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

Enrico Berkes
UMBC

Kristina Manysheva
Columbia University

Martí Mestieri
FRB of Chicago,
UPF, BSE and CREi

September 4, 2024

Abstract

We use a panel of patent data that covers a large range of countries over the past century to study the evolution of innovation and knowledge spillovers across time and space, and the effect of innovation on productivity. We document a substantial rise in international knowledge spillovers since the 1990s. This rise is mostly driven by an increase in citations of US and Japanese patents related to computation, information processing, and medicine. Leveraging the network of patent citations, we provide credible estimates of the elasticity of innovation to international knowledge spillovers. Using this as our first stage, we estimate the causal effect of innovation induced by international spillovers on productivity, and aggregate income per capita. An increase of one standard deviation in log patenting increases sectoral output per worker growth by 1.1 percentage points. We find results of similar magnitude for sectoral TFP growth and income per capita growth.

Keywords: Innovation, Technology Diffusion, Patents.
JEL Classification: O10, O30, O33, O47.

---

*We are grateful to Ufuk Akcigit and Bart Hobijn for the valuable discussions of our paper. We also thank Isaac Baley, Joel David, Matthias Doepke, Ruben Gaetani, Giammario Impullitti, Ben Jones, Nan Li, Joel Mokyr, Dimitris Papanikolau, Sergio Petralia, Thomas Sampson, and seminar attendees at CREI, LSE, UCSD-UCLA-UCB trade conference, Midwest Macro, Nottingham, SED, UB, UCSD, Junior Growth Conference (LBS), Junior Macro Workshop (Philadelphia Fed), NBER SI Growth Meeting, and Northwestern for helpful comments and discussions. The views expressed are the authors' and do not necessarily reflect the views of the Federal Reserve Bank of Chicago or the Federal Reserve System. Any remaining errors are our own. E-mails: eberkes@umbc.edu, km3924@columbia.edu, mestieri.marti@gmail.com.

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 scarce.[1]

In this paper, we use historical patent data spanning a wide range of countries over the past 100 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 (patents count) and knowledge spillovers—i.e., how innovation builds on previous knowledge (patent citations). We document a substantial rise in international knowledge spillovers since the 1990s, mostly driven by the United States and Japan, as well as the rise of importance 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) for two purposes. First, we provide a credible estimate of the elasticity of knowledge creation to international knowledge spillovers. More concretely, exploiting the network of patent citations, we estimate the patenting activity in any given country and field of knowledge induced by ideas created in other countries and fields of knowledge. Second, taking this previous exercise as the first stage of a two-stage least squares regression, we propose a new identification strategy to quantify the effects of innovation induced by knowledge spillovers on productivity and economic growth across countries and industries. To our knowledge, our identification strategy is novel to the endogenous growth literature.

We build our measure of innovation using patent data collected from the Euro-

[1]See Comin and Mestieri (2014) and references therein for an overview of 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 restricted to 25 major technologies since 1780.

pean Patent Office Worldwide Patent Statistical Database (PATSTAT). PATSTAT contains bibliographic and legal information on more than 110 million patent records from the patent offices worldwide and covers both leading industrialized countries and developing countries. To avoid some of the arbitrariness of using broad patent technology classes (Keller, 2002), we classify patents into fields of knowledge obtained through 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 each pair of technology classes. It then uses these probabilities to infer the proximity of technology classes in the knowledge space and to create knowledge clusters.[2]

Armed with our newly defined fields of knowledge, 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 100 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 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, which we measure through patent citations across fields of knowledge and countries. We focus on the post-1970 sample for this analysis, for which we have data on virtually all countries in the world. We show that for the average patent, citations tend to be biased toward domestic 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. On average, more recent patents tend to cite a larger number of patents. This finding aligns with an increase in the burden of knowledge (Jones, 2009).

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, international patents are cited more than twice as much as domestic patents. This

[2]As a robustness check, we also perform a clustering analysis in which the strength of the linkages between different patent classes is based on citations.

finding suggests that 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 citations. The increase is mainly driven by a handful of fields of knowledge that are related to information and communication technologies (ICTs) and medicine. This 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 leverage the different margins of innovation that we have documented to investigate (i) the effect of international knowledge spillovers on innovation, measured by patenting, and (ii) the effect of innovation driven by international knowledge spillovers 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 identification strategy exploits variation in innovation across time and space, as well as the increasing role of international knowledge spillovers documented in the first part of the paper. In the baseline regression, we study the effect of induced innovation 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 and 1960-2016).

Simply correlating innovation and productivity or output per worker is problematic because of measurement error (which would generate attenuation bias), potential reverse causality, and the presence of unobserved factors that affect patenting and the dependent variables simultaneously. 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 corresponds to the spillover function implied by the Kortum (1997) model in a multi-sector-country setting (Comin et al., 2019). It leverages pre-existing knowledge linkages across countries and technologies, measured by the probability of patent citations, to construct the share component of our instrument. The shift component is obtained using lagged patents in other fields of knowledge and countries.

The first stage of our empirical design provides credible estimates of the elasticity of domestic patenting activity to international knowledge spillovers across countries and fields of knowledge. Our analysis reveals that, on average, an increase of 1% in international spillovers is associated with an increase of almost 0.5% in domestic patents. Quantitatively, our estimates are within the 0.2–0.6 range reported by Bottazzi and Peri (2007) for the long-run elasticity of innovation to international knowledge spillovers estimated using co-integration techniques.

In our baseline second-stage regression, the main variable of interest is value added per worker by country and sector (measured using the World Input-Output Database) over the 2000-2014 period. We use patent data starting in 1970 to construct our instrument for this exercise. The regression model includes controls that vary at country-sector-time level. We also control for potential contemporaneous spillovers generated by international input-output linkages. We find a robust effect of innovation on value added per employment growth. A 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), which implies 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—as implied by our theoretical framework.

We conduct several 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 pre-period productivity is uncorrelated with subsequent patent activity predicted by the instrument. In the spirit of Acemoglu et al. (2016), we also "reverse" the network of citations we use 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, anticipated demand. Reassuringly, we find no evidence to support this hypothesis. Finally, our result is also robust to include, as an additional control, the predicted number of patents using input-output linkages instead of citations to ensure that we capture knowledge rather than production spillovers.

We conclude by extending our empirical 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. This allows us to include the patenting activity of virtually all high-income and upper-middle-income countries (as defined by the World Bank). An increase of one residual standard deviation in log patenting implies an increase in the growth of income per capita of between 1.6 and 2.8 percentage points. The implied changes in growth rates represent 24% and 41% of a residual standard deviation of income per capita growth, respectively.

Related Literature This paper relates to the vast and rich literature on the link between innovation and productivity that dates, at least, to 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 the positive effects of innovation on output and productivity growth at 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 particularly 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; also Keller, 2004 and Melitz and Redding, 2021 provide excellent surveys). We contribute to this strand of the literature by documenting an increase in international spillovers since the 1990s, using international linkages to build our shift-share design, and, ultimately, quantify the effect of innovation on productivity.

