemoglu et al. (2016), who use a citation network to percolate sectoral innovations through the innovation network and illustrate how technological progress builds upon itself. Both papers focus on the United States.[4]

2   Data

2.1   Data Sources

In this paper, we measure new ideas through patent data and productivity through value added per worker and TFP. Patent data are collected from the European Patent Office's 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.[5] 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

---

[4]A large number of papers have used more standard shift-share ("Bartik") 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.

[5]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); and Dechezleprêtre et al. (2020).

or more dimensions are often missing for some countries in the years preceding 1970. We therefore split our sample into two groups of countries, which we use at different stages of our analysis. The first group is composed of six major technological leaders – the United States, the United Kingdom, France, Germany, the Soviet Union, and Switzerland – for which all the patents' characteristics required by our analysis are available at least since 1920.[6] The second group includes all the countries covered by PATSTAT and starts in 1970.[7] 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, we use instead the country of the assignee(s) or publication authority. 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 equally to the development of the invention.[8] To avoid double-counting patents that are filed in more than one patent office, we restrict our analysis to patents that are the first in their (DOCDB) family (except for our citation analysis, in which we count all given citations to any patent in a 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 four-digit level (e.g., A01B)—for a total of 650 classes—and we then cluster them into consistent groups following the machine-learning 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.[9]

We supplement the patent data with the World Input Output Database (WIOD, Timmer et al. 2015). 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,

---

[6]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 combine all the patents produced by Armenia, Azerbaijan, Belarus, Estonia, Georgia, Kazakhstan, Kyrgyzstan, Latvia, Lithuania, Moldova, Russia, Tajikistan, Turkmenistan, Ukraine, and Uzbekistan.

[7]For our empirical analysis, we exclude China from our sample because of a substantial rise in the number of Chinese patents since the third revision of the patent law in China in 2008. While we see a sharp increase in the 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 are made up of all the 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.

[8]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.

[9]We discuss in more detail our data construction procedure in Appendix A.1

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.[10] Finally, we use data from the Maddison Project Database (Inklaar et al., 2018) for the historical analysis of income per capita growth presented in the Section 5.3.

2.2 Construction of Fields of Knowledge

Innovation is the process of creating new knowledge, potentially 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.[11]

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 is that because 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. More precisely, we build a probability matrix $T_{642\times642}$,[12] 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.[13] 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.[14]

---

[10]See details in Appendix A.2. In the Appendix, we also discuss the additional database we use (i.e., UNIDO INDSTAT2) for historical data on sectoral manufacturing output by country and the Penn World Data Tables.

[11]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.

[12]Eight IPC subclasses whose second level is 99 (i.e., "Subject Matter not otherwise Provided for in this Section") were excluded from the analysis because they are assigned to patents with no clear identified technology.

[13]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 because different IPC codes might weight differently in terms of their importance and centrality relative to other IPC codes within a given field of knowledge. 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).

[14]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; and Leydesdorff et al., 2014). We also conduct such grouping as a robustness check and find substantial overlap. Another strand of the measures uses the "co-classification" information of patents

We then use a *k-medoids* clustering algorithm to group the IPC subclasses into knowledge clusters. We interpret each resulting cluster as a field of knowledge, and use this classification to analyze the evolution of patenting in the next section. The *k-medoids* algorithm minimizes the distance within clusters by comparing all possible permutations of subclasses, conditional on a specific number of clusters, k. To determine the optimal number of clusters, we first compute the optimal clustering for each possible k and we then "score" each result using the silhouette coefficient. 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.[15] 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.[16]

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. We use the fields of knowledge created in Section 2.2 as our unit of analysis of the technology space.

3.1 Evolution of Fields of Knowledge across Space and Time

We first document the evolution of the major fields of knowledge for the past hundred years and highlight how different countries contributed to their growth at different points in 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

---

(Jaffe, 1986; Engelsman and van Raan, 1994; Breschi et al., 2003; Leydesdorff, 2008; Kogler et al., 2013; and Altuntas et al., 2015). Others used the likelihood of diversification as measures of distance (Hidalgo et al., 2007) and analysis of patent texts (Fu et al., 2012, and Nakamura et al., 2015).

[15]To apply the *k-medoids* algorithm requires the creation of a dissimilarity matrix D, which needs to be symmetric. To obtain such dissimilarity matrix, we apply the following transformation to the inventor probability matrix:

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

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 in our *k-medoids* clustering algorithm to group the IPC subclasses into clusters. More details on the procedure used to construct fields of knowledge can be found in Appendix A.4.

[16]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 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.

number of forward citations.[17] We split our data set 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, especially around the 1990s – approximately 10% of the fields of knowledge account for 60% percent of all patent activity in the 2000s compared with 30% in the first half of the 20th century. Second, there is substantial heterogeneity in the evolution of fields of knowledge over time. At the beginning of the 20th 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 or Medical Preparation). Finally, and as expected, around the mid-1990s the fields of knowledge related to computing and communication techniques started playing a leading role in the innovation landscape.

We 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 do this, we build importance measures that take into consideration country fixed effects or patents that were cited at least once. Table B.2 in the Appendix 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: 1920-1944, 1945-1969, 1970-1994, and 1995-2015. We concentrate our analysis to the seven fields of knowledge that took a 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 patenting share in a certain field of knowledge.[18]

Because of data limitations, for the period 1920-1970, our sample is made up of just six countries: the US, the UK, France, Germany, the Soviet Union, and Switzerland. Figure B.1

---

[17]Note that we are using only the first patent of the family. Moreover, 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.

[18]To account for potential differences in how countries assign patent citations, in this part of the analysis, we use the total number of patents without weighting by the number of citations for better comparability. We also verify that the results are robust to citation-weighted measures.

Figure 1: Evolution of Top Fields of Knowledge
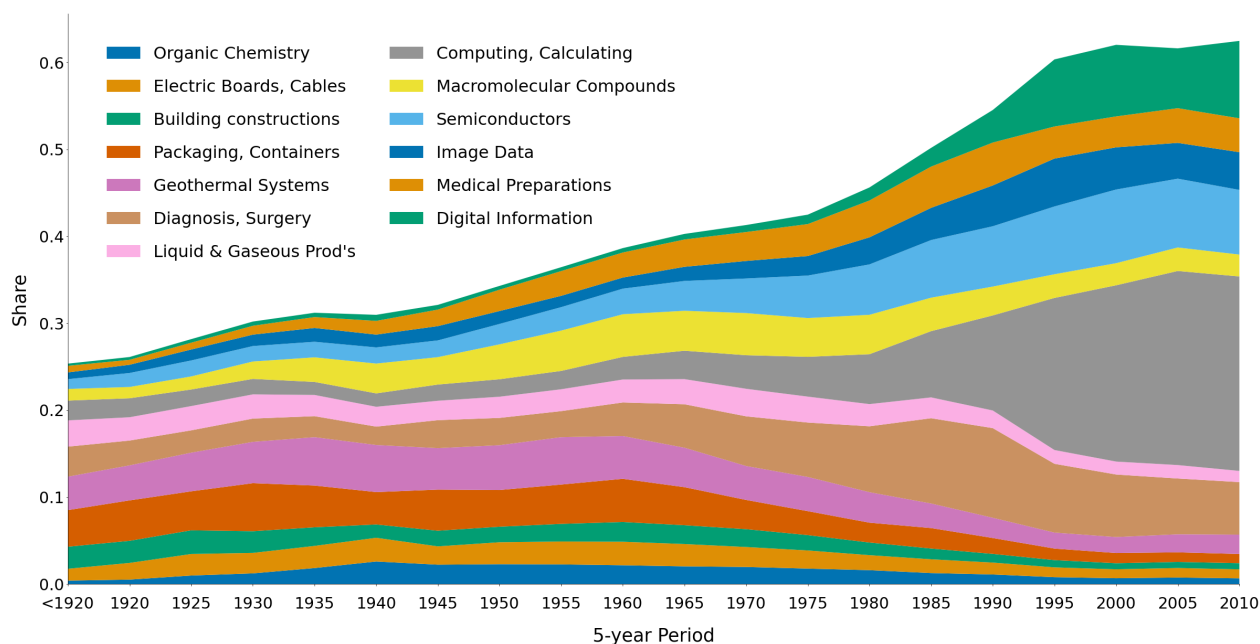

Notes: This figure represents the share of each field of knowledge, measured by the number of first-in-the-family patents weighted by backward citations, in total patent activity across all fields in a given period of time. The width of the colored bars reflects the share of the knowledge field. Exact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

in the Appendix 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 World War II and 1970.

