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

Enrico Berkes — Ohio State
Kristina Manysheva — Northwestern
Martí Mestieri — FRB Chicago & Northwestern

April 5, 2021

Preliminary

Abstract

We use a panel of historical patent data covering the last hundred years and a large range of countries to study the evolution of innovation across time and space and its effect on productivity. We document a substantial rise of international knowledge spillovers as measured by patent citations since the 1990s. This rise is mostly accounted for by an increase in citations to US and Japanese patents in fields of knowledge related to computation, information processing, and medicine. We estimate the effect of innovation induced by international spillovers on TFP in a panel of countries-sectors from 2000 to 2014. We develop a shift-share instrument that leverages pre-existing citation linkages across countries and fields of knowledge, and heterogeneous countries' exposure to technology waves. On average, an increase of one residual standard deviation in patents increases TFP by 0.1 residual standard deviation. The effect on income per capita since 1960 is even larger. An increase in one residual standard deviation in patenting activity induced by international spillovers increases income per capita by 0.28 residual standard deviation.

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

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

1 Introduction

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

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

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

Armed with our newly defined technology classes, we show that their significance – as measured by the share of filed patents that goes to each field of knowledge – has importantly evolved over time. The data reveal substantial technological waves in the last one hundred

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

[2] As a robustness check, we also perform a clustering analysis where the strength of the linkages between different patent classes is based on the cross citation and/or co-appearance.

years. Mechanical engineering accrued the largest share of innovations at the beginning of the twentieth century. Chemistry and physics were the most prominent fields in the mid-century, while medicine and the digital economy appear to be the most important technologies in the past decades. We also show that, while advanced economies account for the bulk of patenting activity, there is substantial variation in terms of countries' specialization across fields of knowledge. Moreover, these patterns of specialization are heterogeneous over time.[3]

Next, we turn our attention to knowledge spillovers. We measure knowledge spillovers through citations across fields of knowledge and countries. For this exercise, we focus on the post 1970 sample for which we have data for virtually all countries in the world. We show that, for the average patent, citations tend to be biased towards domestic, as opposed to international, inventions and towards the same field of knowledge. We also document that, across all these categories, there is an upward trend over time in citations. That is, new patents tend to cite more other patents than older patents.

A striking fact has emerged since the 1990s. Except for the US and Japan, international citations have grown *faster* than domestic citations. After the year 2000, excluding the US and Japan, international citations are more than twice more frequent than domestic citations. This finding suggests that the reliance on knowledge produced elsewhere – and particularly in the U.S. and Japan – has increased over this period of time. Even for technology leaders like Germany or Great Britain, foreign citations now account for most of the citations. This fact may be interpreted as a decline in the prominence of European innovations relative to their U.S. and Japanese counterparts. We also find that most of this increase is driven by a handful of fields of knowledge that are related to ICTs and medicine.

After having laid out these facts, we investigate the effect of innovation (as measured by patenting) on productivity and income. Our baseline exercise studies the effect of innovation induced by international spillovers on productivity in the latest part of the sample (2000-2014) for which we have high quality data on cross-country sectoral TFP.[4] We then extend our analysis back in time and study directly the effect on long run income growth (1960-2016), for which we use the full extent of our patent data.

Simply correlating innovation and productivity (or income) is problematic due to measurement error (which would generate attenuation bias), potential reverse causality, and the presence of unobserved factors affecting simultaneously patenting and the dependent variables. Examples of such factors include financial or external shocks that affect both the output of a country and the amount of innovation produced. In this paper, we address the endogeneity concerns by constructing a shift-share instrument that leverages pre-existing knowledge links

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

[4] We use patent data starting in 1970 to construct our instrument for this exercise.

across countries and technologies and combines it with lagged foreign innovative output in other fields of knowledge. More precisely, our instrument is constructed in two steps – which we discuss now in the context of our baseline exercise studying productivity from year 2000 to 2014 as dependent variable. First, we estimate the strength of the linkages across countries and fields of knowledge (measured by patent citations) between 1970-1990. These constitute our pre-determined *shares*. The *shifts* of our instrument for country c_o and field of knowledge k_o are given by the patents filed in all other countries $c_d \neq c_o$ and fields of knowledge $k_d \neq k_o$ over the years 1990-2000. We are thus implicitly assuming that the probability that patents in (c_d, k_d) generate a patent in (c_o, k_o) can be inferred from the network of patent citations.[5] Applying this procedure recursively, we obtain a predicted number of patents for each country and field of knowledge.

In our main regression, the dependent variable is TFP by country and sector (measured from the World Input Output Database) over the 2000-2014 period. The regression model includes controls that vary at the country-sector-time (e.g., sectoral capital and labor) and includes country-time and sector-time fixed effects to control for differential country and sectoral trends. We find a robust effect of innovation on TFP growth. One residual standard deviation increase in patent growth leads to 0.1 residual standard deviation increase in TFP, after partialling-out fixed effects and controls. We conduct a number of robustness checks to address concerns regarding the validity of the instrument such as the existence of demand-pull, anticipatory effects that might be correlated with the contemporaneous state of the local economy. To do this, among the other things, we "reverse" the network of citations that we used to measure knowledge spillovers and calculate the amount of innovation we would expect to observe *in the past* if the patenting activity was driven by future demand. Reassuringly, we find no evidence supporting this alternative hypothesis.

We conclude the paper by doing two additional exercises. First, we extend our framework to study the effect of innovation over long-run growth. We reconstruct our shift-share using patent data pre-1950 and estimate the effect of innovation on income per capita over the 1960-2016 period. We find a positive, significant coefficient that is very similar in magnitude to the elasticity of patents to TFP that we find for the period 2000-2014. In terms of magnitude, an increase in one standard deviation in patenting activity increases income per capita by 0.28 standard deviation. Second, we illustrate how this shift-share approach can be used in other settings, and we show how it can be used to compute the elasticity of trade flows to sectoral TFP.

[5] In fact, we refine this procedure and extend this logic to higher-order linkages to create our main instrument (see Section 5).

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

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

Our shift-share instrumental approach is related to a number of papers that have used the network structure of citations to construct shift-share instruments. Our approach is most similar to Berkes and Gaetani (2018b), who construct a similar shift-share instrument across US cities and Acemoglu et al. (2016) who use a citation network to percolate innovations downstream and illustrate how technological progress builds upon itself. Both papers use only within country (US) variation.[6]

2 Data

2.1 Data Sources

In this paper, we measure new ideas through patents data, while productivity is measured through TFP and value added data. Patent data are collected from the European Patent Office worldwide Patent Statistical Database (PATSTAT, Autumn 2018 version). PATSTAT contains bibliographical and legal status information on more than 110 million patents from the main patent offices around the world, covering leading industrialized countries, as well as

[6]A large number of papers have used more standard shift-share instruments in the innovation and productivity literature, e.g., Moretti et al. (2019) to estimate the effects of R&D subsidies and Hornbeck and Moretti (2019) estimate the effect of TFP growth in manufacturing across US cities.

developing countries over the period 1782-2018.[7] From PATSTAT, we collect information on patent filing years, inventor and assignee locations, citations, patent families, and technological classes. While PATSTAT provides the most comprehensive coverage of patenting activities worldwide, it also has some limitations (Kang and Tarasconi, 2016). The main limitation for our purposes is data availability in the earlier years. In fact, data along one or more dimensions are often missing for some countries in the years preceding 1970. We therefore split our sample into two groups of countries, that we use at different stages of our analysis. The first group is composed of six major technological leaders (U.S., Great Britain, France, Germany, the Soviet Union, and Switzerland)[8] for which all the patents' characteristics required by our analysis are available since 1920. The second group includes all the countries covered by PATSTAT and starts in 1970.[9] Appendix A provides more information about the composition of the sample and summary statistics.