Our paper also 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 a single country and 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. With respect to providing a global view, our work is perhaps closest to that of Bottazzi and Peri (2003, 2007), who use R&D and patent data for European regions and OECD countries, respectively, 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 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 that of Berkes and Gaetani (2022), who construct a shift-share instrument to leverage the network of patent 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 on itself. Both papers focus on the United States.[3]

2 Data

2.1 Data Sources

In this paper, we measure new ideas based on patent data and productivity based on 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 bibliographic and legal information on more than 110 million patents records from the patent offices around the world and covers both leading industrialized countries and developing countries over the period 1782–2018.[4] 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 has some limitations (Kang and Tarasconi, 2016). The main limitation for our purposes is data availability in the earlier years: Data along one or more dimensions are often missing for some countries in the years preceding 1970. We therefore split our sample into

[3] 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.

[4] PATSTAT has become increasingly popular among researchers in economics because 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. (2018b); Philippe Aghion and Melitz (2018); Bloom et al. (2020); and Dechezleprêtre et al. (2020).

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 patent characteristics required by our analysis are available at least since 1920.[5] The second group includes all countries covered by PATSTAT and starts in 1970.[6] Appendix A provides more information on the composition of the sample and patent variables used in the analysis.

We assign each patent to a geographic unit according to the country of residence of its inventor(s). If this is unavailable, we instead use 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.[7] 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 (with the exception of our citation analysis, in which we count all citations of any patent in a family).

Further, we 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 fields at the four-digit level (e.g., A01B)—for a total of 650 classes—and 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 year instead of the year in which a patent was

[5]Note that to compare consistent geographic units over time, when appropriate, we aggregate patents filed in the German Democratic Republic and the Federal Republic of Germany. Similarly, for the Soviet Union, we combine all patents produced by Armenia, Azerbaijan, Belarus, Estonia, Georgia, Kazakhstan, Kyrgyzstan, Latvia, Lithuania, Moldova, Russia, Tajikistan, Turkmenistan, Ukraine, and Uzbekistan.

[6]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 introduction of the new law, the same pattern is not observed in the number of triadic patents, defined as all 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.

[7]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.

granted.[8]

We supplement the patent data with the World Input-Output Database (WIOD, Timmer et al. 2015). This database provides data on the prices and quantities of inputs, outputs, and trade flows and covers 43 countries and the rest of the world as a "44th country" for the period 2000–2014. The data are classified according to the International Standard Classification Revision 4 (ISIC) for a total of 56 sectors. Using the World Input-Output Tables (WIOT) for each set of countries, sectors, and years, we construct trade flows, gross output, intermediate purchases, and value added expressed in US dollars. 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 trade in intermediate and final goods for each country-sector pair.[9] Finally, we use data from the Maddison Project Database (Inklaar et al., 2018) to collect data on historical income per capita growth across countries.

2.2 Construction of Fields of Knowledge

Innovation is the process of creating new knowledge and potentially building on existing knowledge across different fields. To operationalize our goal of measuring innovation waves across time and space, we build on the rich literature that measures innovative activities through patent data. We cluster finely defined patent classes into broader *fields of knowledge*, which, taken together, constitute what we refer to as the world's *technology space*.

We employ a novel approach to clustering 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 $T_{642\times642}$,[10] in which each element (i, j) is the

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

[9] See details in Appendix A.2. In the appendix, we also discuss the UNIDO INDSTAT2 database and Penn World Data Tables. We use these additional databases to collect historical data on sectoral manufacturing output by country, among other variables.

[10] 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.

probability that an inventor patents in IPC subclass i conditional on also having a patent assigned to subclass j.[11] For example, a mechanical engineer who specializes in brakes will most likely produce inventions related to IPCs B60T (Vehicle Brakes or Parts Thereof) and F16D (Clutches, Brakes), which our algorithm correctly bundles together.[12]

Interpreting this matrix as a distance matrix across technology classes, we then use a *k-medoids* clustering algorithm to group the IPC subclasses into fields 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 then rank each result according to the silhouette coefficient. The silhouette coefficient takes into consideration the distance between elements within a cluster, as well as the distance across clusters, while also penalizing the existence of singletons.[13] The optimal number of clusters implied by the silhouette coefficient is $k = 164$.[14]

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 that emerge from the analysis of our patent data. The stylized facts presented in this section lay the foundation for our empirical analysis. First, we document the variation over time in the importance of different fields of knowledge and the heterogeneous countries' contribution to pushing the technological frontier across fields of knowledge. We will use this time variation across countries and sectors to construct the "shift" component of our instrument. Second, we document the importance of international citations,

[11] The diagonal elements of the matrix, $i = j$, are set to one. Note that the so-obtained matrix does not need to be symmetric. For example, 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).

[12] As a robustness check, we construct the proximity matrix based on citation linkages and apply the same procedure. The results are similar to those obtained using an inventor-based proximity matrix (see the Appendix for further details).

[13] More details on the procedure used to construct fields of knowledge are in Appendix A.4.

[14] The results of clustering algorithm with all subclasses assigned to each cluster are available here.

and particularly its evolution over time and heterogeneity across fields of knowledge. We will use this insight to motivate the use of pairwise citations across countries and fields of knowledge as the "share" component of our shift-share instrument.

3.1 The Evolution of Fields of Knowledge across Space and Time

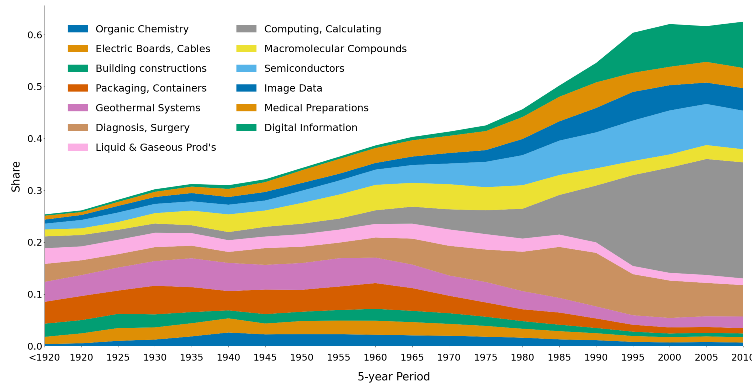
We first document the evolution of the major fields of knowledge for the past 100 years and highlight how different countries contributed to their growth at different points in time. To measure the importance of each field of knowledge at any point, we compute the share of patents that belong to that field of knowledge. Each patent is weighted by the total number of forward citations.[15] We split our dataset into nineteen 5-year periods from 1920 to 2015, plus a period prior to 1920 in which we lump together all patents filed before that year. For each time period, we rank every field of knowledge based on its relative contribution to 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 clear. First, we observe a substantial increase in the concentration of innovation, especially around the 1990s. In the 2000s, approximately 10% of the fields of knowledge account for 60% percent of all patent activity, 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 (e.g., Packaging & Containers; Geothermal Systems) are the most prominent. Starting in the 1950s, we observe a shift toward chemistry and physics (e.g., Macromolecular Compounds). Around the 1980s there was a substantial increase in inventions related to medical and veterinary sciences (e.g., Diagnosis and Surgery or Medical Preparation). Finally, and as expected, around the mid-1990s, fields of knowledge related to computing and communication techniques began to play 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 more

[15]Note that for our analysis we only consider the first patent of the family. If a patent belongs to multiple fields, we add a fraction of the patent to each field proportional to the number of IPC subclasses reported on the patents.