In Figure 2, we consider the whole sample of countries 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. The US also gains substantial prevalence in the second part of the sample. After 1995 other Asian countries, such as South Korea, start rising to the forefront of the technological frontier. In this period, France experiences a decrease in 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 extend our analysis beyond the top 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's ranking is similar to the one of the US in 1970 and it declines subsequently, while Asian countries such as

Figure 2: Countries Shares in Top Fields, 1970-2015
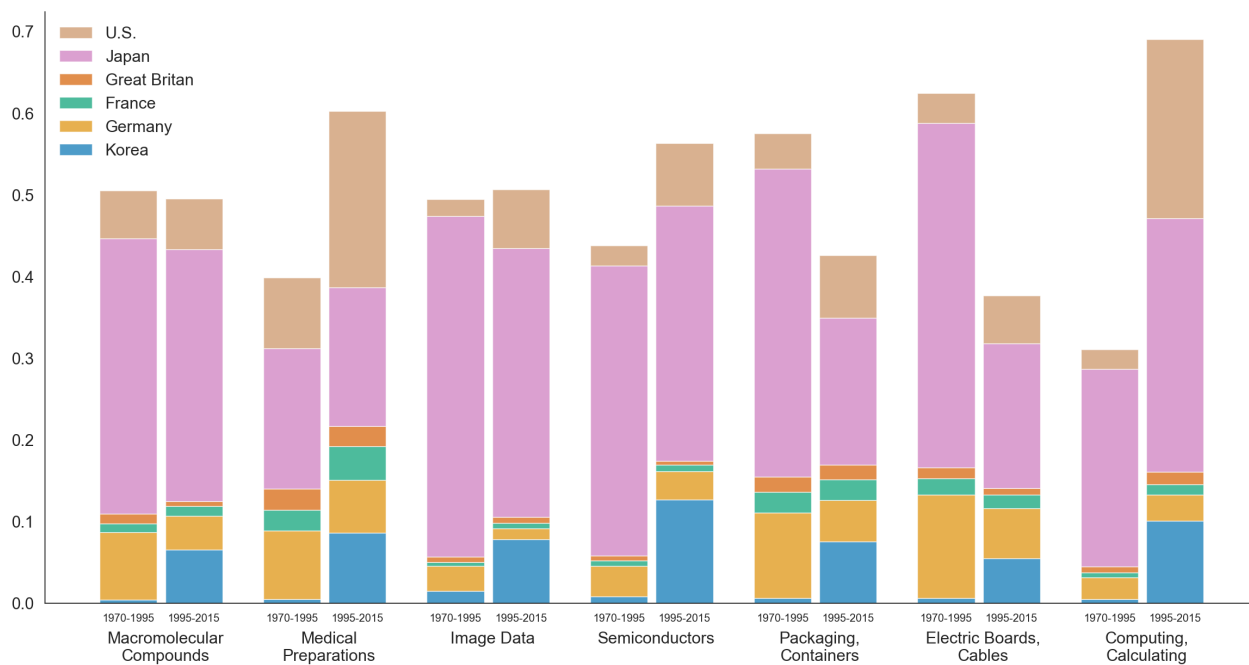
Taiwan gain prominence after the 2000s. See Section B in the Appendix for further details and discussion of this exercise.[19]

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 the composition of the technological output and in terms of the 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, and 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.

We focus our analysis on the post-1970 sample, for which we have data on filed patents

---

[19]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 gross domestic product (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.

for virtually all countries in the world. We compute the citations given by these patents 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 patent citations tend to be more prominent than international patent citations: domestic patents are cited at roughly double the rate that of international patents are. Panel (b) breaks down these trends by additionally looking at whether citations belong to the field of knowledge (or FoK, as noted in Figure 3) of the citing patent.[20] The plot shows that citations tend to be concentrated not only geographically (i.e., domestic patents being cited relatively more), but also technologically (i.e., patents in the same field of knowledge being cited relatively more). Moreover, these gaps appear to have widened over the past decades.

An important pattern that is revealed by our analysis 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, as we have already seen in Figure 2, for the period 1970-2015 Japan and the United States are responsible for the largest share of patents produced worldwide. Panels (c) and (d) of Figure 3 separately depict 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, make 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, at least according to the data on patent citations.[21]

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 on the backward citations received by the five leading fields of knowledge over the past five 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, to a lesser extent, Transmission of Digital Information. What is perhaps even more striking is the fact that most citations to

---

[20]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.

[21]Decomposition of citations for other countries, namely, Germany, France, and the UK, are reported in Figure B.2. The plots for these three frontier countries show how they moved from mostly relying on domestic knowledge in the early periods to foreign knowledge later in the sample.
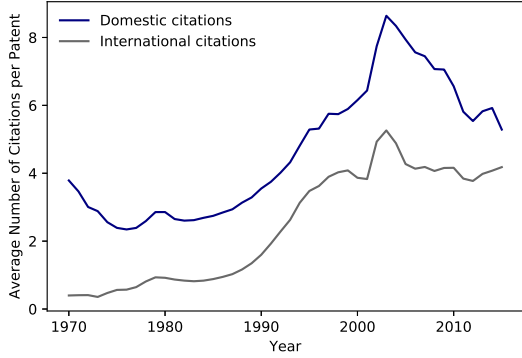
(a) Domestic vs. International Citations
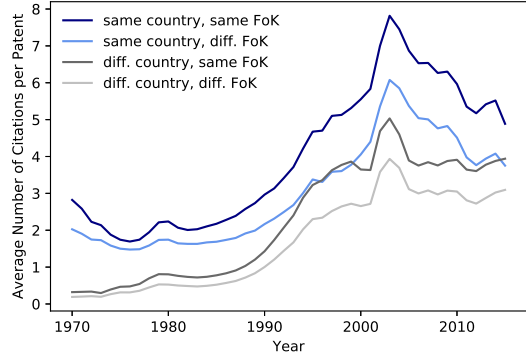
(b) Citation Decomposition

(c) US and Japan
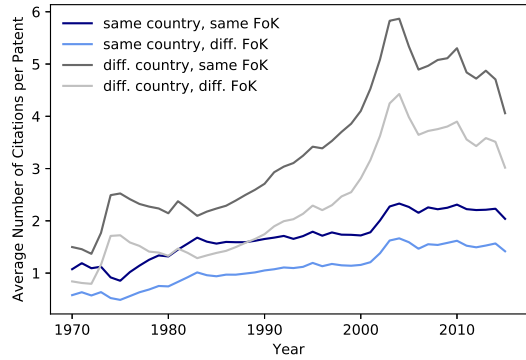
(d) All Countries except US and Japan

Figure 3: Citation Dynamics, 1970-2015

this field of knowledge are given to US and Japanese patents, as illustrated by Figure 4.[22]

Taken together, the evidence presented in this section paints a picture consistent with the view that knowledge spillovers have increasingly become an important component of the innovation process in the past few decades. Although spillovers that originate from the same country and field of knowledge are still the most relevant, internati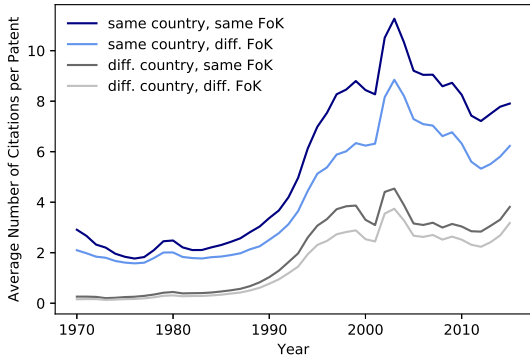onal knowledge spillovers have been steadily gaining importance over the past few decades. This increase is visible when considering spillovers coming both from the same field of knowledge and from other fields of knowledge, and it is mainly driven by a dramatic increase in the citations received by US and Japanese patents, especially in the fields of knowledge related to computing, information processing, and medicine.