We assign each patent to a geographical unit according to the country of residence of its inventor(s). If this information is not available, then the country of the assignee(s) or publication authority is used, instead. When a given patent is associated to 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. 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. To avoid double-counting patents that are filed in more than one patent office, we restrict most of our analysis to patents that are the first in their (DOCDB) family. We further collect the full distribution of technology classes associated to each patent based on the International Patent Classification (IPC). For our analysis, we first consider all the fields at the subclass level (e.g., A01B) – for a total of 650 classes – and we then cluster them into consistent groups following the procedure outlined in Section 2.2. Finally, to capture when an idea was completed and abstract from potential bureaucratic delays that are

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

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

[9]For our empirical analysis, we exclude China from our sample due to a substantial rise in the number of Chinese patents since the 3rd revision of Patent law in China in 2008. While we see a sharp increase in total number of Chinese patents after the implementation of the new law, the same pattern is not observed in the number of Triadic patents which include patents filed jointly in the largest patent offices, that is 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.

orthogonal to innovative activities, in our analysis we use the patent filing years instead of the years in which patents were granted.[10]

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

2.2 Construction of Fields of Knowledge

Innovation is the process of creating new knowledge building on existing knowledge across different fields. To operationalize our goal of measuring innovation waves across time and space, we build on the vast existing literature that measures innovative activities through patent data. We propose grouping finely-defined patent classes into "fields of knowledge," which taken together constitute what we refer to as the *technology space* of the world. This conceptualization also provides a mapping between our patent data and the analytical framework developed in Section 4.[12]

In this paper, we employ a novel approach to group patent technology classes based on inventors' information. Our procedure is based on the likelihood that the same inventor produces inventions associated to different patent subclasses. The idea is that, since knowledge is embedded in people, it is possible to cluster fields of knowledge based on the IPC subclasses in which the same inventors tend to patent.[13] More precisely, we build a probability matrix $T_{642\times642}$,[14] where each element (i, j) is the probability that an inventor patents in IPC subclass i conditional on having patented in subclass j.[15] For example, a mechanical engineer specialized

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

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

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

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

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

[15]The diagonal elements of the matrix, $i = j$, are set equal to one. The so-obtained matrix is not symmetric. For example, manufacture of dairy products (A01J) is closest to dairy product treatment (A23C), while dairy

in brakes will most likely patent in IPCs B60T "Vehicle Brakes or parts thereof" and F16D "Clutches, Brakes", which our algorithm correctly bundles together.[16]

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

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

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

3 Some Stylized Facts on World Innovation

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

product treatment is closest to foods, foodstuffs, or non-alcoholic beverages (A23L)

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

[17]More details on the procedure used to construct fields of knowledge can be found in the Appendix A.4.

[18]As a robustness check, we also construct the proximity matrix based on the citation linkages, instead, and apply the same procedure. The results are similar to the ones obtained with our proximity matrix: (i) the percentage of pairwise IPC subclasses that are in the same cluster is 50.6 (excluding singleton clusters, which accounts for 22.6 percent of all clusters); (ii) the percentage of pairwise IPC subclasses that are in the same cluster weighted by importance, measured by the number of patents in the respective subclass relative to all patents, in the sample is 51.9 (excluding singletons); (iii) the percentage of the same clusters' centers is 67.1.

3.1 Evolution of Fields of Knowledge across Space and Time

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

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

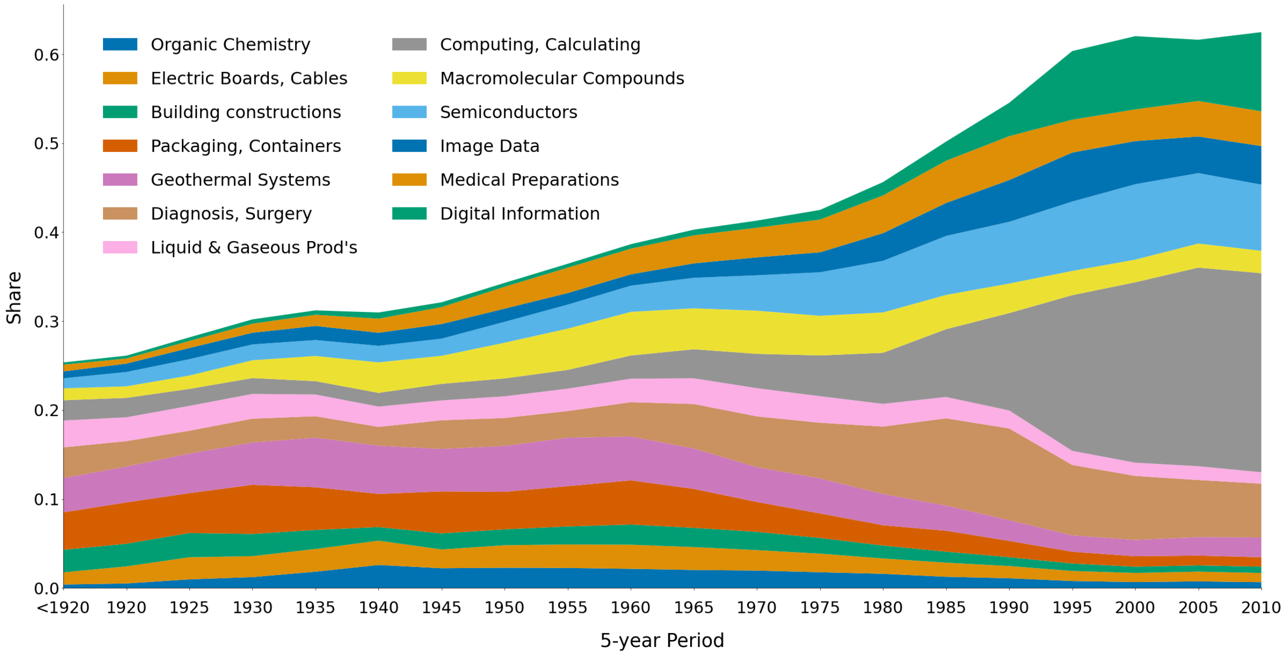
We also perform the same exercise using alternative measures of importance that address possible concerns related for example to heterogeneous patenting practices across countries or the strategic patenting behavior that gained more prominence in the past few decades. To do this, we build importance measures that take into consideration country fixed effects, or are only based on patents that were cited at least once. Table B.2 shows that these measures are highly correlated to our baseline.

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

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

[20] In this part of the analysis, we use the total number of patents without weighting by the number of citations

Figure 1: Evolution of Top Fields of Knowledge



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

.

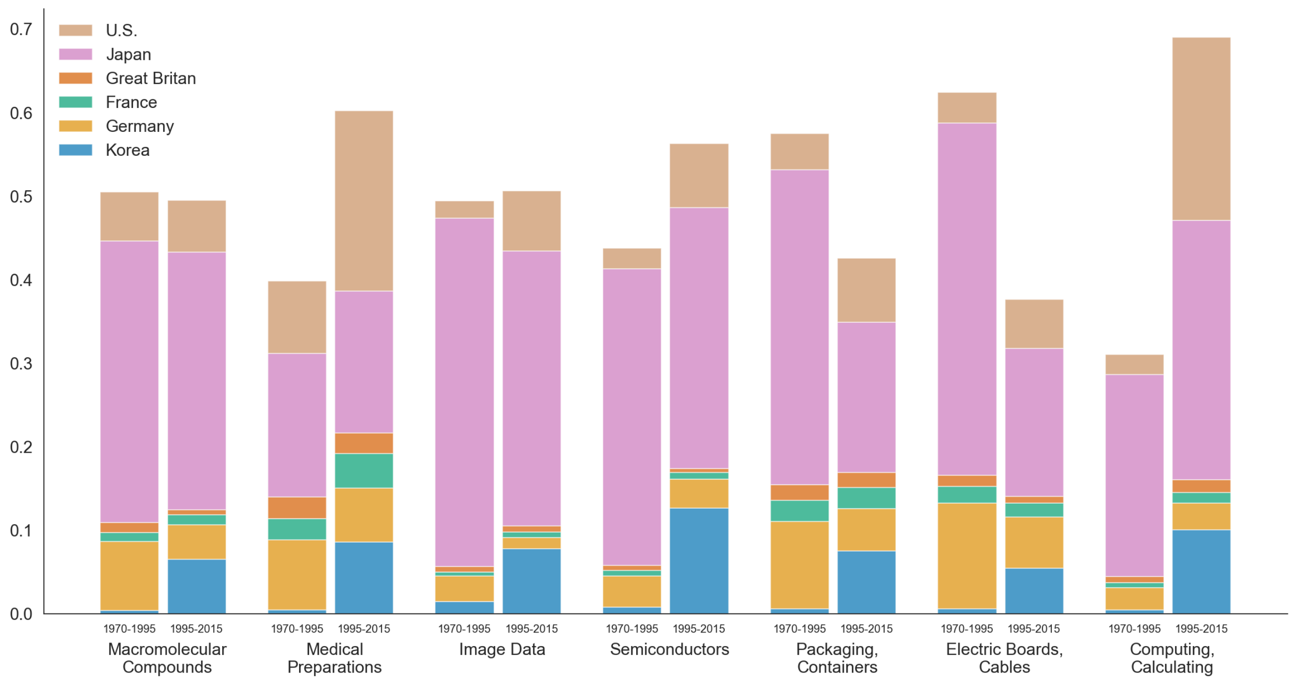
For the period 1920-1970, our sample is limited to six countries: the U.S., Great Britain, Germany, Switzerland, France, and the USSR. Figure B.1 shows that during this time period, the leading innovating role in major fields of knowledge was split between the U.S. and Germany, followed by the UK and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

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

We can extend our analysis beyond the chosen fields of knowledge and compute an overall ranking by averaging the country ranking across all fields of knowledge. This exercise paints a picture similar to the one in Figure 2. Japan and the US are the technological leaders from

for better comparability across countries. Different countries use different procedures to assign citations, which is likely to bias our results.