Figure 1: Evolution of Top Fields of Knowledge



Notes: This figure represents the share of each field of knowledge, measured by the number of first-in-the-family patents weighted by 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.

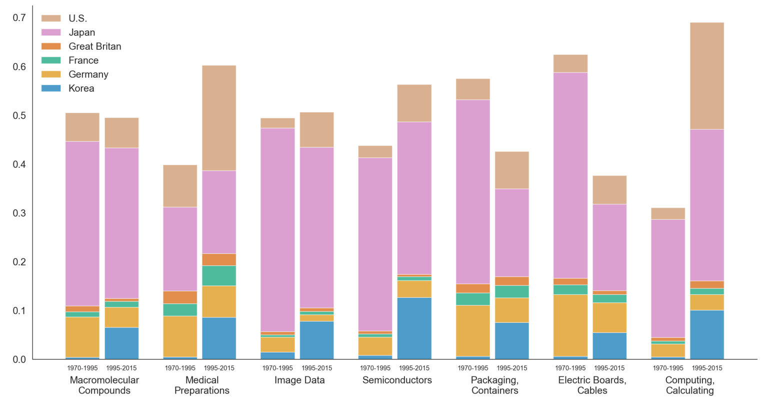
recent decades. To do this, we build importance measures that take into consideration country fixed effects or only patents that were cited at least once. Table B.1 in the Appendix shows that our baseline results are confirmed when using these alternative measures.

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 on the seven fields of knowledge that played a leading role in the innovation space 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.[16]

Because of data limitations, for the first two periods (1920–1944 and 1945–1969), our sample comprises six countries: the US, the UK, France, Germany, the Soviet Union, and Switzerland. Figure B.1 in the Appendix shows that during this period, the leading innovating role in the most prominent fields of knowledge was split between the US and Germany, followed by the UK and France. In fact, Germany overtook the US in every prominent field in the period between the end of World War

[16]To account for potential differences in patent citation requirements and practices across countries, in this part of the analysis we do not weight patents by the number of citations for better comparability. We verify that the results are robust to citation-weighted and other measures, as shown in Table B.1. Figure B.2 represents countries' shares only using patents that were internationally cited.

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



II and 1969.

In Figure 2, we consider the post-1970 sample, for which we have data on filed patents for virtually all countries in the world. Between 1970 and 1995, there are three 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 during these years. After 1995 other Asian countries, such as South Korea, begin moving to the forefront of the technological frontier. In this period, France experiences a decrease in importance in the innovation landscape. Asian countries are particularly innovative in fields related to computing, engineering, and digital information, while their role in chemistry and medicine is less pronounced.

Next, 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 that of the US in the 1970s and subsequently declines, while Asian countries such as Taiwan gain prominence after the 2000s. See Section B in the Appendix for further details and discussion of this exercise.[17]

[17]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, geographic distance, and production of technologies. We find that changes in patenting shares across fields of knowledge are correlated across countries that are geographically and linguistically close to each other.

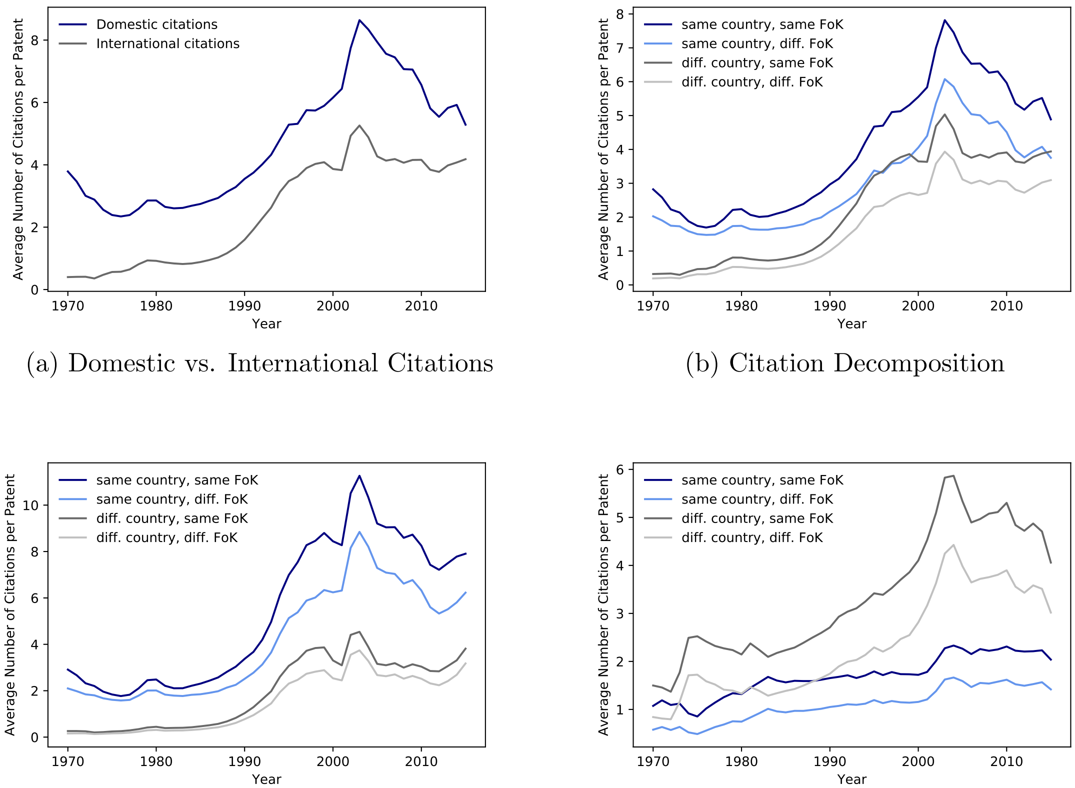
3.2 Using Citations to Measure Spillovers across Time and Space

So far, we have shown that there is substantial time variation in terms of the composition of the technological output and in terms of the geographic contribution to worldwide innovation. We now turn our attention to knowledge spillovers. We measure spillovers through patent citations across fields of knowledge and countries. An abundant literature studies 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 scarcer. 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 again focus our analysis on the post-1970 sample and collect data on all citations of patents filed after 1900. Panel (a) in Figure 3 shows the evolution of the average number of citations, which experienced an important increase starting around the 1980s. Domestic citations continue to increase until 2002 and then they show a marked decline. On the other hand, 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 of international patents. Panel (b) further breaks down these trends by fields of knowledge (FoK) of the citing and cited patents.[18] The plot shows that citations tend to be concentrated not only geographically (i.e., domestic patents are cited relatively more), but also technologically (i.e., patents in the same field of knowledge are cited relatively more). Moreover, these gaps appear to have widened in recent years.

