Journal of Monetary Economics 137 (2023) 128–145

Contents lists available at ScienceDirect

Journal of Monetary Economics

journal homepage: www.elsevier.com/locate/jmoneco

Trade and diffusion of embodied technology: an empirical analysis☆

Stephen Ayerst[a], Faisal Ibrahim[b], Gaelan MacKenzie[c,*], Swapnika Rachapalli[d]

[a] European Department, International Monetary Fund, 3001 11th ST NW Unit 3, Washington, 20001, DC USA
[b] Department of Economics, University of Toronto, 150 St. George St., Toronto, M5S 3G7, ON, Canada
[c] International Economic Analysis Department, Bank of Canada, 234 Wellington St. W, Ottawa, K1A 0G9, ON, Canada
[d] Sauder School of Business, University of British Columbia, 2053 Main Mall, Vancouver, V6T 1Z2, BC, Canada

ARTICLE INFO

Article history:
Received 29 April 2023
Accepted 2 May 2023
Available online 5 May 2023

JEL classification:
O33
F14
O31
O19
F61

Keywords:
Research spillovers
Technology diffusion
Trade
Patents
Innovation

ABSTRACT

Using global patents, citations, inter-sectoral sales, and trade data, we examine the international diffusion of technology through imported inputs. We use citations and sales data to characterize knowledge and production input-output tables for individual countries. Using these tables, we construct a measure of the flow of knowledge-weighted and production-weighted technology embodied in inputs imported from the US. We develop an instrumental variable strategy to establish that increases in embodied technology imports lead to increased innovation and knowledge diffusion in sectors within importing countries. Effects are substantially larger for knowledge-weighted imports of embodied technology.

1. Introduction

Innovation and R&D activity are concentrated in a relatively small number of advanced economies. Recent work demonstrates the quantitative importance of international technology diffusion for the gains from trade and aggregate growth (see, for example, Buera and Oberfield, 2020; Cai et al., 2022; Sampson, 2023). However, little direct empirical evidence exists on the significance of specific channels through which ideas spread across borders. In this paper, we examine the diffusion of technology across countries and sectors through technology embodied in imports of goods from the US using evidence from global patents and citations data.

☆ We thank our discussants, Anna Ignatenko and Simone Lenzu, and participants at many conferences and seminars for helpful comments on this paper. We have also benefited from feedback from Murat Celik, Kevin Lim, Peter Morrow, Diego Restuccia, and Daniel Trefler. The views expressed herein are those of the authors and do not necessarily represent those of the Bank of Canada or its Governing Council, nor of the International Monetary Fund, its Executive Board, or its management.

* Corresponding author.
E-mail addresses: stephen.b.ayerst@gmail.com (S. Ayerst), faisal.ibrahim@mail.utoronto.ca (F. Ibrahim), gmackenzie@bankofcanada.ca (G. MacKenzie), s.rachapalli@sauder.ubc.ca (S. Rachapalli).

https://doi.org/10.1016/j.jmoneco.2023.05.002
0304-3932/

We focus on this channel for three reasons.[1] First, new innovations often manifest themselves as new products or enhancements to existing products and many of these new or enhanced products are traded. These product flows potentially convey information about the innovations embodied within them to the users of the products. Second, the foundational knowledge on which new innovations are based originates from many distinct sectors; these sources vary across sectors and need not be related to sectors' sources of production inputs. Since countries' patterns of trade depend in part on patterns of comparative advantage, their imports of technology embodied in trade flows affect innovation in different sectors in those countries in different ways. Incorporating variation in sectors' sources of knowledge and production inputs is necessary to assess the impacts of a given amount of technology embodied in a set of trade flows on different sectors. Third, accounting for technology embodied in traded inputs has important policy implications. The effects of trade policies go beyond the well-studied impacts of tariffs on, for example, static intermediate and final goods prices, since they can also affect the flow of information and technology across countries and sectors. Because new innovations often build on existing knowledge, changes in technology flows due to changes in trade policy can have effects on innovation activities not accounted for by the policy-induced responses of innovation to import competition and market access.[2]

The first contribution of our paper is to estimate the extent to which trade is a channel of international technology diffusion. We do this by investigating the effects of embodied technology imports on innovation and diffusion outcomes. This channel underlies many theoretical and quantitative models of international technology spillovers (e.g., Alvarez et al., 2013; Buera and Oberfield, 2020; Grossman and Helpman, 1991). We start by developing a conceptual framework to guide our empirical analysis. In the conceptual framework, firm innovations depend on a combination of R&D investment, domestic knowledge spillovers, and international spillovers from the technological frontier. International spillovers depend on embodied technology imports—the import-weighted stock of frontier knowledge—and the relevance of cross-sector knowledge to the innovating sector.