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



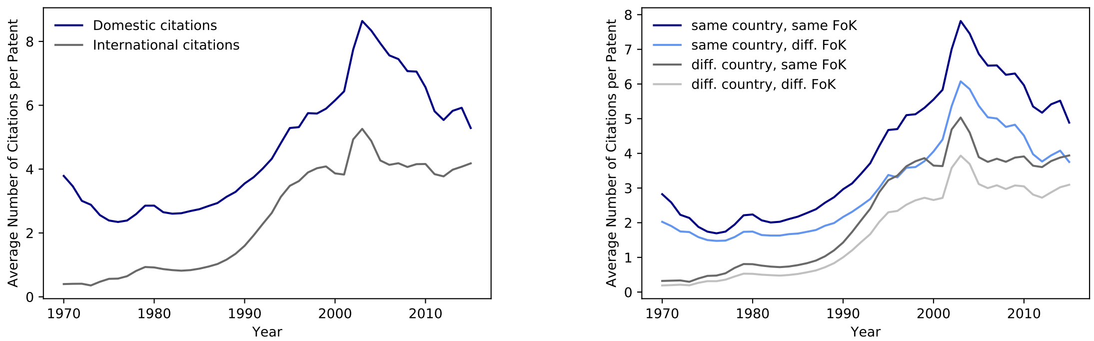
1970 until 1995, with Japan falling behind after the 2000s. The Soviet Union has an average ranking very similar to the US in 1970 but it falls behind subsequently, while Asian countries such as Taiwan gain prominence after the 2000s.[21]

3.2 Using Citations to Measure Spillovers across Time and Space

So far, we have shown that there is substantial time variation in the most prominent fields of knowledge and which countries contribute to their growth. Next, we turn our attention to patent citations across fields of knowledge and countries as a proxy for knowledge spillovers. There is an abundant literature studying within country spillovers using patent citations (e.g., Jaffe et al., 1993, Murata et al., 2014 for the United States). By contrast, the evidence on cross-country knowledge spillovers is more scarce. Despite being an imperfect measure of knowledge spillovers, patent citations provide a useful quantifiable benchmark that can be easily measured and used in our empirical exercises.

For this exercise, we focus on the post 1970 sample, for which we have data for virtually all

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



(a) Domestic vs International

(b) Citation Decomposition

Figure 3: Citation Dynamics, 1970-2015

countries in the world. We compute citations given to patents filed after 1900. Figure 3 shows the evolution of the average number of backward citations per filed patent since 1970. We observe that citations rise substantially after 1990 and decline somewhat after 2010. Panel (a) shows that domestic citations tend to be more prominent than citations given to international patents: domestic patents are cited at a rate that is roughly double the one for international patents. The time path is however similar for both. In panel (b) we decompose citations given to the patents coming from (i) the same country and the same field of knowledge, (ii) same country and different field of knowledge, (iii) different country, but the same field of knowledge, and finally (iv) different country and field of knowledge.[22] We see that the four components play a similar role important in driving the overall trend and with a constant relative ranking where domestic, same field of knowledge tend to be the most cited group and foreign different field of knowledge patents are the least cited.

To formally decompose the relative contribution of these four components across countries and fields of knowledge, we propose the following exercise. Using citation patterns over the period 1998-2018 we construct a network that captures linkages across countries and fields of knowledge. Specifically, each edge in the network corresponds to the number of citations given from field of knowledge k_o and country c_o to another field of knowledge k_d and country c_d with a lag $l \in \{1, ..., 10\}$.[23] We generate the predicted number of patents in the period 2000-15 by interacting (i)-(iv) components of this network structure with the patent growth in the respective field of knowledge and country.

[22] It is important to notice, that the sum of (i)-(iv) does not equal to the total backward citations since there is double-counting due to the fact that cited patents belong to multiple fields of knowledge and (more rarely) to multiple countries.

[23] To compute actual linkages we also normalize the number of backward citations in a number of ways to account for the overall number of citations given by each patent and overall trends in patent activity. See Section 5.1 for a formal discussion.

Table 1: Decomposition of Knowledge Spillovers

	Dependent Variable is: $\ln(1+pat)_t$				
	Same Country		Different Country		All
	Same FoK	Diff. FoK	Same FoK	Diff. FoK	
$\ln(1+\hat{pat})_t^{scsk}$	0.524				0.124
	(0.055)				(0.039)
$\ln(1+\hat{pat})_t^{scdk}$		0.889			0.383
		(0.156)			(0.122)
$\ln(1+\hat{pat})_t^{dcsk}$			0.563		0.195
			(0.026)		(0.010)
$\ln(1+\hat{pat})_t^{dcdk}$				0.792	0.494
				(0.032)	(0.036)
Country×Year FE	Y	Y	Y	Y	Y
FoK×Year FE	Y	Y	Y	Y	Y
Obs.	447,677	447,677	447,677	447,677	447,677
R^2	0.86	0.87	0.87	0.88	0.88

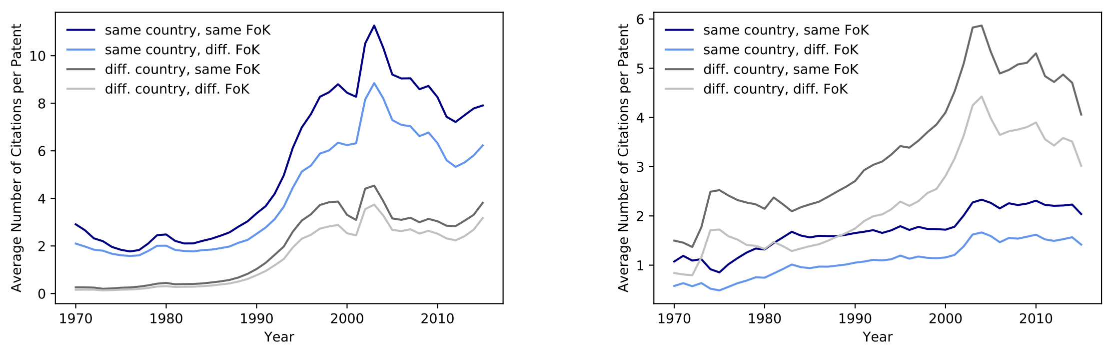
Notes: FoK stands for Field of Knowledge. *scsk* stands for same country, same field of knowledge. Regressors are predicted contributions given citation linkages as explained in the main text.

We use predicted innovation to analyze how much of the aggregate variation in patenting levels across countries and fields of knowledge can be explained by each component. Table 1 reports the results. We find that all four components play a substantial role. To take into account the presence of country and fields of knowledge leaders, as well as trends in innovation activity, we include country-year and field of knowledge-year fixed effects. The higher coefficients on the foreign component relative to the domestic one suggests that new innovations build on past achievements of technological leaders. At the same time, the coefficients are higher for the component that comes from different field of knowledge, which reflects the complexity of innovation.

As we have discussed in the previous section, one important feature of the patent data is that most knowledge (as measured by patent filings) is produced by a handful of countries, the "technological leaders". Specifically, for the period 1970-2015 two countries – Japan and the United States – are the major innovating economies. Figure 4 separately depicts citation dynamics by Japan and the U.S. and the rest of the world. While we observe increase in average number of citations per patent, there are two important differences between panel (a) and panel (b). First, the United States and Japan on average give more citations per patent than the rest of the world. Second, most of the citations in the U.S. and Japan, as shown in panel (a),

Figure 4: Citation Dynamics, 1970-2015

(a) US and Japan (b) Non-US and Non-Japan



are given to domestic patents, while the rest of world mostly relies on knowledge produced in other countries.[24]

Both Figures 3 and 4 depict a rapid increase in the average number citations per patent. To understand what lies behind this increase, we first, look at whether there any changes in citations to different field of knowledge. We observe a substantial increase in number of citations that are given to patents belonging to one particular field of knowledge – "Computing, Calculating, Counting". What is perhaps more striking is the fact that most citations to this field of knowledge are given to American and Japanese patents. As we show in Figure 5, this pattern is observed among domestic citations in the U.S. and Japan, as well as among international citations in the rest of the world.

Taken together, the evidence presented in this section paints a picture consistent with international knowledge spillovers increasing their prominence in the last decades. This increase is visible when considering spillovers both in the same field of knowledge and other fields of knowledge. This increase in international knowledge spillovers is driven by a dramatic increase in the citations received by the U.S. and Japan, especially in the fields of knowledge related to computing, information processing and medicine.

4 Conceptual Framework

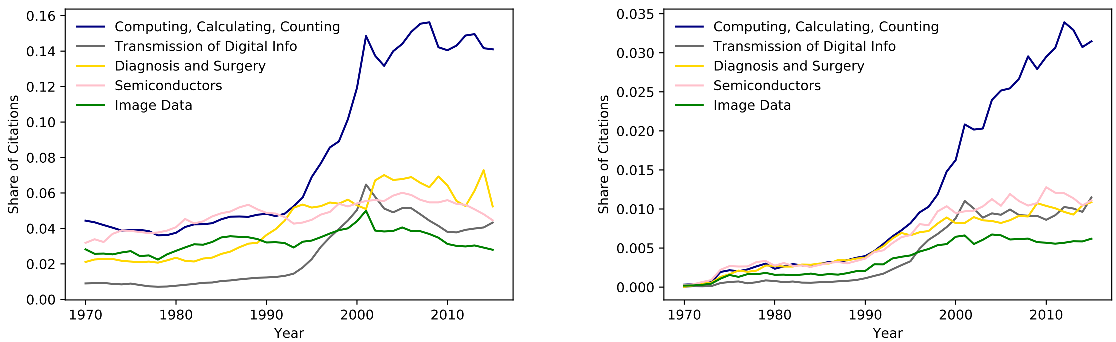
In this section, we present a framework that incorporates the elements of our data analysis in the previous sections and that serves as a guide for empirical exercises. Our framework builds on the canonical growth literature. The fundamental element of our analysis is the production function

[24] Decomposition of citations for other countries – US, Germany, France and Great Britain – are reported in Figure B.2.