As shown in Figure 2, most knowledge (as measured by patent filings) is produced by a handful of countries—what we refer to as the "technological leaders." We now want to see whether this leadership role also translates into a larger influence in the terms of international knowledge spillovers. Panels (c) and (d) of Figure 3 separately depict citation dynamics for Japan and the United States, 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

[18]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.



(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

Japan, on average, tend to receive more citations per patent than the rest of the world. Second, most of the citations in the US and Japan are of domestic patents, while the rest of the world mostly relies on knowledge produced in other countries.[19]

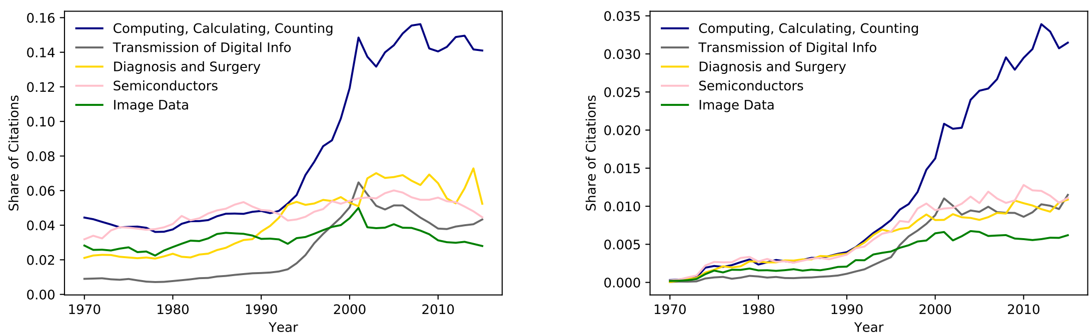
Figure 3 depicts a rapid increase in the overall average number of citations per patent. To better understand what lies behind this increase, we concentrate on the backward citations received by the five leading fields of knowledge over the past five decades. Figure 4 shows that the substantial increase 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 these fields of knowledge

[19]Decomposition of citations for other countries Germany, France, and the UK are reported in Figure B.3 in the Appendix. 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.

Figure 4: Citations shares to patents from US and Japan

(a) US and Japan

(b) All except US and Japan



Notes: Share of citations of US and Japanese patents by FoK, 1970-2015. Each line in the plots represents the share of citations of 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.

are of US and Japanese patents.[20]

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 prevalent, international knowledge spillovers have steadily been gaining importance over the past few decades. This increase is mainly driven by a dramatic increase in the citations received by US and Japanese patents, especially in fields of knowledge related to computing, information processing, and medicine.

The next section presents a conceptual framework that can be used to analyze the patenting patterns documented so far through the lens of endogenous growth theory. Section 5 builds on these stylized facts and the theoretical framework, and proposes a shift-share instrumental variable strategy that allows us to estimate the effect of innovation induced by knowledge spillovers across countries and sectors on productivity. The construction of our proposed instrument leverages the heterogeneity in the reliance on foreign knowledge, as measured by international patent citations, and

[20] 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.

the time variation in innovation, as measured by patent filings. Moreover, the first stage of our regression, interpreted through the lens of our framework, corresponds to the elasticity of innovation on international knowledge spillovers.

4 Conceptual Framework

In this section, we present the framework that will guide our empirical analysis. The framework incorporates patents and patent citations in a standard, multi-sector growth model.[21] Importantly, our framework only specifies the production side of the economy, and does not assume the existence of a balanced growth path of output or productivity at sectoral (or aggregate) level.[22]

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 $S_{csk}(\mathbf{N}_t)$ 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

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

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. Besides being a linear approximation, Equation (2) corresponds

[21]Our formulation builds on previous studies that have examined the patent network of citations, such as Acemoglu et al. (2016). Relative to Acemoglu et al. (2016), we introduce additional model elements to relate our results to TFP and output per capita. We also extend the model to a multi-country setting.

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

to the spillover function of the multi-sector extension of Kortum (1997) developed in Comin et al. (2019). Under this particular micro-foundation, $\alpha_{c's'k't}$ captures the inverse of the bilateral "resistance" in knowledge diffusion between $c's'k'$ and csk at time t. We leverage the structure of this equation and this particular micro-foundation when we construct our instrumental variable. Note that we purposely state Equation (1) generically so that it subsumes 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).[23]

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.[24] 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 countries 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 that 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.

In our framework, there is a representative firm in each country-sector that produces sectoral output that combines 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 term can be parameterized as a full set of dyadic fixed effects, $\phi_{cst} = \tilde{\delta}_{ct} + \tilde{\delta}_{st} + \tilde{\delta}_{cs}$. This parameterization captures the fact that the productivity of ideas (and/or other sources of productivity

[23]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 $\phi \leq 1$.

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

differences not in the model) may differ across country-sector-time pairs because (i) some country-sector pairs may be better in certain sectors than others (captured by $\tilde{\delta}_{cs}$); (ii) some global technology trends may affect certain sectors (captured by $\tilde{\delta}_{st}$); (iii) 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, 2009). That is, we assume there is a positive relationship between ideas produced and sectoral TFP growth. Moreover, as TFP grows and new production methods are adopted, we allow for the existence of adjustment costs that scale up with (a power function of) total output. Adjustment costs capture 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 isoelastic relationship between TFP growth, ideas, and adjustment costs,

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

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. By combining the idea production function, Equation (1), with 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.[25]

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

[25]Given our multi-sector, multi-country set-up, we find it 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 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). Suppose 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 23). Then, we find that TFP growth is $\frac{N_{ct+1}}{N_{ct}} - 1 = N_{ct}^{\phi} R_{ct}$. For $\phi = 1$, the model corresponds to 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).

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

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 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 $\tilde{\delta}_{cs}$ 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 that correspond to sectoral capital-labor ratios (under the assumption of competitive markets for capital and labor across sectors). Since the assumption of competitive factor markets may seem somewhat stringent, we present empirical specifications that also include as direct controls sectoral capital and labor.[26]

5 Empirical Analysis

In this section, we empirically study the effect of innovation on productivity. We begin by 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—in which the dependent variable is output per capita.

[26]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 that use aggregate data.

5.1 Estimating Equations, Identification Strategy, and Data Construction

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 the average annual output per worker between period $t+1$ and $t+n$; 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. We take the average annual output per worker between $t+1$ and $t+3$ as our baseline measure.[27] We follow this approach to smooth out short-term fluctuations in the variable of interest and concentrate on longer-run trends, as is common in the empirical growth literature; e.g., Arcand et al. (2015).