---

[22]Relatedly, Liu and Ma (2021) document a high reliance on domestic knowledge in both the US and Japan using Google Patents' global patent data for 40 countries during the period 1976-2020.

(a) US and Japan	(b) All Countries except US and Japan

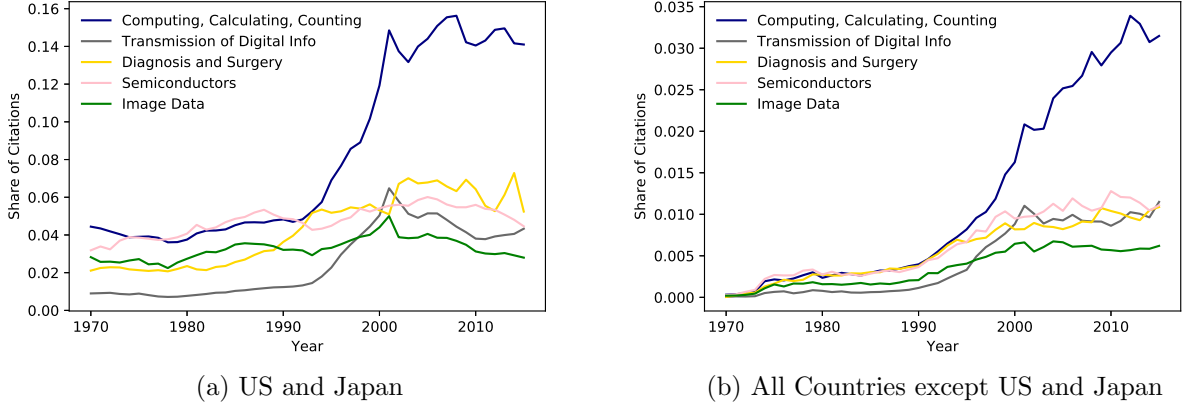Figure 4: Share of citations to US and Japanese patents by FoK, 1970-2015. Each line in the plots represents the share of citations to US and Japanese patents that belong to a given field of knowledge. Panel (a) depicts the shares of domestic citations given by US and Japanese patents, and panel (b)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

4 Conceptual Framework

In this section, we present the framework that will guide our empirical analysis. This framework incorporates patents and patent citations to a standard, multi-sector growth model.[23] Importantly, our framework only specifies the production-side of the economy, and it does not assume the existence of a balanced growth path of output or productivity at the sectoral (or aggregate) level.[24]

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. 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 $\mathbf{N}_t \equiv (N_{111t}, \ldots, N_{cskt}, \ldots, N_{CSKt})$. 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, ΔN_{cskt}; the current stock of knowledge, $\mathbf{N}_t$; and inputs devoted to generate new ideas, R_{cskt};

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

where Δ denotes the time difference operator between $t + 1$ and t. The spillover function

---

[23]Our formulation builds upon previous studies that have been applied to the study of the patent network of citations, such as 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.

[24]Unbalanced sectoral growth is indeed the empirically relevant case for the United States and other advanced economies (Comin et al., 2019).

$S_{csk}(\mathbf{N}_t)$ captures how the current world stock of knowledge $\mathbf{N}_t$ helps generate new ideas in country c, field of knowledge k, and sector s. We assume the spillover function to be

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

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. We leverage this structure to construct our instrumental variable. Note that we purposefully state Equation (1) generically so that it subsumes the first generation endogenous growth models as in Romer (1990) or Aghion and Howitt (1992), semi-endogenous growth models as in Jones (1995), Kortum (1997), or Segerstrom (1998), or second generation models as in Aghion and Howitt (1998), Young (1998), or Peretto (1998).[25]

Since ideas are to a large extent non-rival (Romer, 1990), the vast majority of endogenous growth theories resort to intellectual protection in the form of patents to ensure that investments in new ideas can be recovered with future profits.[26] 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 hard to obtain 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, in our instrumental variables strategy. In practice, however, not all ideas are patented, and not all ideas which a patent builds on are cited. We thus think of patents as a *proxy* for new ideas, ΔN_{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.

In our framework, there is a representative firm in each country-sector that produces sectoral output combining physical inputs (labor and capital) according to the best production methods available in that country-sector at time t, which are summarized by sectoral TFP, denoted TFP_{cst}. Sectoral value added per worker, y_{cst}, is given by the Cobb-Douglas production function $\log y_{cst} = \phi_{cst} + \log TFP_{cst} + \alpha \log k_{cst}$, where k_{sct} denotes capital per worker, $0 < \alpha < 1$, and ϕ_{cst} denotes potential additional sources of variation of total productivity that are not captured by our framework. To obtain the baseline empirical specification, we assume that this effect can be parameterized as a full set of dyadic fixed effects, $\phi_{cst} = \tilde{\delta}_{ct} + \tilde{\delta}_{st} + \tilde{\delta}_{cs}$. This parametrization captures the fact that the productivity of ideas (and/or other unmodeled sources of productivity differences) may differ across country-sector-time pairs because

---

[25]For example, one specification extensively used in the literature (e.g., Romer, 1990, and 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 relationship, $I = N_{ct}^{\phi} R_{ct}$ with $\phi \leq 1$.

[26]See, among others, Aghion and Howitt (1998), Acemoglu (2009a), and the references therein.

some country-sector pairs may be better suited at certain sectors than others (captured by $\tilde{\delta}_{cs}$), there may be some global technology trends affecting certain sectors (captured by $\tilde{\delta}_{st}$), or there may be some country-specific shocks (captured by $\tilde{\delta}_{ct}$).

Following the endogenous growth literature, we assume that the role of ideas is to increase firms' productivity by developing and improving methods of production (e.g., Acemoglu, 2009b). That is, we assume that there is a positive relationship between ideas produced and sectoral TFP growth. Moreover, as TFP grows and new production methods are implemented, we allow for the existence of fixed costs of adjustment scaling up with (a power function of) total output. This adjustment cost stands in for production disruptions related to the adoption of new technologies (e.g., as in Perla and Tonetti, 2014 or Comin and Gertler, 2006). In particular, our empirical specification assumes an iso-elastic relationship between TFP growth, ideas, and adjustment costs,

$$\log\left(\frac{TFP_{cst+1}}{TFP_{cst}}\right) = \phi_0 + \phi_N \log(1 + \Delta N_{cst}) - \phi_Y \log y_{cst}, \tag{3}$$

where $\phi_0, \phi_N, \phi_Y \geq 0$ and $\Delta N_{cst} = \sum_{k=1}^{K} \Delta N_{cskt}$ denotes the total number of ideas generated in country c and sector s at time t across all fields of knowledge. Combining the idea production function, Equation (1), with the TFP Equation (3), we can readily verify that our framework nests a number of cases often considered in the literature, such as endogenous and semi-endogenous growth models.[27]