Figure 5: Share of citations going to the US and Japanese patents by FoK, 1970-2015

(a) US and Japan

(b) Non-US and Non-Japan



Notes: Each line on the plots represents the share of citations going to the U.S. and Japan patents that belong to a given field of knowledge. Panel (a) depicts domestic citations given by American and Japanese patents, and panel (b) depicts international citation to the patents filed in the U.S. and Japan given by other countries.

of ideas and its link to patenting activity. We choose our formulation of the idea production function to remain relatively parsimonious so that it encompasses alternative formulations of endogenous growth theory (see, e.g., Jones, 1999 for a discussion).[25]

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

We denote by N_{cskt} the stock of ideas available in country c, sector s, field of knowledge k and time t. The state of world ideas at time t is thus summarized by the vector $\mathbf{N}_t \equiv (N_{111t}, \ldots, N_{cskt}, \ldots, N_{CSKt})$. 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},

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

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

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 in field of knowledge k for sector s. We take the spillover function to be

$$S_{csk}(\mathbf{N}_t) = \sum_{c\in C}\sum_{s\in S}\sum_{k\in K} \alpha_{cskt} N_{cskt}, \tag{2}$$

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. Note that we purposefully state Equation (1) generically so that it subsumes the first generation of endogenous growth models as in Romer (1990) or Aghion and Howitt (1992), semi-endogenous growth as in Jones (1995), Kortum (1997) or Segerstrom (1998), or second generation as Aghion and Howitt (1998), Young (1998) or Peretto (1998).[26]

Since ideas are to a large extent non-rival (Romer, 1990), the vast majority of these theories resort to intellectual protection in the form of patents to ensure that investments in new ideas can be recovered with future profits.[27] This observation motivates our empirical strategy to proxy the generation of new ideas through patent filings. Patents provide a quantifiable measure over time and space that is arguably very hard (or impossible!) to replicate with other measures of ideas or innovation. Moreover, through citations, patents also provide provide an empirical measure of reliance on existing ideas across space and fields of knowledge–knowledge spillovers. 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 a patent builds on are cited. We thus think of patents as a *proxy* for new ideas, ΔN_{cskt} and citations as *proxy* for spillovers. We discuss in the next section how our empirical specification addresses these potential discrepancies between idea generation and patenting.

Regardless of their vintage, endogenous growth theories argue that there is a positive, monotonic relationship between the ideas produced and sectoral TFP growth TFP_{cst+1}/TFP_{cst}. However, they differ on the implied effect of the current stock of ideas on the generation of new ideas: first-generation theories emphasize the standing on the shoulders of giants effect, while semi-endogenous theories allow for fishing-out effects. To build a connection with our empirical specification, we assume a flexible, isoelastic relationship between ideas and TFP growth

$$\log TFP_{cst+1} = \phi_0 + \phi_A \log TFP_{cst} + \phi_N \log(1 + \Delta N_{cst}), \tag{3}$$

with $\phi_0, \phi_A, \phi_N \geq 0$ and $\Delta N_{cst} = \sum_{k=1}^{K} \Delta N_{cskt}$ denoting the total number of ideas generated

[26] For example, one specification extensively used in the literature (cf., Romer, 1990; Jones, 1995) ignores cross-country spillovers, assumes that $S = K = 1$ and $S_c(\mathbf{N}_t) = N_{ct}$ and postulates a log-linear relationship, $I = N_{ct}^{\phi} R_{ct}$ with $\phi \leq 1$.

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

in country-sector s at time t across all fields of knowledge.

Equation (3) nests a number of cases often considered in the literature and constitutes the basis of our empirical specification in the next section. For example $\phi_0 = 0$ and $\phi_A = \phi_N = 1$ generates building-on-the-shoulders-of-giants dynamics, whereby the growth rate of TFP_{cst} is directly controlled by the number of ideas produced at time t. In this case, if no ideas are produced at time t, $\Delta N_{cst} = 0$, there is no TFP growth. Letting $\phi_A < 1$ introduces the fishing-out-of-the-same-pond effect in the sense that more ideas become necessary over time to sustain constant TFP growth.

Finally, we extend our framework to output per worker–which we also study as an indirect proxy for productivity. Suppose that output per worker, y_{cst}, is given by a Cobb-Douglas production function, $\log y_{cst} = \log TFP_{cst} + \alpha \log k_{cst}$, where k_{sct} denotes capital per worker and $0 < \alpha < 1$. Under the assumption of competitive markets, firm optimization implies that the ratio of sectoral output per worker between two sectors, s and s', is proportional to their TFPs,

$$\log y_{sct} - \log y_{s'ct} = \log TFP_{sct} - \log TFP_{s'ct}. \tag{4}$$

Equation (4) implies that the differential growth rate in output per worker across sectors coincides with the differential growth rate in sectoral TFPs.[28] We use this result as a robustness check when TFP data are available and, more importantly, for instances when only GDP per capita data are available. For this latter case, the case in point is the study for very long-run growth trajectories (1960-2016).[29]

The empirical specification we use when considering output per worker builds on the standard growth regression specification obtained by log-linearizing around the steady-state a Solow model,[30]

$$\begin{aligned} \log y_{cst+1} &= \log y_{cst} + \Delta \log TFP_{cst} + \beta\left(\log y_{ct} - \log TFP_{cst}\right) + \theta \log(1 + \Delta N_{cst}) + \delta_{cs} \\ &= \beta_N \log(1 + \Delta N_{cst}) + \beta_Y \log y_{cst} + \beta_K \log k_{cst} + \delta_{cs}, \end{aligned} \tag{5}$$

where δ_{cs} is a country-sector specific intercept that absorbs the steady-state output per worker of the sector. We have used Equation (3) to go from the first to the second line. The noteworthy feature of Equation (5) relative to Equation (3) is that the level of output per worker also

[28] If we allow α to be sector specific, we have that the difference in output per worker growth rates has an additional term that depends on factor prices weighted by factor share differences which can be absorbed using a country-time fixed effect. Letting R_{ct} denote the price of capital, we would have the term $(\alpha_s - \alpha_{s'}) \log R_{ct}$ appearing in addition to $\log TFP_{sct} - \log TFP_{s'ct}$ in Equation (4).

[29] For this exercise, we even omit sectoral considerations and focus on an aggregate production function, since sectoral output data is not consistently available.

[30] See Barro (1991); Barro and Sala-i Martin (1992); Barro et al. (2004); Acemoglu (2009a); Durlauf et al. (2005) for a detailed derivation and further discussion.

appears on the right-hand-side. This term controls for convergence effects and its analysis has been the focus of empirical growth theories in the last decades. By contrast, however, the focus of our analysis will be on the elasticity of patenting on output growth, β_N, rather than the convergence term β_Y.

5 Empirical Analysis

This section presents the main empirical exercises of the paper to study the effect of innovation on productivity. We begin analyzing cross-country panel data on sectoral TFP. We present our identification strategy and report our baseline results. We finalize the section presenting two extensions. First, we extend our baseline estimation to longer time horizons where the dependent variable is output per capita (thus loosing sectoral variation). Second, we illustrate how our IV strategy may be useful in other contexts and show how to apply it to estimate the elasticity of trade flows to differences in productivity.