The main parameter of interest is the coefficient on patenting, ϕ_N. It captures how changes in the number of patents at country-sector level in a given year translate into changes in output per worker in the following years. More precisely, it corresponds to the elasticity of output per worker growth to patenting. The presence of 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. 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-cycle fluctuations at country level (e.g., a financial crises).[28]

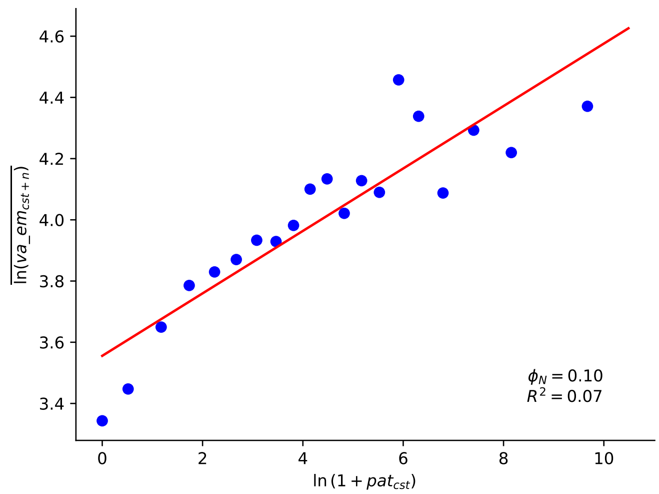
The dependent variable in our main specification is value added per worker measured using the World Input-Output Database (WIOD).[29] The data span the period 2000 through 2014 and cover 36 countries and 20 sectors (see Appendix A for more details). Figure 5 shows the binscatter plot of the raw correlation between patent

[27] We also show in the appendix that the results are robust to selecting any of these years in isolation, $n \in \{1, ..., 3\}$.

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

[29] We also use TFP measures derived from the WIOD as part of our robustness exercises.

Figure 5: Unconditional Correlation between Value Added per Worker and Number of Patents



activity, $\log(1 + pat_{cst})$, and value added per employment, $\overline{\log(va\_em_{cst+n})}$, over our sample period. On average, a 1% increase in the number of patents is associated with a 0.10% increase in future output per worker averaged over the next 3 years.

The relationship depicted in Figure 5, although interesting, cannot be interpreted as causal because a wide range of factors might be affecting innovation activity and productivity at the same time. For example, the obsolescence of some industries might decrease their innovative capacity and productivity at the same time. 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 suffer from attenuation bias due to the presence of measurement error, given that patents are an imperfect measure of ideas and innovation.

5.1.1 Instrument Construction

To address these identification concerns, we build an instrument for patenting activity in a given country and sector. Our instrument is based on the idea that it is possible to predict the number of patents in a country and sector of interest based on existing knowledge linkages. Intuitively, this approach mirrors that of an input-output model for idea production, except that it incorporates the non-rival nature of ideas (i.e., an idea in one country-sector can potentially be used by multiple country-sector pairs).

We rely on a shift-share design that leverages the predetermined network of patent citations from 1970 to 1990 to identify cross-country, cross-sector knowledge links. These will constitute the "share" component of our shift-share instrument. We construct the "shift" component using a mix of the observed and predicted number of patents in other countries and sectors starting from 1980 on a rolling basis. Interacting the shares with the shifts and adding them up, we obtain the predicted number of patents for the period 2000–2014 that we use as our instrument. Our instrument predicts patenting activity in the current period based on knowledge spillovers from other countries and sectors and is a special case of the linear knowledge spillover function presented in Equation (2) in Section 4. More concretely, our mapping of the shifts and shares to patents and citations probabilities, respectively, is motivated by the result in the multisector extension of Kortum (1997) by Comin et al. (2019), who show that changes in knowledge spillovers of any country-sector pair c, s are given by the sum of patenting activities in all country-sector pairs c', s' weighted by the probability of a patent in c', s' being cited by sector pair c, s.[30]

Before delving into the details of the instrument, it is worth emphasizing that our proposed shift-share design differs from a standard Bartik design. The reason is that we exploit the fact that the network of citations is a directed network. This allows us to construct one-directional linkages across country-sector pairs and use shift terms that also vary at country-sector level. By contrast, a standard Bartik variable 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 worldwide industry or technological trends that also affect productivity.

To compute the *shares* of our instrument, we gather patent information on the country of origin, technological field, and 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.[31]

[30] More precisely, a discrete-time version of Comin et al. (2019), implies that the log of spillovers in country-sector cs at time $t+1$, $\ln S_{cs,t+1} = \ln S_{cs,t} + \ln \left(1 + \sum_{c' \in C} \sum_{s' \in S} \pi_{cst \to c's't} \cdot pat_{c's't}, \right)$ where $\pi_{cst \to c's't}$ denotes the probability of a patent in cs citing a patent from $c's'$ at time t, and $pat_{c's't}$ denotes the flow of patents in $c's'$ at time t. One important dimension that Comin et al. (2019) abstract from is the lagged response of innovation to patents, which we incorporate in our empirical analysis by allowing for lags in knowledge diffusion.

[31] We use Eurostat correspondence tables (Van Looy et al., 2014). Patents assigned to multiple sectors are "shared out" as described in Section 2.

The idea is to measure knowledge flows across countries and sectors based on the share of citations that each patent produced in country c_o and sector of origin s_o gives to patents in country and sector of destination, c_d and s_d, respectively. These shares captures the reliance of a certain country and sector pair to ideas generated in other countries and sectors. More precisely, for each patent of sector s_o filed in 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 this 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 elements of the adjacency matrix of the knowledge network are given by

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

where $s_{p \to (c_d,s_d,t-\Delta)}$ denotes the share of citations patent p gives to patents of sector s_d produced in country c_d filed at time $t - \Delta$; $\mathcal{P}\left(s_o, c_o, t\right)$ 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). The resulting objects, $m_{c_o,c_d,s_o,s_d,\Delta}$ capture the reliance of the country and sector of destination, d, on ideas produced in the country and sector of origin, o, and constitute the shares in our shift-share instrument.[32]

Note that our network approach 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 lag $\Delta \in \{1, \cdots, 10\}$. In

[32]As discussed in Section 2, we restrict our sample to patents that are the first in their family. However, we count all cited patents irrespective of their sequence within the family to capture all innovations on which any given patent builds. Berkes and Gaetani (2022) show that the network of patents in the United States is stable over a time frame which roughly coincides with ours. Note that $m_{c_o,c_d,s_o,s_d,\Delta}$ do not need to add up to 1, 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.

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.

For the *shift* terms, we use 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 flows between country-sector dyads is mediated through pre-determined channels across country-sectors which can be proxied by patent citations (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.