To derive our baseline empirical specification, we take the time difference in log-sectoral output between two adjacent time periods, t and $t+1$. Combining the resulting expression with the law of motion for TFP, Equation (3), we find that

$$\log y_{cst+1} = \phi_N \log(1 + \Delta N_{cst}) + \phi_A \log y_{cst} + \delta_{ct} + \delta_{st}, \tag{4}$$

where δ_{ct} and δ_{st} denote country-time and sector-time fixed effects and $\phi_A = 1 - \phi_Y$. The focus of our analysis is on the effect of patenting on value added per worker. This effect is

---

[27]Given our multi-sector, multi-country set-up, we find useful to separate the idea production function, Equation (1), which relates the evolution of the stock of knowledge across $cskt$ bins from the law of motion for TFP, Equation (3). Most models in endogenous growth theory do not present these equations separately. To relate our framework to the standard endogenous growth models, consider a one-country, one-sector and one-field of knowledge economy (or alternatively, a multi-country, multi-sector economy without spillovers across sectors and countries). Assume further that $TFP_{ct} = N_{ct}$, $\phi_0 = \phi_Y = 0$, $\beta_N = 1$ and that the idea production function (1) is $I = N_{ct}^{\phi} R_{ct}$ (as discussed in footnote 25). Then, we find that TFP growth is $\frac{N_{ct+1}}{N_{ct}} - 1 = N_{ct}^{\phi} R_{ct}$. For $\phi = 1$, the model generates the first-generation building-on-the-shoulders-of-giants dynamics (Romer, 1990), whereby the growth rate of TFP_{cst} is directly controlled by the number of ideas produced at time t with an elasticity of one. Letting $\phi < 1$ introduces the semi-endogenous growth fishing-out-of-the-same-pond effect so that increasingly more ideas become necessary to sustain constant TFP growth (Jones, 1995).

captured by ϕ_N, which corresponds to the elasticity of value added per worker growth on patenting. Note also that the country-sector fixed effect $\tilde{\delta}_{cs}$ appearing in our specification of the production function drops from Equation (4) because we take the time difference of log-sectoral output. In addition, note that the country-time fixed effect δ_{ct} absorbs the terms corresponding to sectoral capital-labor ratios (under the assumption of competitive markets for capital and labor across sectors). Since the assumption of factor markets being competitive may seem somewhat stringent, we present empirical specifications that also include as direct controls sectoral capital and labor.[28]

5 Empirical Analysis

In this section, we empirically 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. In Section 5.3, we extend our baseline estimation to a longer time horizon – at the expense of losing sectoral variation – where the dependent variable is output per capita.

5.1 Estimating Equations and Identification Strategy

Our baseline regression model closely follows Equation (4) and is specified as follows,

$$\overline{\log y_{cst+n}} = \phi_N \log\left(1 + pat_{cst}\right) + \phi_A \log y_{cst} + \phi_0 X_{cst} + \delta_{ct} + \delta_{st} + \epsilon_{cst}, \tag{5}$$

where $\overline{\log y_{cst+n}}$ is future annual output per worker in period $t + n$; X_{cst} denotes a set of controls for country c, sector s, and time t; δ_{ct} and δ_{st} denote country-time and sector-time fixed effects; and ϵ_{cst} is the error term. The number of ideas in our model framework ΔN_{cst} is proxied by the number of first-in-the-family patents filed in cst. Thus, there is one departure relative to the model presented in the analytical framework. Rather than looking at one period ahead from t, we look at a measure of output per worker n years ahead of period t. In particular, we take the three-year average annual output per worker as our baseline measure (but we also show in the appendix that the results are robust to selecting any of these years in isolation, $n \in \{1, ..., 3\}$). We follow this approach, since it is common in the empirical growth

---

[28]Our framework implies that the lagged level of sectoral output per worker appears on the right-hand-side of Equation (4) with a coefficient $\phi_A = 1 - \phi_Y < 1$. This result follows from the lagged structure of the TFP, Equation (3), and it is not due to a log-linearization result around a steady state. The coefficient on lagged output per worker has been the focus of much of the cross-country growth literature. This coefficient is typically interpreted as proxying for convergence effects in regressions using aggregate data.

literature to smooth out short-term fluctuations in the variable of interest and concentrate on longer-run trends (e.g., Arcand et al., 2015). Moreover, using this three-year average also alleviates the concern that the effect of a new patent may not be (fully) realized one year after its filing year.

The main coefficient of interest of our empirical equation, Equation (5), is the coefficient on patenting, ϕ_N. It relates changes in the number of patents at the country-sector level in a given year to changes in output per worker in the following years, and it corresponds to the elasticity of output per worker growth to patenting. The presence of the fixed effects in Equation (5) follows from our conceptual framework. Intuitively, the inclusion of sector-year dummies controls for the fact that different industries may differently rely on innovation, as well as the fact that this relationship may vary over time. Sector-year dummies allow us to control for the presence of technological waves and other sectoral shocks that are common across all countries. Finally, the inclusion of country-year fixed effects controls, first, for the fact that different countries have different propensities to innovate, and, second, for any business cycles fluctuations at the country level (e.g., a financial crisis).[29]

Our main specification uses value added per worker from the World Input Output Database (WIOD). We also use TFP measures derived from the WIOD as part of our robustness exercises. The data used in our baseline analysis span from 2000 through 2014, and covers 36 countries and 20 sectors (see Appendix A for more details). Figure 5 shows the binscatter plot of the raw correlation between patent activity, $\log(1 + pat_{cst})$, and value added per employment, $\overline{\log(va\_em_{cst+n})}$, 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 at conventional levels.[30]

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 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 the consequence, of higher innovation activity in a given sector. Finally, estimates might be suffering from attenuation bias, due to presence of measurement error, given that patents are an imperfect measure of ideas and innovation.

---

[29]As we showed in Section 4, country-sector fixed effects are differenced out.

[30]Standard errors are clustered at the country and sector level.

Figure 5: Unconditional Correlation between Value Added Per Worker and Number of Patents
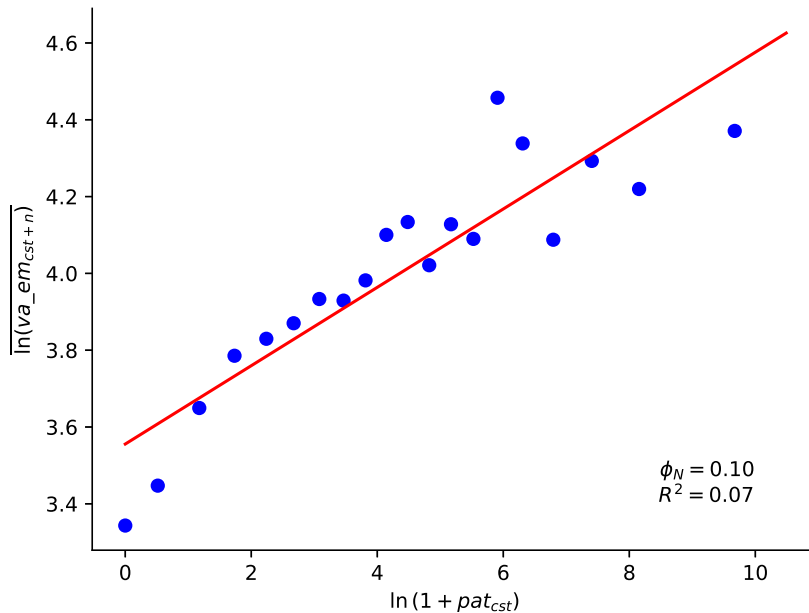
5.1.1 Instrument Construction

To deal with these threats to identification, we build an instrumental variable for patenting activity in a given country and sector. 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. We exploit the pre-determined network of patent citations during the period 1970-90 to identify knowledge links and construct the "share" component of our shift-share instrument. We then construct the "shifts" for the period 1990-2014 using a mix of the observed and predicted number of patents in other countries and sectors starting from the year 1980 on a rolling basis. 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.[31] Thus, our instrument predicts patenting activity in the current period based on 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

---

[31]As we will discuss in detail below, we only use "predicted" patents coming from lagged, pre-2000 patenting data as shifts to generate the instrument for our baseline sample.