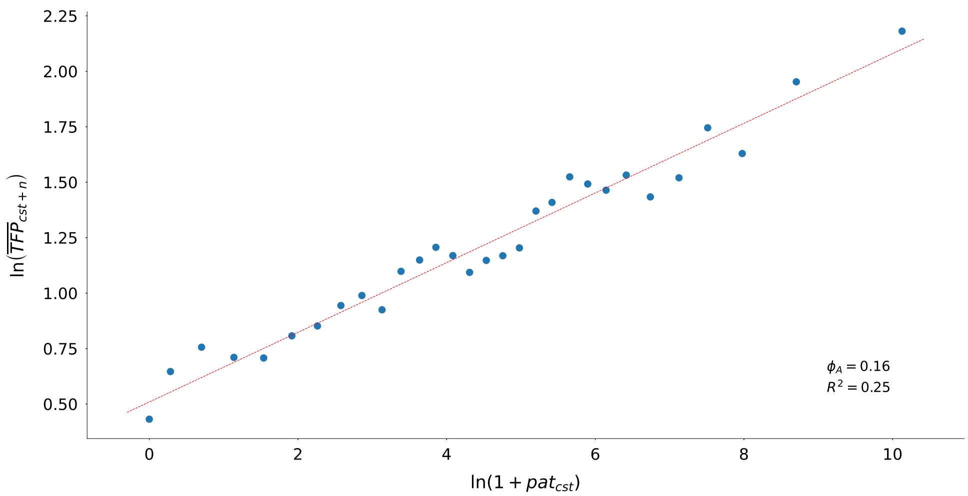
Our empirical model is based on Equation (3) from our analytical framework,

$$\ln(\overline{TFP}_{cst+n}) = \phi_A \ln(TFP_{cst}) + \phi_N \ln(1 + pat_{cst}) + \phi_0 X_{cst} + \delta_{ct} + \delta_{st} + \epsilon_{cst} \tag{6}$$

where $\ln(\overline{TFP}_{cst+n})$ is an average of future TFP spanning n consecutive years (i.e., from $t+1$ to $t+n$), X_{cst} denotes a set of controls for country c and sector s and δ_{ct} and δ_{st} denote country-time and sector-time fixed effects. Thus, relative to the model presented in the analytical framework, there are two departures. First, rather than looking one period ahead, we look at an average over a window of n years (we take $n = 3$ as our baseline, and show that the results are robust to $n \in \{1, ..., 5\}$). We do this, as it is common in the growth literature (e.g., Arcand et al., 2015), to smooth out short-term fluctuations in the patenting activity and concentrate on long-run trends. Second, we unpack the constant ϕ_0 in our analytical framework to allow for controls that are country-sector-time specific (e.g., capital, employment), and country$\times$time and sector$\times$time to allow flexible differential trends across countries and sectors.

Our main results use the TFP measures derived from the World Input Output Database. The data used in our baseline analysis span from year 2000 through 2014, and covers 36 countries and divides national economies into 20 sectors. As we discuss below in more details, we use the 1970-2000 patent data to construct our instrument. Figure 6 shows the binscatter plot of the correlation between patent activity $(1 + pat_{cst})$ and productivity $\ln(\overline{TFP}_{cst+n})$ during the period 2000-2014. In the cross-section of countries and sectors, a one percent increase in tge number of patents is associated with a 0.16 percent increase in future TFP averaged over the next three years. The coefficient is statistically significant.

Figure 6: Unconditional Correlation between TFP and Number of Patents



5.1 Identification and Threats to Validity

Equation (6) is our baseline model to study the effects of innovation on productivity. The coefficient of interest is ϕ_A that relate changes in number of patents at a country-sector in a given year to changes in TFP in the following years.

The inclusion of sector-year dummies accounts for the fact that different industries rely differently on innovations, and such relationship can vary over time. In addition, we want to take into account the presence of technological waves that are demand-driven or some other shocks, that are common across all countries. The inclusion of country-year fixed effects accounts, first, for the fact that different countries have different propensities to innovate, and, second, for any business cycles fluctuations at a country level, like financial crisis.

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

To deal with these issues, we use build an instrumental variable for the actual number of patents. We start our analysis with a simple and intuitive Bartik-like approach to construct the instrument. Specifically, this instrument predicts the number of patents in a country c_o and sector s_o by interacting the share of citations that this country-sector gives to patents in other sectors s_d in pre-sample period with subsequent patent flows in these sectors across other countries. In other words, the instrument relies on the exposure of patent activity in a specific country and sector to patent activity in other sectors measured by citations.[31] To compute the shares, we use data on patents in 1990 and citations given up to 10 year lag. We exclude exposure of patenting activity to innovations in the same country and the same sector to take into consideration potential endogeneity concerns arising from country or sector specific factors, such as industrial policy. In addition to this instrument, we compute predicted number of patents, that not only takes into account exposure of patent activity to innovation in other sectors, but also to innovation in other sectors and countries.

The results of our regression analysis using the instrument described above are in the Table C.1 and suggest that there is no statistically significant relationship between patent activity and productivity. And while the instrument(s) described above has the advantage of being intuitive, and the approach used to construct it has been widely used in the literature, identification concerns are still present. Despite the fact that we use shocks originated in other countries to construct predicted number of patents, it is still possible that endogeneity concerns are not fully addressed due to presence of countries leaders who determine in which sectors most of innovation activity is going to happen. As a result, the shocks that we use in the construction of the instrument are not orthogonal both to patent activity and productivity. Moreover, this instrument does not take into account the lagged nature of knowledge spillovers.

We propose an instrument for patent activity at the country and sector level that is used to tackle above mentioned challenges, and which we use in our preferred specification. Specifically, we adopt a shift-share strategy similar to the one proposed by Berkes and Gaetani (2018a). The idea is to exploit a pre-determined network, rather than just shares, of patent citations that were given along the period 1970-90 to identify knowledge links. Then to diffuse the observed patents filed in the period 1990-1999 to predict the patenting activity in this period. And finally, use this predictions along with the citation network to construct our instrument.

The construction of the instrument follows several steps. First, we collect all the patents filed in the pre-sample period 1970-1990 along with the information about the country, technological field, backward and forward citations, and the sequence of the patent within its family for those patents based on the procedure described in the data section. We use correspondence from technological fields to industry codes to assign each patent one or multiple NACE codes, with respective weights in the latter case.

Second, exploiting the network of patent citations, we build a network of knowledge links.

[31] Formally, predicted number of patents are computed in the following way:

$$d_{c_o,s_o,s_d,t_1} = \frac{\sum_{t'\in[1,T]} \sum_{c_d \neq c_o} citations_{c_o,s_o,t_1 \to s_d,c_d,t_1-t'}}{\sum_{s_d \neq s_o} \sum_{t'\in[1,T]} \sum_{c_d \neq c_o} citations_{c_o,s_o,t_1 \to s_d,c_d,t_1-t'}}$$

and

$$\hat{pat}_{c_o,s_o,t} = \sum_{s_d \neq s_o} d_{c_o,s_o,s_d,t_1} \times \sum_{c_d \neq c_o} pat_{s_d,t}$$

The underlying idea is to proxy the knowledge flows across countries and sectors from the share of citations that each patent produced in the country and sector of origin given to patents in the destination country and sector. For each patent of sector s_o belonging to country c_o at time t, we calculate the share of citations that it gives to patents produced in sector s_d country c_d at time $t-\Delta$. We then sum those shares, and to avoid size effects due to the fact that some locations or sectors tend to patent more for idiosyncratic reasons, we normalize this share by the total number of patents produced there. Our network also takes into account that the speed at which ideas diffuse might differ across locations and sectors. We formally capture this effect by allowing the weights in our network to be time specific. In other words, the strength of the links depends on how many years are between the time cited and citing patents were filed. Formally, the adjacency matrix of the knowledge network is defined as follows:

$$d_{c_o,c_d,s_o,s_d,\Delta} = \begin{cases} 0 & c_o = c_d \\ 0 & s_o = s_d \\ \dfrac{\sum\limits_{t=1970}^{1990} \sum\limits_{p \in (\mathcal{S},\mathcal{N},\mathcal{T})} s_{p \to (c_d,s_d,t-\Delta)}}{\sum\limits_{t=1970}^{1990} \sum\limits_{q} \mathbb{I}_{\{q \in (c_d,s_d,t-\Delta)\}}} & otherwise \end{cases} \qquad \text{for } \Delta \in \{1,\ldots,10\}$$

where $s_{p \to (c_d,s_d,t-\Delta)}$ is the share of citation that patent p gives to patents of sector s_d produced in country c_d at time $t-\Delta$. Note that to avoid endogeneity concerns coming from the fact that edged that link the same geographical area or sector might be correlated with future shocks, we discard citations coming from the same country and from the same sector. In addition, to capture knowledge creation originated in a particular country we restrict our sample to patents that are only the first in their family. Since those patents that are not the first in the family are mostly being filed for protection reasons on other territories than original one and so have a negative association with productivity, including them might downward the estimator of interest. However, for cited patents we include all patents irrespective of their sequence in their family to capture all innovations on which patent is build on.

Finally, we diffuse the observed patents filed in the period 1990-1999 to predict the patenting activity we expect to observe in the other countries and sectors if the pre-determined network of ideas was the only thing that mattered for the production of knowledge.