We use patents data as our primary measure of innovation in our analysis. Patents document innovations that result in new products, new components of existing products, or new methods of producing products. The second contribution of our paper is to construct a novel dataset on country-sector-level innovations and trade. We leverage the Google Patents database to construct detailed patent outcomes for a wide range of countries. In particular, the database allows us to construct measures of patenting based on the locations of innovators and measures of cross-country citation flows. We also use import data from the Centre d'Études Prospectives et d'Informations Internationales (CEPII) database of international trade flows and cross-sector sales from the Bureau of Economic Analysis (BEA) in our analysis. Finally, we map data into consistent sector definitions using a series of concordances.

The third contribution of our paper is to construct empirical measures of embodied technology imports. As a first step, we construct measures of the cross-sector relevance of knowledge. We use cross-sector citations and sales data to construct knowledge and production input-output (IO) tables as a measure of the relevance of cross-sector knowledge. Within a country, we construct the knowledge IO table using the share of citations from each sectors' patents to each other sectors' patents. We similarly construct the production IO table using the share of sales between sectors. Unlike with patents, data to construct the production IO table is only available for the US, which we take as the frontier economy in our analysis. Since knowledge and production IO linkages could, in principle, be similar for many sectors, we demonstrate that the US knowledge and production IO tables are distinct.[3] In particular, we document that knowledge and production IO linkages are not highly correlated on average, that knowledge IO linkages are less concentrated than production IO linkages for the average sector, that the sectors that are key economy-wide sources of inputs differ between the knowledge and production IO tables, and that the sources of knowledge inputs are more persistent than the sources of production inputs. We also find that knowledge IO linkages are less persistent on average in sectors in non-US countries than in the US.

Along with US imports, we use the IO tables to develop measures of embodied technology imports. We develop two measures based on the knowledge and production IO tables that we refer to as the knowledge-weighted and production-weighted embodied technology imports. Specifically, we aggregate US import-weighted knowledge stocks using Cobb-Douglas weights from the knowledge and production IO tables. We exclude the own-sector component in the construction of embodied technology as imports and innovation activity in a country-sector can be correlated due to domestic demand shocks or import competition effects. The knowledge-weighted measure is directly related to our mechanism of interest since it relies on knowledge flows across sectors. The production-weighted measure is also included as potentially important transfers of technology can occur through production interactions. A key outcome of our analysis is then to measure the relative strength of spillovers from embodied technology weighted by knowledge and production linkages.

Our main empirical specification involves regressing measures of innovation and diffusion outcomes on knowledge-weighted and production-weighted embodied technology imports. The main innovation outcomes are patents, forward citations, and forward citations per patent while our main diffusion outcomes are US backward citations, US backward cita-

[1] Other channels include technology licensing, foreign direct investment, knowledge transfers within multinational firms, immigration, trade in services, and cross-border scientific or technical collaborations (see Keller, 2004; 2010; 2021, for surveys of empirical evidence of different channels).

[2] Shu and Steinwender (2019) survey the empirical literature examining evidence of the effects of import competition and market access on innovation. Existing work that, like us, focuses on effects that are present in patents data includes Bloom et al. (2016), Bombardini et al. (2017), and Autor et al. (2020) for import competition, and Coelli et al. (2022) and Aghion et al. (2022) for market access.

[3] Though not the focus of our paper, we are among the first to provide a descriptive comparison of the knowledge and production IO tables of an economy. Concurrent work in Hötte (2023) and Liu and Ma (2023) construct similar knowledge and production IO tables and compare them.

tions per patent, and the US backward citation share.[4] We also include controls for the US knowledge stock, each country's own knowledge stock—constructed using country-specific knowledge IO linkages—and own-sector imports. Additionally, our long panel of data, spanning from 1995 through 2015, allows us to control for high-dimensional fixed effects. We include country-sector fixed effects to account for differences in the propensity to patent and innovate across countries and sectors, country-year fixed effects to account for changes in patenting markets and development within countries over time, as well as aggregate sector-year fixed effects to control for structural changes in sectoral technologies over time.