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" would only use as sources of variation the own country-sector exposure (shares) and the world patenting activity in a sector (shifts). 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.[32]

Next, we discuss in detail the steps we follow to construct our proposed instrument. To compute the "share" terms of our instrument, we gather patent information on the country of origin, technological field, backward and forward citations 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.[33] The underlying idea is to measure knowledge flows across countries and sectors through the share of citations that each patent produced in the country c_o and sector s_o of origin o gives to patents in country of destination d, c_d, and sector, s_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 and 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 destination country d.

Formally, the entries of the adjacency matrix of the knowledge network for a citation lag Δ are given by,

$$m_{c_o,c_d,s_o,s_d,\Delta} = \frac{\sum_{t=T_0^{share}}^{T_1^{share}} \sum_{p \in \mathcal{P}(c_o,s_o,t)} s_{p\to(c_d,s_d,t-\Delta)}}{\sum_{t=T_0^{share}}^{T_1^{share}} |\mathcal{P}(c_d, s_d, t - \Delta)|}, \tag{6}$$

where $s_{p\to(c_d,s_d,t-\Delta)}$ 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 (c_o, s_o) at time t, and $|\mathcal{P}(\cdot)|$ denotes the total number of patents in the set (i.e., the 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

---

[32]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 to either patent activity or productivity.

[33]We use Eurostat correspondence tables (Van Looy et al., 2014).

country-sector (c_d, s_d) at time $t - \Delta$, and normalize by the patent count in the destination country-sector at time $t - \Delta$. We use the resulting object $m_{c_o,c_d,s_o,s_d,\Delta}$ to construct the "shares" in our shift-share instrument.[34] Note that the "share" terms $m_{c_o,c_d,s_o,s_d,\Delta}$ do not need to add up to one, since their levels capture the number of citations from (c_o, s_o) that are typically received by patents filed in (c_d, s_d) with a lag Δ.

Our network analysis also takes into account the fact 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 $\Delta \in \{1, \cdots, 10\}$. In other words, we allow for the strength of the links to depend on how many years have passed between when the cited and citing patents were filed. In sum, our share terms are allowed to vary by country-sector citing-cited pairs and by time lag between cited and citing patents.

Our shift-share design is based on the idea that it is possible to predict the number of patents in a country and sector of interest based on pre-determined knowledge linkages. Intuitively, this approach mirrors the one of an input-output model for idea production except that it recognizes the non-rival nature of ideas (an idea in one country-sector can potentially spillover to multiple country-sector pairs). To carry out this approach, we then use as shift terms patents filed Δ 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 filed in the country-sector of interest. 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_0 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 rescaling term that ensures that predicted number of patents is equal to the

---

[34] As discussed 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 (2022) show that the network of patents in the United Stated is stable in the time frame they consider, which roughly coincides with ours.

actual number of patents in period t worldwide and $pat_{c_d,s_d,1990-\Delta}$ is the actual number of patents filed in $c_d, s_d, 1990 - \Delta$.[35] Between 1991 and 1999 we construct the predicted number of patents using the previously computed predicted number of patents for years *since* 1990, and the 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 only leveraging the *predicted* patenting computed in the 1990s described above:

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

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. 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). Last but not least, we discard citations coming from the same country and from the same sector when we construct predicted patents. In other words, when calculating the $m_{c_o,c_d,s_o,s_d,\Delta}$ terms in Equation (6), we set the own-country and own-sector terms to 0,

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

We exclude own country and sector to avoid endogeneity concerns arising from the fact that the links that connect the same country or sector might be correlated with future shocks (despite being at least 10 years apart).[36]

Figure 6 visually compares the actual and predicted number of patents through a binscatter plot. The two variables are strongly but not perfectly correlated: the coefficient of the regression is 0.65 and $R^2 = 0.50$. The Kleibergen-Paap Wald F statistics in the benchmark regression is 34, which rules out weak instrument concerns.

To provide evidence in support of our instrument, we report in the next section tests for a

---

[35]Figure C.1 in the appendix shows a simple example of this procedure.

[36]For example, Cai and Li (2019) document the importance of multi-sector firm innovation using US patents, suggesting that some firms are able to internalize knowledge spillovers across sectors.

Figure 6: Unconditional Correlation between Actual and Predicted Patents
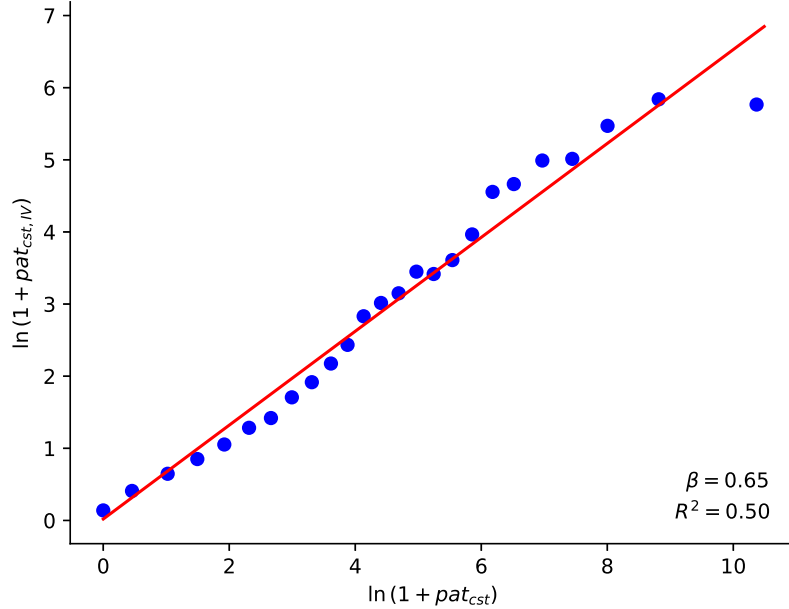
number of assumptions underlying the identification restrictions of shift-share designs, along the lines of Tabellini (2020).[37] First, the validity of the shift-share instrument rests on the assumption that countries and sectors giving more citations (to other sectors and countries) in the period between 1970 and 1990 are not on different trajectories in terms of the evolution of output per worker in the analysis period (2000-2014). We test this assumption in two ways: i) by regressing productivity in 1990 against average patent activity in the period of 2000-14 predicted by the instrument and ii) by checking that results are unchanged when controlling for an average level of patent activity in the period 1970-90.[38]

Second, we rule out the possibility that the links of knowledge diffusion used to construct the instrument capture 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 a 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 in other country-sector pairs. More precisely, we start by constructing the pre-determined network of citations, this time

---

[37]The analysis of the validity of our instrument 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.

[38]We use value added per employment obtained from United Nations Industrial Development Organization (UNIDO) data as a measure of productivity.

using forward citations instead. 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 have been necessary to rationalize the 2000-2014 period. Finally, we include this predicted number of patents in our baseline regression as an additional control. These 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 according to our idea generation empirical model.

5.2 Innovation and Productivity

In this section, we explore the effect of innovation on productivity. As we have just discussed, our identification strategy relies on pre-determined network knowledge linkages. They allow us to predict country- and sector-specific shocks to innovation activity (measured by patent filings) due to knowledge created in other geographical areas and sectors.