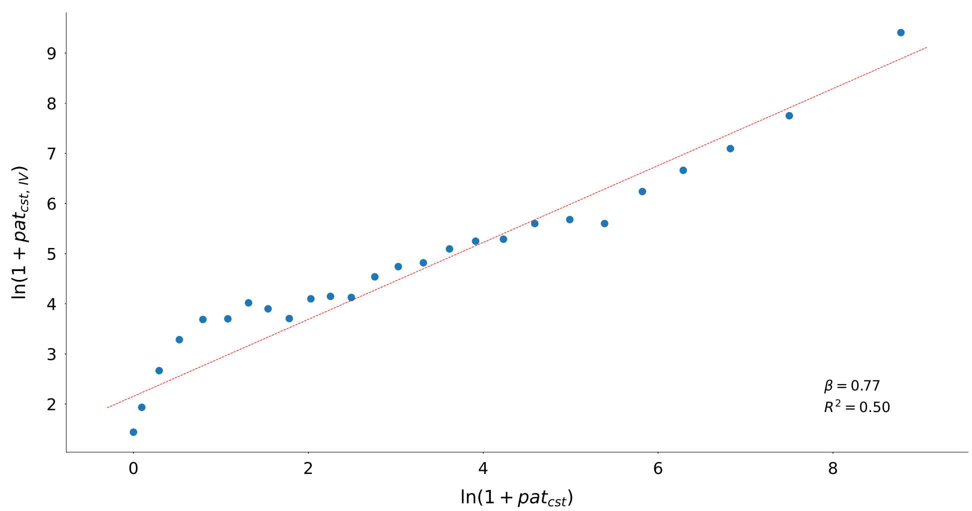
$$\hat{pat}_{c_o,s_o,t} = a_t \sum_{\Delta=1}^{10} \sum_{s_d \in \mathcal{S} \setminus s_o} \sum_{c_d \in \mathcal{N} \setminus c_o} \left(d_{c_o,c_d,s_o,s_d,\Delta}\right) pat_{c_d,s_d,t-\Delta}$$

where a_t is a rescaling term that ensures that estimated number of patents is equal to actual number of patents in period t worldwide, and $pat_{c_d,s_d,t-\Delta}$ is actual number of patents in year $t-\Delta$ filed in country c_d in sector s_d[32]. It is important to mention, that throughout entire instrument construction when number of patents in a given country and given sector is used – country and sector weight within each patent is taken into account.

The intuition behind this approach mirrors the one of an input-output model. In this case, ideas patented in the past are used as inputs and combined to produce new inventions under the assumption that the innovation process remains stable over time. Berkes and Gaetani (2018) show that the network of patents in the United Stated is indeed stable in the time frame they consider.

Figure 7 visually compares actual and predicted number of patents. The two variables are strongly but not perfectly correlated: the coefficient of the regression is 0.77 and $R^2 = 0.50$. The Cragg-Donald Wald F statistics in the benchmark regression is 2,070, which rules out weak instrument concerns.

Figure 7: Unconditional Correlation between Actual and Predicted Patents



This instrument belongs to a family of shift-share instruments: weighted averages of a common set of shocks, with weights reflecting heterogeneous shock exposure. The validity of shift-share instrumental variable regressions must rely on some assumptions about the shocks, exposure shares, or both. Borusyak et al. (2018) and Goldsmith-Pinkham et al. (2020) provide a technical discussion of those assumptions.

To provide evidence in support of the instrument validity in our setting, we test for a number of assumptions. First, the validity of the shift-share instrument rests on the assumption that countries and sectors giving more citations (to other sectors and countries) in the period between

[32]Figure C.1 represents a simple example of described procedure

1970 and 1990 are not on different trajectories for the evolution of TFP in the period of analysis (2000-2014). We test this assumptions in two ways: i) regressing productivity in 1990 against average patent activity in the period of 2000-14 predicted by the instrument, ii) we check that results are unchanged when controlling separately for an average level of patent activity in the period 1970-90 and productivity in 1990.[33]

Second, we rule out the possibility that the links of knowledge diffusion used to construct the instrument capture a demand pull from the destination country and sector, rather than a supply push from the origin country and sector. We start by constructing pre-determined network but now using forward citations instead of backward. Then, we use data from our sample and the network to generate predicted number of patents in period 1970-1990, that is included in the baseline regression as additional control. In other words, these predicted patents are patents that should have been filed in period of 1970-1990 to generate patent activity in the period 2000-14 that we observe in the data.

5.2 Innovation and Productivity

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

Table 2 shows our benchmark estimates of the relationship between TFP, estimated using dual approach and lagged innovation instrumented with predicted innovation.[34] We use average value of our productivity measure in the period of three years to smooth out any business cycle fluctuations. We also include in regressions capital, employment and intermediate imports on a country-sector level as controls. In addition, regressions include country-year and sector-year fixed effect to capture country and sector specific shocks at a given point of time. We cluster our standard errors at a country level. Our benchmark regression used data from years 1970-90 to compute pre-determined network linkages, and the period of our analysis is 2000-2014.

The coefficient on innovation activity is positive and statistically significant. Moreover, the coefficient obtained using instrumental variable approach is almost twice higher than the one using OLS. The regression results suggest that a 1% increase in patenting between 2000 and 2014 leads to 0.016% increase in TFP. Given presence of fixed effects in our regression, it is important to interpret the coefficient as a change in TFP caused by the growth rate of

[33]Since we do not have data on TFP for the period before 2000, we use value added per employment obtained from UNIDO data as a measure of productivity.

[34]Results for productivity measured by TFP, estimated using primal approach, as well as value added per employment are reported in Table C.2 in the Appendix.

Table 2: 2SLS Estimates: 2000-2014

	$\ln(\overline{\text{TFP}_{t+n}})$					
	OLS	IV	OLS	IV	OLS	IV
$\ln(patent_t)$	0.008	0.017	0.007	0.016	0.007	0.016
	(0.005)	(0.006)	(0.006)	(0.007)	(0.005)	(0.007)
$\ln(\text{tfp}_t)$	0.952	0.949	0.967	0.967	0.972	0.972
	(0.011)	(0.011)	(0.015)	(0.015)	(0.018)	(0.018)
$\ln(capital_t)$			-0.026	-0.027	-0.031	-0.031
			(0.007)	(0.007)	(0.010)	(0.010)
$\ln(employ_t)$			0.025	0.023	0.021	0.019
			(0.004)	(0.005)	(0.007)	(0.007)
$\ln(int\_import_t)$					0.008	0.008
					(0.008)	(0.008)
Country-Year FE	Y	Y	Y	Y	Y	Y
Sector-Year FE	Y	Y	Y	Y	Y	Y
# obs.	8,169	8,169	8,169	8,169	8,169	8,169
# countries	36	36	36	36	36	36
	First-stage estimates					
Predicted		0.478		0.471		0.471
$\ln(patent_t)$		(0.068)		(0.081)		(0.081)
CD Wald F		2,070		2,141		2,142

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

innovation activity that is beyond average growth rate of innovations across the world in a given sector and average growth rate of innovation in a country in a given period of time. Bringing those numbers to actual data means that 1 standard deviation increase in increase in annual number of patents leads to 0.1 standard deviations increase in TFP after partialling out fixed effects and controls used in our baseline regression.

The 2SLS estimated are larger than the ones obtained in the OLS regression. This increase is consistent with the likely scenario in which our OLS estimates suffer from attenuation bias because patents are an imperfect measure of innovation activity. Another possible explanation for the bias could be an increase in market concentration–a trend observed in most advanced countries since 2000. Higher market concentration leads to slowdown in productivity, while stimulates innovation activity due to the fact that leader(s) don't want to give up their leading role (Akcigit and Ates, 2021).

Table 3: Checking for Pre-trends

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

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

5.2.1 Robustness Checks

The key identifying assumption behind the instrument can be violated if the characteristics of countries and sectors that give more citations to particular sectors and countries in the period 1970-90 had persistent effects on patent activity as well as on changes in the outcomes of interest (beyond our regression controls). We test this assumption in a variety of ways. First, we test for pre-trends, by showing that the pre-period productivity is uncorrelated with subsequent patent activity predicted by the instrument. Table 3 presents the results of regressing average value of productivity during the pre-sample period against average annual number of patent in period 2000-14.[35] The coefficients of this regression, reported in Columns (3) and (4), are not statistically significant. Importantly, they are quantitatively different from the estimates obtained for the period used in main exercises, reported in Columns (1) and (2).

Second, in Columns (2) and (3) of Table 4, we check that results do hold when we also control for an average level of patent activity in the period 1970-90 and level of productivity in 1990, measured by value added per employment. In the case, when we add separately historical level of productivity, the results are unchanged. However, when we add average level of historical patent activity, the coefficient of interest becomes twice as large (in absolute value).

[35] As a measure of productivity we use value added per employment data as data on TFP for historical periods is not available. We also averaged all the variables in order to suppress the time dimension as the left-hand side and right-hand side of our regression belong to different time periods.

Yet, statistically we can not distinguish it from the baseline level.

Next, we rule out the possibility that the links of knowledge diffusion used to construct the instrument capture a demand pull factors from the destination country and sector, rather than a supply push from the origin. To do that, we include in our baseline regression number of patents that should have been filed in pre-sample period to explain actual number of patents observed in the sample in the period of study given citations linkages in pre-sample.[36] Results presented in Column (4) of Table 4 are very stable and the coefficient remains statistically significant and quantitatively close to the baseline.