Our instrument 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 scaling term that ensures that the predicted number of patents is equal to the actual number of patents in period t worldwide and $pat_{c_d,s_d,1990-\Delta}$ is the actual number of patents filed in $c_d, s_d, 1990 - \Delta$. 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 by only leveraging the *predicted* patenting computed in the 1990s, as 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 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)[33] we discard citations that come 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,

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

To provide evidence in support of our instrument's plausibility, in the next section we conduct several empirical tests for the assumptions that underlie the identification of shift-share designs, along the lines of Tabellini (2020).[34]

5.1.2 First-stage Estimates and Knowledge Spillovers

Before turning to our main empirical specification, we discuss the relationship between the actual and predicted number of patents, which is the first stage of our empirical specification. The relationship between these two variables captures the importance of knowledge spillovers in generating innovation and is therefore an object of interest in its own. Through the lens of our framework, our first stage corresponds to the elasticity of domestic innovation to the stock of relevant knowledge in other country-sector pair. This is an important elasticity to be estimated in economic growth theory. In the context of a closed economy, whether its value is less than 1 or not was the center of the well-known debate regarding endogenous vs. semi-endogenous growth. In a multi-country model, its value is key to understanding technology diffusion and optimal R&D policy.[35]

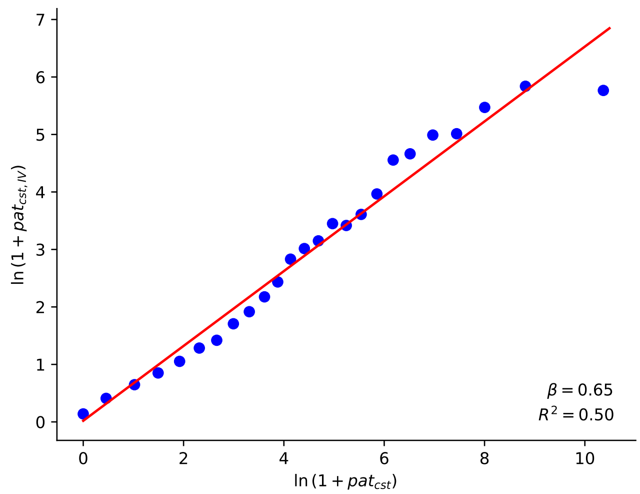
Figure 6 visually compares the actual and predicted number of patents using a binscatter plot. The two variables are strongly correlated: The coefficient of the

[33]For example, Cai and Li (2019) document the importance of multi-sector firm innovation using US patents, which suggests that some firms can internalize knowledge spillovers across sectors.

[34]The analysis of the validity of our instrument falls within the shift-share instrumental variable framework and relies on 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.

[35]For example, whether subsidizing green innovation in one country can offset the lack of innovation in other countries depends crucially on this elasticity (Hémous, 2016). See also Akcigit et al. (2018a) and Sampson (2023) for other applications in which international spillovers play a key role.

Figure 6: Unconditional Correlation between Actual and Predicted Patents



regression is 0.65, and the $R^2 = 0.50$. Table 1 shows that this positive relationship is confirmed when controlling for country-year and sector-year fixed effects, as well as predicted patents obtained through different citation networks. These estimates inform us about the average knowledge spillovers from other country-sector pairs to a given country-sector pair. Column (1) reports the results using the predicted number of patents constructed with the baseline knowledge network described above. In magnitude, a one residual standard deviation increase in the logarithm of predicted patents outside country-sector (c, s) implies an increase of 0.46 residual standard deviations in patenting in (c, s).[36]

As we showed in Section 3, not only are the US and Japan among the top innovating countries in the last decades, but the two countries are also the largest recipients of citations since the mid-1990s. In columns (2) and (3), we investigate whether the first-stage is robust to splitting the knowledge network between the US and Japan, and the rest of the world when constructing our instrument. To this end, we estimate the predicted number of patents in constructing our instrument using citation linkages (i) excluding those from the US and Japan and (ii) including only those coming from the US and Japan. Quantitatively, both estimates imply similar magnitudes of the effect: A one-standard-deviation increase in international (and inter-sectoral)

[36]We residualize all variables with all regression fixed effects before computing the standard deviations.

Table 1: Actual and Predicted Patents

	$\log(1+pat_{cst})$		
	(1)	(2)	(3)
$\log(1+\widehat{pat}_{cst})$	0.452	0.470	0.599
	(0.047)	(0.048)	(0.061)
Country-Year FE	Y	Y	Y
Sector-Year FE	Y	Y	Y
# obs.	31,292	31,292	31,292
# countries	198	198	198

Notes: The period of the analysis is 2000-2014. Standard errors (in parentheses) are clustered at the country-sector level. Column (1) reports results using our original instrument. Column (2) excludes citations to US and Japanese patents when buildling the instrument. Column (3) only includes citations to US and Japanese patents. We report estimates using all countries in our sample. Table C.1 reports the same regression restricting the sample to countries in the second-stage analysis.

patenting generates an increase of around 0.5 standard deviations in local patenting.

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 predetermined 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 geographic areas and sectors.

Table 2 shows our benchmark estimates of the relationship between value added per employment and innovation instrumented with predicted innovation. We use a 3-year average of output per worker to remove short-term business-cycle fluctuations.[37] Our benchmark regression uses data from the years 1970-90 to compute predetermined 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, our baseline specifications, we add to our empirical model lagged capital and employment as controls, to account for differences in inputs across countries and sectors.[38] The

[37] Our baseline specification $\log(1+pat)$ allows us to retain observations with zero patenting. The results are robust to using the inverse hyperbolic sine transformation of the number of patents, instead. Results for alternative log transformation of patents and forward lags for the dependent variable are reported in Table C.3 in the Appendix.

[38] Results with both lagged and contemporaneous capital and employment as controls are very

Kleibergen-Paap Wald F-statistic in the benchmark regression is 34, which rules out weak instrument concerns.

Finally, we investigate whether our estimate of interest may be capturing other contemporaneous spillovers unrelated to patenting. We do this by adding two additional controls to our regression model. First, we add 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 the diffusion of ideas or intangible knowledge through trade (Ayerst et al., 2020). Second, we include a measure of predicted gross output in country c and sector s at time t using the input-output relationship across countries and sectors. We construct this measure analogously to our predicted patents measure, but we use as shares the input-output relationships and as shifts sectoral output in other countries. This addresses the concern that there might be contemporaneous economic spillovers through input-output linkages (rather than citations) that may generate knowledge diffusion and affect domestic productivity. Columns (5) and (6) show that our coefficient of interest remains stable and hardly changes when we include these controls.[39]

The coefficient on innovation activity is positive and statistically significant across specifications. The magnitude of the 2SLS estimates is also stable across specifications. According to our framework, the coefficient in column (4) implies that a 1% increase in patenting leads to a 0.017% increase in value added per worker. This estimated elasticity implies that a 1 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 residual standard deviation in output per worker growth in our sample.[40] In terms of magnitudes, our estimates imply that an interquartile range increase in the log of the number of patents gener-

similar and are reported in Table C.4 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.