Table 1 shows our benchmark estimates of the relationship between value added per employment and innovation instrumented with predicted innovation. As discussed above, we use a three-year average of output per worker to remove short-term business cycle fluctuations.[39] Our benchmark regression uses data from the years 1970-90 to compute pre-determined network linkages, and the period of our analysis is 2000-2014. The first two columns report the estimated results when we only include lagged value added as a control, as well as country-year and sector-year fixed effects. In the third and fourth columns, we add to our empirical model lagged capital and employment as controls, to account for differences in inputs across countries and sectors. We find similar results to the regressions in columns (1) and (2). We use the specification in columns (3) and (4) as our a baseline.[40]

Finally, in the fifth and sixth columns, we exploit the trade linkages given by the world input-output structure and add as controls the value of intermediates imported by each country-sector pair to explore the possibility that foreign imports of intermediates may disproportionately contribute to value added per worker, perhaps because of diffusion of ideas or intangible knowledge (Ayerst et al., 2020). We find no support for this hypothesis: the estimated coefficient on patenting hardly changes relative to our baseline.

---

[39]Our baseline specification $\log(1 + pat)$ allows us to retain the observations with zero patenting. The results are robust to using the inverse hyperbolic sine transformation of the number of patents instead of $\log(1+pat)$. Results for alternative log transformation of patents and forward lags for the dependent variable are reported in Table C.2 in the Appendix.

[40]Results with both lagged and contemporaneous capital and employment as controls are very similar. They are reported in Table C.3 in the Appendix. The fact that the inclusion of these controls does not change the estimated coefficient on patenting is consistent with our conceptual framework – which suggests that, with competitive factor markets, capital labor ratios across sectors are equalized and thus absorbed by the country-time fixed effects.

Table 1: 2SLS Estimates: 2000-2014

	OLS (1)	2SLS (2)	OLS (3)	2SLS (4)	OLS (5)	2SLS (6)
			$\overline{\log(va\_em_{cst+n})}$ $n \in \{1,2,3\}$			
$\log(1+pat_{cst})$	0.006	0.019	0.004	0.017	0.005	0.019
	(0.003)	(0.008)	(0.003)	(0.008)	(0.003)	(0.007)
$\log(va\_em_{cst})$	0.919	0.917	0.942	0.937	0.934	0.928
	(0.012)	(0.012)	(0.016)	(0.016)	(0.016)	(0.015)
$\log(capital_{cst})$			-0.016	-0.014	-0.015	-0.014
			(0.008)	(0.008)	(0.008)	(0.008)
$\log(employ_{cst})$			0.020	0.015	0.012	0.006
			(0.010)	(0.010)	(0.011)	(0.010)
$\log(int\_import_{cst})$					0.009	0.010
					(0.009)	(0.009)
Country-Year FE	Y	Y	Y	Y	Y	Y
Sector-Year FE	Y	Y	Y	Y	Y	Y
# obs.	8,357	8,357	8,357	8,357	8,357	8,357
# countries	36	36	36	36	36	36
			First-stage estimates			
Predicted		0.496		0.461		0.461
$\log(1+pat_{cst})$		(0.082)		(0.079)		(0.079)
F-stat		36.7		33.9		34.3

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 (in parentheses) are two-way clustered at the country and sector levels. Columns (1), (3), and (5) report the results using OLS, and Columns (2), (4), and (6) report the results obtained with 2SLS. Kleibergen-Paap Wald F-stat is reported for the first stage.

The coefficient on innovation activity is positive, and statistically significant across the board. The magnitude of the two-stage least squares regressions is also stable across specifications. The coefficient in column (4) suggests that a 1% increase in patenting leads to 0.017% increase in value added per employment. Using the structure of our simple framework, we can rewrite the estimating equation by subtracting the current level of log value added per worker to also conclude that the estimated elasticity implies that a 1% increase in patenting leads to a 0.017% increase in the growth of value added per worker. This estimated elasticity implies that a one residual standard deviation increase in log patenting generates an increase in value added per employment growth of 1.1 percentage points. This change in valued added growth represents 7.8% of the standard deviation in output per worker growth

in our sample.[41] To have a sense of magnitudes, 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 from the level of innovation observed in Australia or France to the level observed in the US in 2000.

If we used interquartile range changes to quantify our results instead of standard deviation changes, we would obtain similar results. A one interquartile range increase in the log of the number of patents implies an increase of 10.4% of the interquartile range in value added per employment growth. 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 3.1% and 2.9%, respectively.

The estimated 2SLS coefficients are larger than the ones obtained with 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 downward bias could be an increase in market concentration—a trend observed in most advanced countries since the 2000s. In particular, Akcigit and Ates (2021) and Olmstead-Rumsey (2019) have argued that higher market concentration leads to a slowdown in aggregate productivity growth while stimulating the innovation activity of market leaders to maintain their technological advantage.

First-Stage Estimates and Knowledge Spillovers Before turning to the 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 about 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 one standard deviation increase in predicted patents outside country-sector (c, s) implies a 0.46 increase in actual patenting in country sector (c, s) in a sample period.[42]

---

[41]Note that these results are calculated using residual standard deviations, that is, standard deviations obtained after partialling out the full set of controls in column (4). Without doing that, we would obtain larger effects. In fact, a one standard deviation increase in log patents implies an increase in log value added per employment (or value added per employment growth) of 4.4 percentage points.

[42]We residualize all variables with all regression controls before computing the standard deviations. An analogous exercise without partialling out the controls would imply a 0.43 standard deviation increase.

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

	$\overline{\Delta \log(y_{cst+n})}$ $n \in \{1,2,3\}$								
	VA/EMP			Primal TFP			Dual TFP		
	2SLS (1)	2SLS (2)	2SLS (3)	2SLS (4)	2SLS (5)	2SLS (6)	2SLS (7)	2SLS (8)	2SLS (9)
$\log(1 + patent_{cst})$	0.011	0.009	0.009	0.007	0.010	0.010	0.004	0.008	0.008
	(0.004)	(0.004)	(0.003)	(0.004)	(0.006)	(0.005)	(0.004)	(0.003)	(0.003)
$\log(y_{cst})$	-0.044	-0.031	-0.033	-0.017	-0.010	-0.011	-0.018	-0.009	-0.009
	(0.009)	(0.007)	(0.007)	(0.010)	(0.009)	(0.008)	(0.009)	(0.005)	(0.005)
$\log(capital_{cst})$		-0.005	-0.005		-0.023	-0.022		-0.031	-0.031
		(0.003)	(0.003)		(0.003)	(0.003)		(0.003)	(0.004)
$\log(employ_{cst})$		0.005	0.002		0.021	0.022		0.026	0.023
		(0.004)	(0.005)		(0.004)	(0.004)		(0.004)	(0.003)
$\log(int\_import_{cst})$			0.003			-0.002			0.001
			(0.004)			(0.005)			(0.003)
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
$\log(1 + pat_t)$	(0.085)	(0.079)	(0.079)	(0.081)	(0.080)	(0.080)	(0.085)	(0.083)	(0.083)
F-stat	30.5	33.9	34.3	34.5	32.5	32.4	38.1	35.0	34.9

Notes: Period of the analysis is 2000-14 using the pre-determined matrix based on the data from 1970-90. First-stage estimates include all controls. Standard errors (in parentheses) are two-way clustered at the country and sector levels. y_{cst} is a respective measure of productivity (in columns (1)-(3) y_{cst} is value added per employment, and in columns (4)-(9) y_{cst} stands for TFP measured using either the primal or dual approach). In the case of primal TFP for our baseline specification (Column (5)), the main coefficient of interest is significant at the 10% level with p=0.09. Kleibergen-Paap Wald F-stat is reported for the first stage.

Alternative Growth Specification and TFP Regressions To assess the robustness of our findings, we extend our analysis to using TFP growth instead of output per worker as our dependent variable.[43] Table 2 shows our estimates for two measures of TFP growth, as well as value added per employment growth (rather than in levels, as in our baseline specification). The coefficient on innovation activity is positive, statistically significant across different measures, and quantitatively consistent with our baseline results.[44] Moreover, when comparing the coefficient on patenting, ϕ_N, across different specifications, e.g., columns (3), (6), and (9), we see that, as implied by our simple framework, its magnitude is similar regardless of whether we use value added or TFP as the dependent variable.[45]

---

[43] 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).