We repeat all these robustness checks using two other measures of productivity and obtain similar results. These results are reported in Table C.3 in the Appendix. Finally, to check for outliers driving our results, we show that our results remain unchanged if we exclude one country or sector at a time.[37]

5.3 Innovation and Long-term Development

We extend now our analysis to longer-time periods. One challenge of looking at long-term outcomes is that high quality TFP panel data spanning a large number of countries and sectors is not available. To circumvent this problem, we adapt our empirical strategy to study the relationship between innovation activity and real GDP per capita at the aggregate country level since 1960.[38] That is, we depart in two dimensions relative to our baseline exercise. First, we abstract from setoral variation. Second, we use real GDP per capita rather than TFP as our outcome variable. Since we have shown in the robustness section that our results go through and have a similar magnitude with sectoral value added per worker, we have some confidence of using output per capita as a proxy for productivity.

To obtain our shift-share instrument in this setup, we start from our baseline instrument and then collapse all the sectoral variation, to have only country-time variation. That is, we sum the number of predicted patents within each country and time across all sectors:

$$\widehat{total\_pat}_{c_o,t} = \sum_{s_o \in \mathcal{S}} \widehat{pat}_{c_o,s_o,t}.$$

We note that the instrument is constructed as in our baseline but using different pre-sample

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

[37] The largest change in magintude that we obtain in ϕ_A 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 0.016 to 0.021.

[38] Data for real GDP per capita is from Maddison Project Database (Inklaar et al., 2018)

Table 4: 2SLS Estimates: Robustness

	$\ln(\overline{\text{TFP}_{t+n}})$			
	(1)	(2)	(3)	(4)
$\ln(patent_t)$	0.016	0.018	0.028	0.029
	(0.007)	(0.009)	(0.013)	(0.010)
$\ln(va\_em_{1990})$		0.019		
		(0.010)		
$\ln(\overline{patent_{1970-90}})$			-0.009	
			(0.009)	
$\ln(\widehat{patent}_{t-30})$				-0.010
				(0.007)
Controls	✓	✓	✓	✓
Country-Year FE	Y	Y	Y	Y
Sector-Year FE	Y	Y	Y	Y
# obs.	8,169	6,222	8,169	8,169
	First-stage estimates			
Predicted	0.514	0.520	0.273	0.397
$\ln(patent_t)$	(0.086)	(0.086)	(0.054)	(0.060)
CD Wald F	1,867	1,902	503	803

Notes: Column (1) shows the results of our baseline regression, Column (2) and (3) include separately to baseline regression historical levels of productivity and average patent activity, respectively. Column (4) includes predicted number of patents driven by demand pull factors to the baseline regression. All regressions include (log) values for TFP, capital, employment, and intermediate imports as controls. Standard errors are clustered at a country level in parentheses.

time periods. In particular, we use the pre-1950 data to construct the pre-existing linkages across country-sectors, and country-sector patenting activity during the period 1950-1959 to construct our shift components.

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

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

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

Table 5: 2SLS Estimates: Innovation and Long-term Development: 1960-2016

	Dependent Variable is: $\ln(gdp\_cap_{t+n})$			
	All Countries		HI & UMI Countries	
	OLS	IV	OLS	IV
$\ln(patent_t)$	0.003	0.016	0.005	0.013
	(0.002)	(0.008)	(0.003)	(0.007)
$\ln(gdp\_cap_t)$	0.913	0.884	0.893	0.872
	(0.019)	(0.026)	(0.030)	(0.036)
Country FE	Y	Y	Y	Y
Year FE	Y	Y	Y	Y
# obs.	3,942	3,942	2,558	2,558
# countries	118	118	58	58
	First-stage estimates			
Predicted		-1.704		-1.711
$\ln(patent)$		(0.585)		(0.625)
CD Wald F		66.4		62.4

Notes: Period of the analysis is 1960-2016 using pre-determined matrix based on the data prior 1950. Standard errors are clustered at a country level in parentheses.

specification we use $n = 3$, results for $n = 5$ are similar both qualitatively and quantitatively.

Table 5 shows the results estimated using all available citation and patent data for the period prior 1950 to generate the knowledge network, and the period of the analysis is 1960-2016. Given that comprehensive data for the period prior 1950 are available mostly for advanced countries, we also report the results only for High Income and Upper Middle Income countries based on the World Bank classification.[39] We find a positive, significant coefficient that is very similar in magnitude to the elasticity of patents to TFP that we find for the period 2000-2014. Moreover, this result is mostly driven by high income and upper middle income countries. Bringing those numbers to the actual data implies that one standard deviation increase in annual number of patents leads to 0.28 standard deviations increase in annual GDP per capita for the whole sample and to 0.23 standard deviations for the sub-sample of higher income countries after partialling out fixed effects and controls used the regression.

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

5.4 Innovation and the Trade Elasticity to TFP

The shift-share instrument we propose in the paper can be applied in a variety of other settings. In this section, we illustrate this point by using our instrument to estimate the elasticity of cross-country, cross-sector TFP differences on trade flows. That is, we quantify the importance of Ricardian comparative advantage following the estimating equation derived in Costinot et al. (2012). The only difference relative to Costinot et al. is that we extend the analysis to a panel setting (in addition to use our shift-share instrument, rather their instrument which is R&D expenditures in a given year). As in Costinot et al., the dependent variable is the log of bilateral "corrected exports" disaggregated by sectors and adjusted for openness of a country and a sector (this dependent variable follows from computing trade flows in a standard Ricardian model). The estimating equation is the following specification

$$\ln \tilde{x}_{ijt}^k = \theta \ln z_{it}^k + \delta_{ijt} + \delta_{jt}^k + \varepsilon_{ijt}^k$$

where $\tilde{x}_{ijt}^k$ denotes corrected exports (as discussed above), $\tilde{x}_{ijt}^k = x_{ijt}^k / x_{iit}^k$, z_{it}^k is exporter TFP, δ_{ijt} and δ_{jt}^k importer-exporter-time fixed effects and importer-time-industry fixed effects, respectively. Table 6 documents the results using average corrected exports in the three years period on the right hand side, and TFP measures in the analogous period instrumented by the lagged level of predicted patents on the left hand side.[40] As in Costinot et al. we find that the OLS estimation is downward bias. After instrumenting, the elasticity parameter is around 2.6. This value is somewhat lower than what they find and in the lower range of trade elasticities (but within a plausible range).[41]

6 Conclusion

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

[40] The results reported in Table 6 are for TFP estimated with dual approach, the results for TFP estimated with primal approach are analogous and reported in Table C.4.

[41] As pointed out by Boehm et al. (2020), the estimation of trade elasticities in panel data with the inclusion of time dummies interacted with importer-sector fixed effects and importer-exporter tends to lead to lower trade elasticities.

Table 6: 2SLS Estimates: 2000-2014

Dependent Variable is: $\overline{\text{Adjusted exports}}_{c^{Ex}c^{Im}st+n}$		
	OLS	IV
$\ln(\overline{TFP})_{c^{Ex}st+n}$	0.106	2.554
	(0.211)	(1.144)
Country^{Ex}-Country^{Im}-Year FE	Y	Y
Country^{Im}-Sector-Year FE	Y	Y
# obs.	307,382	307,382
# countries^{Im}	39	39
# countries^{Ex}	36	36
	First-stage estimates	
Predicted		0.074
$\ln(patent_{c^{Ex}st+n})$		(0.035)
CD Wald F		3,989

Notes: Period of the analysis is 2000-2014 using pre-determined matrix based on the data from 1970-90. Standard errors are clustered at a country of imports, country of exports and sector level in parentheses.

accounted for rising citations to US and Japanese patents in fields of knowledge related to computation, information processing, and medicine.

After having documenting these facts, we propose a shift-share identification that leverages the knowledge spillovers across fields of knowledge and countries (to construct a the shift) and the heterogeneity in exposure of countries to technological waves (to construct the share). We then estimate the effect of innovation on TFP in a panel of countries-sectors for the period 2000-2014 using historical patent data spanning 1970 through 2000. On average, an increase in one standard deviation in patents increases TFP by 0.1 standard deviation. We also estimate the effect of innovation on income per capita since 1960. An increase in one standard deviation in patenting activity increases income per capita by 0.28 standard deviation.

References

Acemoglu, D. (2009a): *Introduction to Modern Economic Growth*, Princeton University Press.

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

Acemoglu, D., U. Akcigit, and W. R. Kerr (2016): "Innovation network," *Proceedings of the National Academy of Sciences*, 113, 11483–11488.

Aghion, P., A. Dechezlepretre, D. Hemous, R. Martin, and J. van Reenen (2016): "Carbon Taxes, Path Dependency, and Directed Technical Change: Evidence from the Auto Industry," *Journal of Political Economy*, 124, 1 – 51.

Aghion, P. and P. Howitt (1992): "A Model of Growth through Creative Destruction," *Econometrica*, 60, 323–51.

——— (1998): *Endogenous Growth Theory*, Cambridge, MA: MIT Press.

Akcigit, U. and S. T. Ates (2021): "Ten Facts on Declining Business Dynamism and Lessons from Endogenous Growth Theory," *American Economic Journal: Macroeconomics*, 13, 257–98.

Akcigit, U., S. Baslandze, and F. Lotti (2018): "Connecting to Power: Political Connections, Innovation, and Firm Dynamics," Working Paper 25136, National Bureau of Economic Research.