A potential concern with estimating the effects of the trade channel of technology diffusion on domestic innovation is that domestic shocks to expected future profits or shocks to R&D productivity can lead to R&D activity in a country-sector and the demand for production and knowledge intermediate inputs being correlated.[5] To address this concern, we use an instrumental variable (IV) strategy to isolate the effects of embodied technology imports on innovation and diffusion outcomes. For each country, we construct a cluster of related countries that fall into the same quintiles of the distributions of the total trade (exports plus imports) to GDP ratio and GDP per capita. We then construct the instrument for each country using US exports to all countries outside of the country's cluster. The instrument isolates US supply shocks by excluding countries that are likely to experience correlated demand shocks.[6]

Using our IV strategy, we find that a 1% increase in the knowledge-weighted embodied technology imports increases patenting by around 0.67%. In comparison, a 1% increase in the production-weighted embodied technology imports increases patenting by 0.003%. To quantify the size of the of the estimated coefficients, we show that a one standard deviation increase of residualized knowledge-weighted and production-weighted embodied technology imports accounts for 13.2% and 1.8% standard deviations of residualized patenting, respectively. The considerably larger estimate of the knowledge-weighted measure is consistent with our expectation that the knowledge IO table better approximates the relevance of knowledge across sectors.

For diffusion outcomes, we find that a 1% increase in embodied technology imports increases US backward citations by a similar proportion for the knowledge-weighted measure and 0.005% for the production-weighted measure. Knowledge-weighted and production-weighted embodied technology imports account respectively for 8.8% and 1.3% of the standard deviation of residualized US backward citations. Despite the elasticity for US backward citations being larger than that for patenting, we do not find consistent evidence that either measure of embodied technology imports increases the rate of US backward citations (US backward citations per patent) or the share of US backward citations (out of total backward citations in a country-sector-year). We expect that foreign backward citations are noisier measures than patenting outcomes and find that the rate of US backward citations becomes positive and statistically significant in some of our robustness exercises.

Our estimated coefficients are robust to a variety of alternative specifications. We find similar coefficient estimates for different lags of the regressors. We consider alternative instruments constructed using a traditional leave-one-out approach and constructed using all other countries within a cluster (as opposed to all countries outside of the cluster as in our baseline). We find economically more significant results when we restrict the sample to the 40 countries with the most patenting activity, consistent with the idea that patenting activity is better measured in these countries or more representative of innovation activity. Finally, we find similar results using alternative constructions of the main variables, alternative innovation and diffusion outcomes, or other controls.

Related Literature. Our work contributes to the empirical literature on the channels of international technology diffusion (most recently surveyed by Keller, 2021), particularly papers that examine the trade channel. This includes work pioneered by Coe and Helpman (1995) and the within-sector analysis of R&D diffusion across borders through both trade and non-trade channels in Acharya and Keller (2009). Our focus on direct evidence for diffusion using citations in new patents is closely related to MacGarvie (2006) and concurrent work by Aghion et al. (2021), both of which use French firm-level data on the extensive margins of trade participation to show that citations to firms' patents increase in foreign markets with which firms interact through trade. We add to this body of evidence by showing with a sector-level analysis that embodied technology imports are a source of technology diffusion.

In doing so, our paper provides evidence for the international technology diffusion that underlies recent growth models featuring trade, diffusion, and innovation (e.g., Buera and Oberfield, 2020; Cai et al., 2022; Sampson, 2023). Most closely related is Cai et al. (2022) who examine inter-sectoral and cross-country technology diffusion. In their model, technology diffuses exogenously within and across borders based on parameters estimated using citation linkages from the US Patent and Trademark Office. We show that this diffusion across countries depends on endogenous trade flows between countries.

The empirical approach we take to evaluate the effects of diffusion of technology across countries is complementary to recent work using patents data to measure international technology diffusion through inter-sectoral networks, including Fons-Rosen et al. (2019), Berkes et al. (2022), and Liu and Ma (2023).[7] To the best of our knowledge, ours is the first paper to include inter-sectoral knowledge IO measures based on these data to estimate the trade channel of technology diffusion.

[4] Forward citations are measured over a five-year period to mitigate truncation issues. US backward citations are measured as the total citations of US patents by all patents applied for in a given country-sector-year.

[5] Data on R&D spending at the level of industry disaggregation used in our analysis is unavailable for most countries in our sample.