[44] As in our baseline specification, the results reported in Table 2 are robust to using the inverse hyperbolic sine transformation of the number of patents instead of $\log(1 + pats)$ and adding forward lags as controls. See Tables C.5 and C.4 in the Appendix.

[45] We also find similar results to our baseline ϕ_N when estimating Equation (5) assuming $\phi_A = 1$ (and, thus, having the growth rate as a dependent variable). These are reported in Table C.1 in the Appendix.

Table 3: Checking for Pre-trends

| | $\log(\overline{va\_emp_{cs}})$ | | | |
| | Sample Period | | Pre-Sample Period | |
	(1)	(2)	(3)	(4)
$\log(\overline{1 + pat_{cs2000-14}})$	0.080	0.102	0.032	0.014
	(0.033)	(0.046)	(0.064)	(0.053)
Controls	✓	✓	✓	✓
Country FE	Y	Y	Y	Y
Sector FE	Y	Y	Y	Y
# obs.	641	433	433	424

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 the dependent variable is the 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 the period 2000-14. Standard errors (in parentheses) are two-way clustered at a country and sector levels.

5.2.1 Robustness Checks

As discussed above, the validity of our shift-share design rests on country-sector pairs that give more citations pre-1990 not being on different trajectories in the terms of output per worker post-2000. This assumption is violated if the characteristics of countries and sectors that give more citations to particular countries and sectors in the period 1970-90 had persistent effects on patenting activity, as well as on changes in the outcomes of interest, and these are not captured by our 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 the average value of productivity during the pre-sample period against the average annual number of patents in the period 2000-14.[46] The coefficients of this regression, reported in Columns (3) and (4), are not significantly different from zero, while the estimates obtained for the period used in the main exercises, reported in Columns (1) and (2), are indeed significant.

Second, in Column (2) of Table 4, we check that our results hold when controlling for the average level of patenting activity in the period 1970-90. The results remain virtually unchanged. The coefficient of interest becomes larger in magnitude (in absolute value), but it is statistically indistinguishable from the baseline level because the standard error also

---

[46] As a measure of productivity we use value added per employment data from UNIDO database, since 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.

increases.

Next, we present evidence consistent with ruling 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 as a control the number of patents that should have been filed in the pre-sample period to explain the actual number of patents observed in the sample in the period of study given the citations linkages in the pre-sample.[47] The results presented in Column (3) of Table 4 are stable. The coefficient of interest remains statistically significant and quantitatively close to the baseline. Column (4) includes both controls simultaneously, i.e., the historical patent activity and the demand-driven number of patents in the past in the baseline regression. The coefficient remains significant and has a similar magnitude.

Finally, to check for whether some outliers are driving our results, we repeat our baseline regression excluding one country or sector at a time. We find that our results remain stable and are essentially unchanged across all these regressions.[48]

5.3 Innovation and Long-term Development

Our baseline analysis studied value added per worker after the year 2000. This section extends our analysis to a longer time frame. 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 are not readily available. To circumvent this problem, we adapt our empirical strategy to study the relationship between innovation activity and GDP per capita at the aggregate country level since 1980 (and later extend it back to 1960), using real GDP per capita data from the Maddison Project Database (Inklaar et al., 2018). We therefore depart from our baseline exercise along two dimensions. First, we abstract from sectoral variation both when we construct our instrument and when we conduct the regression analysis. Second, we use 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 patent data for the period prior to 1970 mostly covers advanced economies and given that for most developing countries we observe little to

---

[47] We construct this variable by using the "reverse" matrix procedure described in the end of Section 5.1. To deal with the time dimension of data, we include in the regression the predicted number of patents that should have been filed 30 years in the past. The results hold for other choices of lags.

[48] The largest change in magnitude that we obtain in ϕ_N is when we exclude the sector called 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. These results are available upon request.

Table 4: 2SLS Estimates: Robustness

| | $\overline{\log(va\_em_{cst+n})}$ $n \in \{1,2,3\}$ | | | |
	(1)	(2)	(3)	(4)
$\log(1 + pat_{cst})$	0.017	0.029	0.025	0.030
	(0.008)	(0.010)	(0.010)	(0.010)
$\log(1 + \overline{pat_{cs1970-90}})$		-0.009		-0.009
		(0.005)		(0.005)
$\log(1 + \widehat{pat}_{cst-30})$			-0.006	-0.001
			(0.007)	(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
$\log(1 + pat_{cst})$	(0.079)	(0.058)	(0.065)	(0.056)
F-stat	33.9	20.9	35.5	29.3

Notes: Column (1) shows the results of our baseline regression. Column (2) and (3) show the regression results when including separately the historical levels of average patent activity and the predicted number of patents driven by demand pull factors, respectively; and Column (4) shows the regression results when including them together. All regressions include (log) values for value added per employment, capital, and employment as controls. Standard errors (in parentheses) are two-way clustered at the country and sector levels. Kleibergen-Paap Wald F-stat is reported for the first stage.

no innovation activity measured in terms of patents prior to 1970, our shift-share design may miss a part of the variation we are interested in capturing. For these reasons, we choose the years 1980-2016 as our baseline time period, while we use the pre-1980 data to construct our instrument (so that we include the 1970s, for which there are data on a substantial number of patents for middle-income economies). The set of countries that we consider are the ones categorized as high-income and upper-middle-income countries according to the World Bank classification,[49] for which we have substantial variation in patenting activity.[50]

---

[49]https://datahelpdesk.worldbank.org/knowledgebase/articles/906519-world-bank-country-and-lending-groups

[50]Our patent data cannot distinguish between zero patenting activity and missing data in a given country, sector, and year. Throughout our analysis, we assume that no records of patenting activity are treated as zero patents. Under this assumption, the average annual number of patents in the period 1960-80 is 21,264 and 1,227 patents for high-income and upper-middle-income countries, respectively. At the same time, the average number of annual patents for the same period for lower-middle-income and low-income countries is 45 patents and 1 patent, respectively. Given the little variation in patenting activity for the historical

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 matrix is computed as

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

and we thus abstract from sectoral variation.[51] We use the citation data observed in the period prior 1980 to construct the pre-existing linkages across countries, along with countries' patenting activity during the period starting in 1970, as shifts to construct our instrument for patenting activity during the period 1980-2016.[52]

The empirical specification we run corresponds to Equation (4) in our motivating framework (without sectoral variation). As a reminder, it is obtained from combining a Cobb-Douglas aggregate production function and our law of motion for TFP. The following specification is used in the analysis:

$$\overline{\log(gdp\_cap_{ct+n})} = \phi_N \log(1 + pat_{ct}) + \phi_A \log(gdp\_cap_{ct}) + \delta_t + \delta_c + \varepsilon_{ct}, \qquad (7)$$

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 our results. As in the previous section, the 2SLS estimates reported in columns (2) and (4) imply a higher elasticity of patenting on income compared with 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 sectoral output per worker that we find for the period 2000-2014. The

---

time period in less developed economies, we focus on high-income and upper-middle-income countries for our long-term analysis. We report a number of robustness checks at the end of the section.

[51] As a robustness check, we also compute our shift-share instrument using cross country and sector variation and then aggregate 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.

[52] Similar to our baseline instrument, we use a mix of actual and predicted patents as shifts. We also do not take into account domestic spillovers when constructing the instrument, i.e., $m_{c_o,c_d} = 0$, when $o = d$ and consider citation lags $\Delta \in \{1, \cdots, 10\}$. However, we no longer have the intermediate 10-year period between the pre-determined matrix and instrument as in our baseline. This is to ensure both that we have a sufficiently long sample size of GDP growth rates and that we include patenting activity of the 1970s to construct our shift-share. We also performed as a robustness check analysis where we use all citations available before 1960/70 to construct the pre-determined matrix of citation linkages, along with 1970/80-2016 as a period for the regression analysis, obtaining similar results.