Akcigit, U., J. Grigsby, and T. Nicholas (2017): "The Rise of American Ingenuity: Innovation and Inventors of the Golden Age," CEPR Discussion Papers 11755, C.E.P.R. Discussion Papers.

Altuntas, S., T. Dereli, and A. Kusiak (2015): "Analysis of patent documents with weighted association rules," *Technological Forecasting and Social Change*, 92, 249–262.

Barro, R. (1991): "Economic Growth in a Cross Section of Countries," *The Quarterly Journal of Economics*, 106, 407–443.

Barro, R., R. Barro, X. Sala-i Martin, X. Sala-i Martin, and M. I. of Technology (2004): *Economic Growth*, The MIT Press, MIT Press.

Barro, R. J. and X. Sala-i Martin (1992): "Convergence," *Journal of Political Economy*, 100, 223–251.

Berkes, E. and R. Gaetani (2018a): "Income Segregation and Rise of the Knowledge Economy," 2018 Meeting Papers 213, Society for Economic Dynamics.

——— (2018b): "Income Segregation and the Rise of the Knowledge Economy," 2018 Meeting Papers 213, Society for Economic Dynamics.

Bloom, N., C. I. Jones, J. Van Reenen, and M. Webb (2020): "Are Ideas Getting Harder to Find?" *American Economic Review*, 110, 1104–44.

Boehm, C. E., A. A. Levchenko, and N. Pandalai-Nayar (2020): "The Long and Short (Run) of Trade Elasticities," Working Paper 27064, National Bureau of Economic Research.

Borusyak, K., P. Hull, and X. Jaravel (2018): "Quasi-Experimental Shift-Share Research Designs," NBER Working Papers 24997, National Bureau of Economic Research, Inc.

Bottazzi, L. and G. Peri (2003): "Innovation and spillovers in regions: Evidence from European patent data," *European Economic Review*, 47, 687–710.

Breschi, S., F. Lissoni, and F. Malerba (2003): "Knowledge-relatedness in firm technological diversification," *Research Policy*, 32, 69–87.

Clark, G. and R. C. Feenstra (2003): "Technology in the Great Divergence," in *Globalization in Historical Perspective*, National Bureau of Economic Research, Inc, NBER Chapters, 277–322.

Coelli, F., A. Moxnes, and K. H. Ulltveit-Moe (2016): "Better, Faster, Stronger: Global Innovation and Trade Liberalization," NBER Working Papers 22647, National Bureau of Economic Research, Inc.

Comin, D. and B. Hobijn (2010): "An Exploration of Technology Diffusion," *American Economic Review*, 100, 2031–2059.

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

Costinot, A., D. Donaldson, and I. Komunjer (2012): "What Goods Do Countries Trade? A Quantitative Exploration of Ricardo's Ideas," *Review of Economic Studies*, 79, 581–608.

Dechezleprêtre, A., D. Hémous, M. Olsen, and C. Zanella (2020): "Automating Labor: Evidence from Firm-Level Patent Data," CEP Discussion Papers dp1679.pdf, Centre for Economic Performance, LSE.

Durlauf, S. N., P. A. Johnson, and J. R. Temple (2005): "Growth Econometrics," in *Handbook of Economic Growth*, ed. by P. Aghion and S. Durlauf, Elsevier, vol. 1 of *Handbook of Economic Growth*, chap. 8, 555–677.

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

Engelsman, E. and A. van Raan (1994): "A patent-based cartography of technology," *Research Policy*, 23, 1–26.

Fu, K., J. Chan, J. Cagan, K. Kotovsky, C. Schunn, K. Wood, and P. Pillar (2012): "The Meaning of "Near" and "Far": The Impact of Structuring Design Databases and the Effect of Distance of Analogy on Design Output," vol. 135.

Goldsmith-Pinkham, P., I. Sorkin, and H. Swift (2020): "Bartik Instruments: What, When, Why, and How," *American Economic Review*, 110, 2586–2624.

Griliches, Z. (1979): "Issues in Assessing the Contribution of Research and Development to Productivity Growth," *Bell Journal of Economics*, 10, 92–116.

——— (1986): "Productivity, R&D, and the Basic Research at the Firm Level in the 1970's," *American Economic Review*, 76, 141–154.

Hidalgo, C., B. Klinger, A.-L. Barabasi, and R. Hausmann (2007): "The Product Space Conditions the Development of Nations," *Science (New York, N.Y.)*, 317, 482–7.

Hornbeck, R. and E. Moretti (2019): "Estimating Who Benefits from Productivity Growth: Direct and Indirect Effects of City Manufacturing TFP Growth on Wages, Rents, and Inequality," IZA Discussion Papers 12277, Institute of Labor Economics (IZA).

Inklaar, R., H. de Jong, J. Bolt, and J. van Zanden (2018): "Rebasing 'Maddison': new income comparisons and the shape of long-run economic development," GGDC Research Memorandum GD-174, Groningen Growth and Development Centre, University of Groningen.

Jaffe, A. (1986): "Technological Opportunity and Spillovers of R&D: Evidence from Firms' Patents, Profits, and Market Value," *American Economic Review*, 76, 984–1001.

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

Jones, C. I. (1995): "R&D-Based Models of Economic Growth," *Journal of Political Economy*, 103, 759–784.

——— (1999): "Growth: With or Without Scale Effects?" *American Economic Review*, 89, 139–144.

Kang, B. and G. Tarasconi (2016): "PATSTAT revisited: Suggestions for better usage," *World Patent Information*, 46, 56–63.

Kay, L., N. Newman, J. Youtie, A. L. Porter, and I. Rafols (2014): "Patent overlay mapping: Visualizing technological distance," *Journal of the Association for Information Science and Technology*, 65, 2432–2443.

Keller, W. (2002): "Geographic Localization of International Technology Diffusion," *American Economic Review*, 92, 120–142.

——— (2004): "International Technology Diffusion," *Journal of Economic Literature*, 42, 752–782.

Keller, W. and S. R. Yeaple (2013): "The Gravity of Knowledge," *American Economic Review*, 103, 1414–1444.

Klenow, P. and A. Rodríguez-Clare (1997): "The Neoclassical Revival in Growth Economics: Has It Gone Too Far?" in *NBER Macroeconomics Annual 1997, Volume 12*, National Bureau of Economic Research, Inc, NBER Chapters, 73–114.

Kogler, D., D. Rigby, and I. Tucker (2013): "Mapping Knowledge Space and Technological Relatedness in US Cities," *European Planning Studies*, 21, 1374.

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

Leydesdorff, L. (2008): "Patent classifications as indicators of intellectual organization," *Journal of the American Society for Information Science and Technology*, 59, 1582–1597.

Leydesdorff, L., D. Kushnir, and I. Rafols (2014): "Interactive overlay maps for US patent (USPTO) data based on International Patent Classification (IPC)," *Scientometrics*, 98, 1583–1599.

Leydesdorff, L. and L. Vaughan (2006): "Co-occurrence matrices and their applications in information science: Extending ACA to the Web environment," *Journal of the American Society for Information Science and Technology*, 57, 1616–1628.

Moretti, E., C. Steinwender, and J. V. Reenen (2019): "The Intellectual Spoils of War? Defense R&D, Productivity and International Spillovers," NBER Working Papers 26483, National Bureau of Economic Research, Inc.

Murata, Y., R. Nakajima, R. Okamoto, and R. Tamura (2014): "Localized Knowledge Spillovers and Patent Citations: A Distance-Based Approach," *The Review of Economics and Statistics*, 96, 967–985.

Nakamura, H., S. Suzuki, I. Sakata, and Y. Kajikawa (2015): "Knowledge combination modeling: The measurement of knowledge similarity between different technological domains," *Technological Forecasting and Social Change*, 94, 187–201.

Nicholas, T. (2010): "The Role of Independent Invention in U.S. Technological Development, 1880–1930," *The Journal of Economic History*, 70, 57–82.

Packalen, M. and J. Bhattacharya (2015): "Cities and Ideas," Working Paper 20921, National Bureau of Economic Research.

Peretto, P. (1998): "Technological Change and Population Growth," *Journal of Economic Growth*, 3, 283–311.

Petralia, S., P.-A. Balland, and D. L. Rigby (2016): "Unveiling the geography of historical patents in the United States from 1836 to 1975," *Scientific Data*, 3, 160074.

Philippe Aghion, Antonin Bergeaud, M. L. and M. J. Melitz (2018): "The Impact of Exports on Innovation: Theory and Evidence," Working papers 678, Banque de France.

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

Segerstrom, P. S. (1998): "Endogenous Growth without Scale Effects," *American Economic Review*, 88, 1290–1310.

von Wartburg, I., T. Teichert, and K. Rost (2005): "Inventive progress measured by multi-stage patent citation analysis," *Research Policy*, 34, 1591–1607.

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

Zitt, M., E. Bassecoulard, and Y. Okubo (2000): "Shadows of the Past in International Cooperation: Collaboration Profiles of the Top Five Producers of Science," *Scientometrics*, 47, 627–657.

Link to Appendix (Click Here)