[6] The commonly-used leave-one-out instrument can be viewed as a case of this strategy in which each cluster includes only a single country. Our results are robust to using this instrument instead.

[7] We also build on work that documents and examines the effects of inter-sectoral patterns of knowledge flows in single-country settings, such as Acemoglu et al. (2016) and Cai and Li (2019).

Fons-Rosen et al. (2019) use patents-based sector-pair measures of technological similarity adapted from Bloom et al. (2013), which are distinct from our citations-based IO measures, to investigate the foreign direct investment channel of technology diffusion. Berkes et al. (2022) show that there has been a large increase in international knowledge spillovers since the 1990s as measured by cross-country patent citations and that the innovations induced by this increase in diffusion lead to an increase in the growth rates of sectoral output per worker and total factor productivity. Closely related is the empirical exercise in Liu and Ma (2023) that documents that global spillovers from past patenting activity that depend on the network of patent citations across countries and sectors lead to increases in innovation.

Our paper is also related to the branch of the trade literature examining the effects of changes in access to intermediate production inputs due to trade policy on many dimensions of firm performance. This line of research includes work that shows that increased openness to trade of production inputs leads to higher productivity (Amiti and Konings, 2007; Topalova and Khandelwal, 2011), increases in product scope and new product introduction (Goldberg et al., 2010), and reductions in marginal costs (De Loecker et al., 2016).[8] Though our analysis is conducted at the sector level rather than the firm level, our results speak directly to the mechanisms through which trade in inputs leads to improvements in performance and suggest that technology diffusion and increases in the generation of new patented technology follow from increases in embodied technology imports.

Outline. The remainder of this paper proceeds as follows. Section 2 describes the data used in our analysis. Section 3 presents the conceptual framework used to guide our empirical analysis. Section 4 describes the constructions of the knowledge and production IO tables. Section 5 describes our empirical strategy and baseline specifications. Section 6 discusses the estimation results and robustness checks. Section 7 concludes.

2. Data

In this section, we provide an overview of the data used for the main analysis. We use data on patent applications and citations, inter-sectoral purchases of inputs by US sectors, and bilateral product-level trade flows from the US into other countries. These data come from a variety of sources and are provided in a range of distinct classifications that compel us to use concordance tables to translate all the data into a consistent classification system. We briefly describe the data and concordances we use below and leave the remaining details of the data collection and variable construction to Online Appendix B.

Patents and citations data. We draw on data collected by Google Patents from a wide range of patent offices around the world. For each distinct patent family, which comprises the set of patent applications for a given innovation filed at one or more patent offices, we identify the earliest date a patent was applied for at any patent office and treat this as the filing date for the patent family. Each application in a patent family contains the following information that we use in our analysis: the technology categories to which the innovation is relevant, which are represented by International Patent Classification (IPC) codes; the set of inventors of the patent application and their countries of residence; and citations to other patents listed in the patent application.[9] Throughout our analysis, we focus on patent applications rather than patent grants as grant dates are unavailable in the Google Patents database for patents applied for at many national patent offices, whereas application dates are available.[10] Furthermore, as we examine technology diffusion and its effects, patent application events better reflect the timing of diffusion than do patent grant events.

We calculate the number of initial applications of patent families filed in each year between 1995 and 2015 in each country and technology subclass (a 4-character IPC code) and refer to these as patent counts.[11] Patents are assigned to countries using fractional counts by computing the share of inventors of each patent from each country.[12] For a subset of patent families, applications are submitted to the three patent offices that throughout our sample period are of global significance, including the European Patent Office (EPO), the Japan Patent Office (JPO), and the United States Patent and Trademark Office (USPTO). We count the number of such triadic patent applications.[13]

In addition to counts of patent families, we use information on citations between patents. To measure the quality of patents filed in each year and each country and technology subclass, we compute the number of citations received by these patents across citing patents applied for each year from 1995 to 2021 in all countries and technology classes and define these as the forward citations of the patents in each year. Backward citations data are used for two purposes. First, as described in Section 4.1, we use backward citations to measure knowledge linkages between sectors. Second, for patents filed each year and in each non-US country and technology subclass, we calculate the number of backward citations to US patents, domestic patents, and other foreign patents filed in each technology subclass in each year.

8 See also the other relevant works surveyed in Shu and Steinwender (2019).

9 We focus our analysis on those patent families with non-missing data for each of these three sets of information. Online Appendix B explains how we select information on these attributes from among the patent applications in a family.