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

| | Dependent Variable is: $\overline{\log(gdp\_cap_{ct+n})}$ | | | |
| | OLS | 2SLS | OLS | 2SLS |
	(1)	(2)	(3)	(4)
$\log(1 + pat_{ct})$	0.013	0.086	0.005	0.034
	(0.004)	(0.021)	(0.003)	(0.012)
$\log(gdp\_cap_{ct})$	0.906	0.735	0.852	0.804
	(0.026)	(0.052)	(0.025)	(0.028)
Country FE	Y	Y	Y	Y
Year FE	N	N	Y	Y
# obs.	1,985	1,985	1,985	1,985
# countries	60	60	60	60
	First-stage estimates			
Predicted		0.771		1.884
$\log(1 + pat_{ct})$		(0.199)		(0.695)
F-stat		15.0		7.3

Notes: Period of the analysis is 1980-2016 using pre-determined matrix based on the data for the pre-1980 period. Standard errors (in parentheses) are clustered at the country level. 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 the number of patents, the regressions in columns (3) and (4) also include year fixed effects. Kleibergen-Paap Wald F-stat is reported for the first stage.

elasticity of patenting to income per capita is 0.034.[53] Quantitatively, this elasticity implies that one standard deviation increase in the logarithm of the annual number of patents leads to 0.41 standard deviations increase in the logarithm of annual GDP per capita, implying an increase of 2.8 percentage points in the growth of GDP per capita.

Income per capita growth over longer horizons. We extend the period of analysis to longer time horizons. Columns (1)-(4) in Table C.7 in the Appendix report the results of running the same specification, Equation (7), using income per capita data spanning the periods 1960-2016 and 1970-2016. In each case, we construct our shift-share instrument in an analogous way to what we have done so far in this section, but now with patenting data pre-1960 or pre-1970, respectively. In both cases, we find a positive and significant first stage, despite our network of innovation being more sparse. We estimate a positive and significant

---

[53]If we run our regression for all countries in our sample rather than only middle and upper income countries, we find an almost identical coefficient of 0.31. However, the first stage is weak and the estimated coefficient is not significant at conventional levels. See columns (5) and (6) of Table C.7 in the Appendix.

effect of innovation on income per capita growth in both regressions. The implied magnitudes suggest that that one standard deviation increase in the logarithm of the annual number of patents generates an increase of 1.64 and 2.15 percentage points in GDP per capita growth for the periods 1960-2016 and 1970-2016, respectively.

6   Conclusion

In this paper, we use a panel of historical patent data spanning the past hundred years and a large range of countries to study the evolution of innovation across time and space and its effect on productivity. In the first part of the paper, we propose a clustering algorithm to classify finely defined patent classes into fields of knowledge based on inventors' patent activity. We then turn to documenting some salient facts of patenting activity since the beginning of the 20th century. We document broad technological waves over the 20th century and in the early decades of the 21st century, and the heterogeneous contribution of countries to these waves. We also document a substantial rise of international knowledge spillovers, as measured by patent citations since the 1990s. This rise is mostly accounted by an increase in citations to US and Japanese patents in fields of knowledge related to computation, information processing, and medicine.

After documenting these facts, we propose a shift-share approach that leverages the directed network of knowledge spillovers across fields of knowledge and countries (to construct the shift terms) and the heterogeneity in exposure of countries to technological waves (to construct the share terms). We then utilize our proposed instrument to estimate the effect of innovation on output per worker and TFP growth in a panel of country-sectors over the period 2000-2014, with our instrument using historical patent data spanning the years 1970 through 2000. We find that, on average, an increase of one standard deviation in patenting implies a 1.1 percentage point increase of output per worker growth.

Finally, we estimate the effect of innovation on long-run income per capita growth and find a positive effect, similar in magnitude to our baseline results. We believe that our shift-share design can be applied to other settings in which the effect of innovation or productivity are of interest. For example, our empirical strategy can be employed in a multi-sectoral Ricardian trade model as in Costinot et al. (2012) to estimate the elasticity of trade flows to productivity differences.

References

ACEMOGLU, D. (2009a): *Introduction to modern economic growth*, Princeton, NJ [u.a.]: Princeton Univ. Press.

——— (2009b): *Introduction to Modern Economic Growth*, Princeton University Press.

ACEMOGLU, D., U. AKCIGIT, AND W. KERR (2015): "Networks and the Macroeconomy: An Empirical Exploration," in *NBER Macroeconomics Annual 2015, Volume 30*, National Bureau of Economic Research, Inc, 273–335.

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. ATES (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.

AYERST, S., F. IBRAHIM, G. MACKENZIE, AND S. RACHAPALLI (2020): "Trade and Diffusion of Embodied Technology: An Empirical Analysis," .

BAQAEE, D. AND E. FARHI (2019): "The Macroeconomic Impact of Microeconomic Shocks: Beyond Hulten's Theorem," *Econometrica*, 87, 1155–1203.

BERKES, E. AND R. GAETANI (2022): "Income Segregation and the Rise of the Knowledge Economy," *American Economic Journal: Applied*, forthcoming.

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.

BUERA, F. AND E. OBERFIELD (2020): "The Global Diffusion of Ideas," *Econometrica*, 88, 83–114.

CAI, J. AND N. LI (2019): "Growth Through Inter-sectoral Knowledge Linkages," *Review of Economic Studies*, 86, 1827–1866.

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 M. GERTLER (2006): "Medium-Term Business Cycles," *American Economic Review*, 96, 523–551.

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 (2014): "Technology Diffusion: Measurement, Causes, and Consequences," in *Handbook of Economic Growth*, ed. by P. Aghion and S. Durlauf, Elsevier, vol. 2 of *Handbook of Economic Growth*, chap. 2, 565–622.

———— (2018): "If Technology Has Arrived Everywhere, Why Has Income Diverged?" *American Economic Journal: Macroeconomics*, 10, 137–178.

COSTINOT, A., D. DONALDSON, AND I. KOMUNJER (2012): "What Goods Do Countries Trade? A Quantitative Exploration of Ricardo's Ideas," *Review of Economic Studies*, 79, 581–608.

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.

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.

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.

KLEINMAN, B., E. LIU, AND S. REDDING (2021): "Dynamic spatial general equilibrium," Lse research online documents on economics, London School of Economics and Political Science, LSE Library.

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.

KOGAN, L., D. PAPANIKOLAOU, A. SERU, AND N. STOFFMAN (2017): "Technological Innovation, Resource Allocation, and Growth," *The Quarterly Journal of Economics*, 132, 665–712.

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.

LIU, E. (2019): "Industrial Policies in Production Networks," *The Quarterly Journal of Economics*, 134, 1883–1948.

LIU, E. AND S. MA (2021): "Innovation Networks and Innovation Policy," NBER Working Papers 29607, National Bureau of Economic Research, Inc.

MELITZ, M. AND S. REDDING (2021): "Trade and Innovation," CEPR Discussion Papers 16264, C.E.P.R. Discussion Papers.

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.

OBERFIELD, E. (2018): "A Theory of Input–Output Architecture," *Econometrica*, 86, 559–589.

OLMSTEAD-RUMSEY, J. (2019): "Market Concentration and the Productivity Slowdown," MPRA Paper 93260, University Library of Munich, Germany.

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.

PERLA, J. AND C. TONETTI (2014): "Equilibrium Imitation and Growth," *Journal of Political Economy*, 122, pp. 52–76.

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.

TIMMER, M., E. DIETZENBACHER, B. LOS, R. STEHRER, AND G. DE VRIES (2015): "An Illustrated User Guide to the World Input–Output Database: the Case of Global Automotive Production," *Review of International Economics*, 23, 575–605.

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)