[39]In the robustness section below, we also show that the coefficient of interest remains stable if we include the predicted number of patents as an additional control using input-output linkages as shares instead of citations. The construction of both variables is detailed in Appendix A.2.

[40]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 1 standard deviation increase in log patents would imply an increase in log value added per worker (or value added per worker growth) of 4.4 percentage points.

Table 2: 2SLS Estimates: 2000-2014

	$\overline{\log(va\_em_{cst+n})}$	$n \in \{1,2,3\}$				
	OLS	2SLS	OLS	2SLS	OLS	2SLS
	(1)	(2)	(3)	(4)	(5)	(6)
$\log(1 + pat_{cst})$	0.006	0.019	0.004	0.017	0.005	0.018
	(0.003)	(0.008)	(0.003)	(0.008)	(0.003)	(0.007)
$\log(va\_em_{cst})$	0.919	0.917	0.942	0.937	0.933	0.927
	(0.012)	(0.012)	(0.016)	(0.016)	(0.016)	(0.015)
$\log(cap_{cst})$			-0.016	-0.014	-0.015	-0.014
			(0.008)	(0.008)	(0.008)	(0.008)
$\log(empl_{cst})$			0.020	0.015	0.010	0.004
			(0.010)	(0.010)	(0.010)	(0.009)
$\log(int\_imp_{cst})$					-0.003	-0.002
					(0.014)	(0.014)
$\log(\widehat{go}\_IO_{cst})$					0.014	0.014
					(0.012)	(0.013)
Coun.-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-statistic		36.7		33.9		33.39

Notes: The period of the analysis is 2000-14. We use the predetermined matrix based on 1970-90 data. First-stage estimates include all controls. Standard errors (in parentheses) are two-way clustered at country and sector levels. Columns (1), (3), (5), and (7) report results using OLS, and Columns (2), (4), (6), and (8) report results obtained with 2SLS. The Kleibergen-Paap Wald F-statistic is reported for the first stage.

ates an increase of 10.4% of the interquartile range in value added per employment growth. For example, if Mexico in 2000 innovated in computer and electronic products and pharmaceuticals at the level of the US, ceteris paribus, output per worker in these sectors would have been 3.1% and 2.9% higher, respectively.

The estimated 2SLS coefficients are larger than those 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.

Alternative Growth Specification and TFP Regressions To assess the robustness of our findings, we extend our analysis by using TFP growth instead of output per worker as our dependent variable.[41] Table 3 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.[42] Moreover, when comparing the coefficient on patenting, ϕ_N, across specifications we see that, as implied by our simple framework, its magnitude is similar regardless of whether we use value added or TFP as dependent variable.[43]

5.2.1 Robustness Checks

The validity of our shift-share design rests on the assumption that country-sector pairs that give more citations pre-1990 are not on different trajectories in terms of output per worker post-2000. This assumption would be 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 characteristics were not captured by our controls. We test this assumption in a variety of ways.

First, we test for pre-trends by showing that pre-period productivity is uncorrelated with future patent activity as predicted by the instrument. Table 4 presents the results of regressing the average value of productivity during the pre-sample period against the predicted number of patents in the period 2000-14.[44] The coefficients

[41]We obtain measures of TFP growth at the country-sector level at a given period of time using "dual" and "primal" approaches, as in Hsieh (1999) and Hsieh (2002).

[42]As in our baseline specification, the results reported in Table 3 are robust to using the inverse hyperbolic sine transformation of the number of patents instead of $\log(1 + pat)$, adding forward lags as controls and adding all set of controls. See Tables C.7, C.6, and C.5 in the Appendix.

[43]We also find results similar to our baseline ϕ_N when estimating Equation (5) assuming $\phi_A = 1$ (and, thus, having the growth rate as a dependent variable). See Table C.2 in the Appendix.

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

Table 3: 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)
$\log(1+pat_{cst})$	0.011	0.009	0.007	0.010	0.004	0.008
	(0.004)	(0.004)	(0.004)	(0.006)	(0.004)	(0.003)
$\log(y_{cst})$	-0.044	-0.031	-0.017	-0.010	-0.018	-0.009
	(0.009)	(0.007)	(0.010)	(0.009)	(0.009)	(0.005)
$\log(cap_{cst})$		-0.005		-0.023		-0.031
		(0.003)		(0.003)		(0.003)
$\log(empl_{cst})$		0.005		0.021		0.026
		(0.004)		(0.004)		(0.004)
Count.-Year FE	Y	Y	Y	Y	Y	Y
Sector-Year FE	Y	Y	Y	Y	Y	Y
# obs.	8,834	8,357	7,931	7,931	8,554	8,336
# countries	36	36	36	36	36	36
	First-stage estimates					
Predicted	0.468	0.461	0.498	0.470	0.498	0.472
$\log(1+pat_t)$	(0.085)	(0.079)	(0.081)	(0.080)	(0.085)	(0.083)
F-statistic	30.5	33.9	34.5	32.5	38.1	35.0

Notes: The period of analysis is 2000-14. We use the predetermined matrix based on the data from 1970 to 1990. First-stage estimates include all controls. Standard errors (in parentheses) are two-way clustered at country and sector levels. y_{cst} is a respective measure of productivity. In columns (1)-(2), y_{cst} is value added per employment. In columns (3)-(6), y_{cst} stands for TFP measured using either the primal or dual approach. In the case of primal TFP for our baseline specification (Columns (3)-(4)), the main coefficient of interest is significant at the 10% level with p=0.09. The Kleibergen-Paap Wald F-stat is reported for the first stage.

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 empirical analysis, reported in Columns (1) and (2), are indeed significant.

Second, in Column (2) of Table 5, we check that our results hold when controlling for the average level of patenting activity in the period 1970-90. The results are virtually unchanged. The coefficient of interest becomes larger in magnitude (in absolute value), but is statistically indistinguishable from the baseline estimate.

since data for earlier years are not available in the WIOD. We also averaged all the variables to suppress the time dimension because the left-hand and right-hand sides of our regression belong to different time periods.

Table 4: 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 using WIOD and UNIDO data, respectively. The latter is included for better compatibility with results in Columns (3) and (4), where the dependent variable is the average value added per employment computed using 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 country and sector levels.

Third, we want to rule out that our results are driven by demand pull factors from the destination country and sector, rather than a supply push from the origin. We do so by directly controlling for this using 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 patenting activity that we observe in the data. More precisely, we first construct our network of citations, this time using forward citations instead of backward citations. Then, using the patenting activity across country-sector pairs during our sample period (2000-2014), we infer the number of patents in the period 1970-1990 that would have been necessary to rationalize the data in the 2000-2014 period.[45] The estimate presented in Column (3) of Table 5 is consistent with our baseline result. 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 baseline regression. The coefficient remains significant and has a similar magnitude.