10 For instance, there are no grant dates available for patents filed at the Israel Patent Office.

11 For families with multiple IPC codes, we count these patents once for each technology subclass.

12 Using information on the countries of the inventors rather than the patent office of the initial application of a patent family allows us to account for innovations developed in one country for which patent protection is first sought in another country. The sample used in our baseline analysis includes data from 82 countries.

13 We also include patents applied for at the JPO, the USPTO, and at the patent offices of France, Germany, and the United Kingdom. This definition of triadic patent families is consistent with the methodology described by Dernis and Khan (2004).

Inter-sectoral input purchases. To measure production input-output relationships, we employ the Bureau of Economic Analysis (BEA) Supplementary Use Tables. These tables are available at five-year intervals and provide the value of purchases by input sector made by US output sectors based on the most up-to-date US industrial classification in use at the time. We use tables that span from 1992 to 2007. Sector classifications are based on US Standard Industrial Classification (SIC) codes for the 1992 Use table, while in more recent vintages they are based on the North American Industry Classification System (NAICS). We describe how we convert the data based on the various SIC and NAICS classifications into a consistent classification in Online Appendix B. The BEA Use tables not only cover a long period of time, they are available at a high level of disaggregation compared to alternative sources of inter-sectoral sales data. Moreover, using US data enables us to examine how sectors in importing countries are affected by the technology embodied in imports of production inputs from the US based on the patterns of how those inputs are used in the US.

Bilateral trade data. Import data from CEPII's Base pour L'Analyse du Commerce International (BACI) database provide the value of imports of different goods from the US into each country. Our analysis uses annual data from 1995 to 2015. Import values are denominated in current US dollars that we convert to constant 2010 US dollars using CPI deflators taken from the OECD. Goods are classified using 1992 Harmonized System (HS) codes at the 6-digit level of disaggregation.

Concordances between classifications. Because the raw data underlying our analysis are categorized using different classification systems, we employ multiple concordances between these classifications to provide a coherent framework for analysis. We choose the most disaggregated sectors in the 2002 BEA data as our endpoint classification system. This classification, in which sectors are defined similarly to those in the 2002 US 6-digit NAICS classification, allows us to retain a high degree of disaggregation in our analysis while avoiding the potential problems that would arise in a crosswalk of our inter-sectoral input purchase data from the BEA sectors into the more numerous HS goods categories.[14]

We implement a concordance methodology that enables us to first construct measures of technology embodied in goods at the same level of disaggregation as the imports data and second to measure the flow of technology embodied in goods imported from different US sectors. The data downloaded from the Google Patents database are classified into different IPC version 8 4-character technology subclasses.

For the first stage, we convert the data on patent counts, forward citations, stocks of knowledge (the measurement of which we describe in Section 5.2), and backward citations between technology subclasses into categories of goods.[15] To do this, we use the concordance developed by Lybbert and Zolas (2014) between technology subclasses and 2002 6-digit HS codes and then crosswalk this data to 1992 6-digit HS codes. This first concordance is based on an algorithm that uses keywords extracted from the 2002 HS code descriptions that are matched with the text of patent titles and abstracts to construct probabilistic links between the IPC technology subclasses of the matched patents and the HS goods categories.[16]

In the second stage, a series of crosswalks between 1992 HS codes and our endpoint 2002 BEA classification that provide us with weights used to map goods into sectors is overlaid on the knowledge stocks, patents, citations, and trade data. The crosswalks used are the following: first from 1992 6-digit HS codes to 1987 4-digit Standard Industrial Classification (SIC) codes, second from 1987 4-digit SIC codes to 2002 6-digit NAICS codes, and third from these NAICS codes into the 2002 BEA classification. In applying the first two of these crosswalks, mappings from 1992 HS codes to 2002 NAICS codes use weights derived from the earliest available breakdown of employment by 2002 6-digit NAICS sector from County Business Patterns (CBP) data.[17] Similar procedures that leverage CBP-based employment weights are used to crosswalk the data underlying the different vintages of the BEA Use tables into the 2002 BEA sector categories.

3. Conceptual framework

Before turning to our empirical analysis, we describe a stylized conceptual framework to guide our analysis. Time is discrete and indexed by t. The economy is populated by a unit mass of identical firms in each sector of each country. Because firms are identical, we refer to them by their country-sector-year (i, h, t) to simplify notation. To be consistent with our data structure and the empirical approach described in Section 5, we define three levels of sectoral aggregation, where we denote n as a summary sector (the highest aggregation), h as a sector (the focus of our analysis), and p as a subsector (or product). We also define $\mathcal{P}^h$ as the set of subsectors p in sector h and $n(h)$ as the summary sector n that contains sector h.[18]

[14] There are no publicly available sources of data on input-output relationships across goods categorized by disaggregated HS codes. The analysis sample used in our baseline specifications includes 292 sectors.

[15] See Online Appendix B for the procedure we use to calculate citations between technology categories.

[16] Related papers that use the concordances introduced by Lybbert and Zolas (2014) and extended to other classifications in Goldschlag et al. (2020) include Kukharskyy (2020) and Hötte (2023), among others. Kukharskyy (2020) uses the concordances with citations data to construct cross-sector knowledge linkages, but applies these linkages to investigating how the applicability of multinational parent firms' knowledge capital for a foreign affiliate affects the ownership stake (the degree of integration) of the parent firm in its affiliate. Hötte (2023) also constructs inter-sectoral knowledge linkages and combines them with production linkages to explore how different network characteristics of the knowledge and production IO tables are associated with the level and growth of US sector-level output and patenting.

[17] The details of this procedure and links to the sources of all concordances used in this paper are provided in Online Appendix B.

[18] $\mathcal{P}^h$ can be thought of as the set of products that are associated with a sector h. Our raw trade data is collected at the product level and we first convert our patents data to this level of aggregation. We introduce the summary sector $n(h)$ since we use fixed effects at this level of aggregation in our empirical application. These details are further explained in Section 5.

Firms in each country produce innovations by investing in R&D, denoted by $R_{i,t}^h$, to earn future profits $\pi_{i,t+1}^h$ per innovation in the following period.[19] Expected profits in period $t+1$ can be written as $\mathbb{E}_t[\pi_{i,t+1}^h] = \bar{\pi}_{i,t} \times \bar{\pi}_t^{n(h)} \times \bar{\pi}_i^h \times e^{u_{i,t}^h}$ where $u_{i,t}^h$ is an independent and identically distributed random variable that is known to firms in period t. We use a broader sector aggregation for profits in $\bar{\pi}_t^{n(h)}$ to be consistent with our empirical specification. A firm (i, h, t) that invests $R_{i,t}^h$ into R&D produces innovations in the next period at rate

$$X_{i,t+1}^h = \left(\frac{R_{i,t}^h}{\psi_{i,t}^h}\right)^{\frac{1}{\zeta}} \left(Z_{i,t}^h S_{i,t}^h\right)^{1-\frac{1}{\zeta}},$$

where $\psi_{i,t}^h$ governs the relative cost of R&D across country-sector-years, $Z_{i,t}^h$ is the domestic stock of relevant knowledge for sector h, and $S_{i,t}^h$ is a spillover from the frontier of knowledge (described below). The R&D cost parameter is equal to $\psi_{i,t}^h = \psi_{i,t} \times \psi_t^{n(h)} \times \psi_i^h \times e^{\nu_{i,t}^h}$, where $\nu_{i,t}^h$ is an independent and identically distributed random variable that, like $u_{i,t}^h$, is known to firms in period t. The variable $X_{i,t}^h$ can be interpreted as the average quality of innovations or the quality-adjusted rate of innovations.

Domestic knowledge $Z_{i,t}^h$ depends on the stocks of knowledge in different sectors of the domestic economy and the relevance of those stocks of knowledge as inputs into innovation for the innovating sector h. Domestic knowledge is given by

$$Z_{i,t}^h = \prod_l \mathcal{G}_Z\left(\sum_{p \in \mathcal{P}^l} K_{i,t}^p\right)^{\kappa_{i,t}^{l,h}},$$

where $\mathcal{G}_Z(\cdot)$ is a monotonic function that dictates the strength of spillovers from domestic knowledge in an input sector, which we set to $\mathcal{G}_Z(x) = (1+x)^{\eta_Z}$.[20] The other variables are the knowledge stock $K_{i,t}^p$ of country-subsector-year (i, p, t) and the relevance of knowledge from sector l for producing innovations in country-sector-year (i, h, t), denoted by $\kappa_{i,t}^{l,h}$.