[45] 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.

Table 5: 2SLS Estimates: Robustness

	$\overline{\log(va\_em_{cst+n})}$ $n \in \{1,2,3\}$				
	(1)	(2)	(3)	(4)	(5)
$\log(1+pat_{cst})$	0.017	0.029	0.025	0.030	0.017
	(0.008)	(0.010)	(0.010)	(0.010)	(0.008)
$\log(1+\overline{pat_{cs1970-90}})$		-0.009		-0.009	
		(0.005)		(0.005)	
$\log(1+\widehat{pat}_{cst-30})$			-0.006	-0.001	
			(0.007)	(0.006)	
$\log(1+\widehat{pat}\_IOlink_{cst})$					-0.001
					(0.006)
Controls	✓	✓	✓	✓	✓
Country-Year FE	Y	Y	Y	Y	Y
Sector-Year FE	Y	Y	Y	Y	Y
# obs.	8,357	8,357	8,357	8,357	8,357
	First-stage estimates				
Predicted	0.461	0.264	0.388	0.305	0.459
$\log(1+pat_{cst})$	(0.079)	(0.058)	(0.065)	(0.056)	(0.078)
F-statistic	33.9	20.9	35.5	29.3	34.8

Notes: Column (1) shows the results of our baseline regression; Columns (2) and (3) show regression results when separately including the historical levels of average patent activity and the predicted number of patents driven by demand pull factors, respectively; Column (4) shows regression results when including them together; and Column (5) when we include the predicted number of patents using input-output linkages. All regressions include (log) values for value added per employment, capital, and employment as controls. Standard errors (in parentheses) are two-way clustered at country and sector levels. The Kleibergen-Paap Wald F-statistic is reported for the first stage.

Fourth, in Column (5), we test for the possibility that citations capture factors that link countries and sectors other than knowledge spillovers, such as input-output linkages. To do so, we include to our regression model the predicted number of patents constructed using as shares the world input-output matrix rather than the citation network. The coefficient of interest remains stable. Moreover, the coefficient on the new regressor is insignificant both in the first- and second-stage, which suggests that our instrument does not purely capture input-output relationships across countries and sectors.

Finally, to check 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 of these regressions.[46]

5.3 Innovation and Long-term Development

Our empirical analysis so far has 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 that span 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 time period that spans as many years as possible. On the other hand, given that comprehensive patent data for the period prior to 1970 mostly cover advanced economies, and given that for most developing countries we observe little to 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, and we use the pre-1980 data to construct our instrument. The set of countries we consider are those categorized as high-income and upper-middle-income countries according to the World Bank classification, for which we have substantial variation in patenting activity.

The regression model that we estimate corresponds to Equation (4) in our motivating framework (without sectoral variation)[47]:

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

[46]The largest change in magnitude we obtain in ϕ_N is when we exclude the sector "Manufacture of wood and of products of wood and cork, except furniture; manufacture of articles of straw and plaiting materials." In this case, it increases from 0.017 to 0.023.

[47]As a reminder, it is obtained from combining a Cobb-Douglas aggregate production function and our law of motion for TFP.

where on the left-hand side variable is the average level of GDP per capita over $n = 3$ years after t to smooth out business cycle fluctuations and other idiosyncratic shocks.

Table 6 reports 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 to 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, as predicted by our theoretical framework. The elasticity of patenting to income per capita is 0.034.[48] Quantitatively, this elasticity implies that 1 residual standard deviation increase in the logarithm of the annual number of patents leads to a 0.41 residual standard deviation increase in the logarithm of annual GDP per capita, which implies 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.9 in the Appendix report the results of running the same specification, Equation (7), using income per capita data that span 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 innovation's network is more sparse. We estimate a positive and significant effect of innovation on income per capita growth in both regressions. The implied magnitudes suggest that a 1 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 that span the past 100 years and a large range of countries to study the evolution of innovation across time and space

[48]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.9 in the Appendix.

Table 6: 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-statistic		15.0		7.3

Notes: The period of the analysis is 1980-2016. We use the pre-determined matrix based on data for the pre-1980 period. Standard errors (in parentheses) are clustered at country level. Columns (1) and (3) present the results for OLS, and Columns (2) and (4) present 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, regressions in columns (3) and (4) also include year fixed effects. The Kleibergen-Paap Wald F-statistic is reported for the first stage.

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 document 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 in international knowledge spillovers, as measured by patent citations since the 1990s. This rise is mainly accounted for by an increase in citations of 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 and heterogeneity in the exposure of countries to technological waves. We then use our proposed instrument to estimate the causal effect of innovation on output per worker and TFP growth in a panel of country-sectors over the period 2000-2014. We find that, on average, an increase of 1 standard deviation in patenting implies a 1.1 percentage point increase in output per worker growth. Moreover, the first stage of our empirical setting informs us on the elasticity of innovation on international knowledge—a central parameter for growth theory—which we find to be around 0.5.

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. An increase in one standard deviation in patenting activity increases income per capita by 0.28 standard deviation. We believe that our shift-share design can be applied to other settings in which the effect of innovation or productivity is 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. (2009): *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 Macro Annual 2015, Vol. 30*, NBER, 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. T. Ates, and G. Impullitti (2018a): "Innovation and Trade Policy in a Globalized World," NBER Working Papers 24543, National Bureau of Economic Research, Inc.

Akcigit, U., S. Baslandze, and F. Lotti (2018b): "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.

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.

——— (2007): "The International Dynamics of R&D and Innovation in the Long Run and in The Short Run," *Economic Journal*, 117, 486–511.

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 , 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 , LSE.

Eaton, J. and S. Kortum (1999): "International Technology Diffusion: Theory and Measurement," *International Economic Review*, 40, 537–570.

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.

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.

Hémous, D. (2016): "The dynamic impact of unilateral environmental policies," *Journal of International Economics*, 103, 80–95.

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, GGDC, U. of Groningen.

Jaffe, A., M. Trajtenberg, and R. Henderson (1993): "Geographic Localization of Knowledge Spillovers as Evidenced by Patent Citations," *Quarterly Journal of Economics*, 108, 577–598.

Jones, B. F. (2009): "The Burden of Knowledge and the "Death of the Renaissance Man": Is Innovation Getting Harder?" *The Review of Economic Studies*, 76, 283–317.

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.

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 Macro Annual 1997, Vol. 12*, NBER, 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.

Kortum, S. S. (1997): "Research, Patenting, and Technological Change," *Econometrica*, 65, 1389–1420.

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 WP 26483.

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.

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," WP 678, BdF.

Romer, P. (1990): "Endogenous Technological Change," *Journal of Political Economy*, 98, S71–102.

Sampson, T. (2023): "Technology Gaps, Trade, and Income," *American Economic Review*, 113, 472–513.

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," *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*.

Young, A. (1998): "Growth without Scale Effects," *Journal of Political Economy*, 106, 41–63.

Link to Appendix (Click Here)