Spillovers from the frontier economy depend on the stocks of knowledge embodied in traded goods coming from the frontier economy. The sectoral flow of knowledge coming into sector h from other sectors l depends on a Cobb-Douglas aggregator given by

$$S_{i,t}^h = \prod_l \mathcal{G}_S\left(\sum_{p \in \mathcal{P}^l} (m_{F,i,t}^p \times K_{F,t}^p)\right)^{\gamma_{F,t}^{l,h}},$$

where $\mathcal{G}_S(\cdot)$ is a monotonic function that dictates the strength of spillovers from the embodied frontier technology that we set to $\mathcal{G}_S(x) = (1+x)^{\eta_S}$. The value of $\gamma_{F,t}^{l,h}$ captures the usefulness of knowledge from sector l to sector h in the frontier economy in period t. We allow for this parameter to change over time to capture dynamics in knowledge linkages over time. We also assume that the relevance of frontier knowledge in sector l for innovating in sector h is determined in the frontier economy, whereas the relevance of domestic knowledge for innovating is specific to the domestic economy. This could be thought of as reflecting how different types of goods result in knowledge spillovers to different countries.

The spillover of ideas from the frontier economy depend on two components: $K_{F,t}^p$ is the frontier stock of knowledge in subsector p and $m_{F,i,t}^p = M_{F,i,t}^p / Y_{i,t}^{h(p)}$ is the imports from the frontier economy to the domestic economy $M_{F,i,t}^p$ divided by the subsector output $Y_{i,t}^h$. Unlike with domestic knowledge, we scale frontier knowledge by the relative abundance of frontier goods in the domestic economy, as measured by $m_{F,i,t}^p$. Intuitively, embodied technology can be thought of as capturing the probability that a domestic innovator encounters a frontier good by chance multiplied by the probability that the innovator realizes a new innovation given the encounter (as in, for example, Bloom et al., 2013; Buera and Oberfield, 2020; Lucas and Moll, 2014; Perla and Tonetti, 2014). Whenever $\gamma_{F,t}^{l,h} > 0$ and $\eta_S > 0$, increased domestic abundance of frontier goods in subsectors $p \in \mathcal{P}^l$ (higher $m_{F,i,t}^p$) and increased technology embodied in those goods (higher $K_{F,t}^p$) both increase innovation in sector h.

The problem of a firm is to maximize net expected profits by choosing R&D expenditure. Equivalently, the firm's problem can be written as choosing the innovation rate

$$X_{i,t+1}^h = \arg\max_X X\pi_{i,t+1}^h - \psi_{i,t}^h X^\zeta \left(Z_{i,t}^h S_{i,t}^h\right)^{1-\zeta},$$

[19] We simplify the environment by assuming that firms only earn profits in the next period, but the model would be equivalent if firms earned a stream of profits proportional to expected profits.

[20] This specification of $\mathcal{G}_Z(x)$ is consistent with our treatment of zeros in the empirical analysis.

where the second term is the R&D cost paid by the firm for a given innovation rate. Solving the problem implies that firms innovate at rate

$$X^h_{i,t+1} = \tilde{\zeta} S^h_{i,t} \times Z^h_{i,t} \times \left[\frac{\bar{\pi}_{i,t}}{\psi_{i,t}} \times \frac{\bar{\pi}^{n(h)}_t}{\psi^{n(h)}_t} \times \frac{\bar{\pi}^h_i}{\psi^h_i} \times e^{u^h_{i,t} - \nu^h_{i,t}} \right]^{\frac{1}{\zeta-1}}, \tag{1}$$

where $\tilde{\zeta} = \zeta^{-1/(\zeta-1)}$. Taking the log of Eq. (1) and grouping variables implies

$$\ln X^h_{i,t+1} = \ln S^h_{i,t} + \ln Z^h_{i,t} + f_{i,t} + f^{n(h)}_t + f^h_i + \epsilon^h_{i,t},$$

where $f_{i,t} = (\ln \bar{\pi}_{i,t} - \ln \psi_{i,t})/(\zeta - 1)$, $f^{n(h)}_t = (\ln \bar{\pi}^{n(h)}_t - \ln \psi^{n(h)}_t)/(\zeta - 1)$, $f^h_i = (\ln \bar{\pi}^h_i - \ln \psi^h_i)/(\zeta - 1)$, and $\epsilon^h_{i,t} = (u^h_{i,t} - \nu^h_{i,t})/(\zeta - 1)$. Substituting our functional forms for the domestic and frontier spillovers, $\mathcal{G}_S$ and $\mathcal{G}_Z$, we get:

$$\ln X^h_{i,t+1} = \eta_S \ln EmbTech^h_{i,t} + \eta_Z \ln OwnTech^h_{i,t} + f_{i,t} + f^{n(h)}_t + f^h_i + \epsilon^h_{i,t}, \tag{2}$$

where

$$EmbTech^h_{i,t} = \prod_l \left(1 + \sum\nolimits_{p \in \mathcal{P}^l} (m^p_{F,i,t} \times K^p_{F,t}) \right)^{\gamma^{l,h}_{F,t}}$$

$$OwnTech^h_{i,t} = \prod_l \left(1 + \sum\nolimits_{p \in \mathcal{P}^l} K^p_{i,t} \right)^{\kappa^{l,h}_{i,t}}.$$

The expression in Eq. (2) provides the foundation for our empirical strategy. The parameters, η_S and η_Z, are the elasticities of the spillover functions $\mathcal{G}_S$ and $\mathcal{G}_Z$ respectively. The rest of the paper focuses on identifying these elasticity parameters which modulate the impact of frontier and domestic knowledge spillovers on innovation outcomes in a country-sector. In the next sections, we construct variables that correspond to the values of *EmbTech* and *OwnTech*, which stand for "embodied technology imports" and "own-country technology", respectively.

The conceptual framework highlights the relationship between embodied technology imports and innovation outcomes. The assumptions on the nature of expected profits and investment costs are relatively flexible and capture many macroeconomic differences across countries and sectors that may otherwise be of concern in estimating the relationship. This would include, for example, country-specific business cycles, sector-specific trends, such as digitalization, and time-invariant differences in the comparative advantage of countries across different sectors. However, difficulties may arise if there are persistent country-sector-specific shocks that drive both an increase in imports and innovation. To deal with these issues, we separate the own-sector and cross-sector effects, since we expect these issues to be most severe within sectors, and we develop an IV strategy. These remedies are discussed in detail in Section 5.

4. Inter-sectoral technology linkages

We use the conceptual framework as a roadmap for the empirical analysis. We start by developing measures of the inter-sectoral relevance of technology inputs. For $\kappa^{l,h}_{i,t}$, which parameterizes domestic knowledge spillovers in country-sector-year (i, h, t) from sector l, we base this on patent citation relationships between those sectors. Spillovers from technology embodied in imports from the frontier may depend on both the inter-sectoral knowledge flows captured by these citations as well as patterns of inter-sectoral input purchases, which we measure using production relationships between sectors, so we allow $\gamma^{l,h}_{F,t}$ to depend on both types of relationships. We also use this section to highlight key differences between the knowledge and production IO tables that are constructed using these relationships to shed light on how we separately identify the effects of imported embodied technology that operate through these two channels.

4.1. Construction of knowledge and production IO tables

Our analysis estimates the effects of embodied technology imports on patenting outcomes. We focus on two candidates to describe the relevance of knowledge in each sector for generating innovations in other sectors. The first is *knowledge input-output linkages*, which describe the relative flow of patent citations across sectors. This measure is tightly linked with our focus on innovation outcomes since patent citations represent a direct report of flows of technology. The second is *production input-output linkages*, which describe the relative flow of intermediate inputs across sectors. While less directly linked to innovation outcomes, the use of intermediate inputs captures another channel through which technology can diffuse within and between sectors. We collect these measures of technology relevance into separate knowledge and production IO tables and document patterns of inter-sectoral technology flows.

We denote the number of citations of country-sector-year (j, l, s) patents by country-sector-year (i, h, t) patents as $Cites^{l,h}_{j,i,s,t}$. This variable captures the reported flow of knowledge from (j, l, s) to (i, h, t).[21] The set of sectors is denoted by $\mathcal{H}$ and the set of countries by $\mathcal{I}$.

[21] Similarly to the allocation of patents to countries, we weight each citation by the product of the cited and citing patents' fractional country weights based on their respective inventor country compositions. In this notation, each year refers to the filing year of the relevant patents.

Knowledge IO linkages, which measure the relevance of knowledge produced in each input (cited) sector for each output (citing) sector, are constructed using the backward citations made by patents. More specifically, let $\kappa_{i,t}^{l,h}$ denote the knowledge IO linkage between sectors l and h in country i in year t. We allow for this relationship to change over time and base the relationship in year t on patents filed between years $t-\bar{\tau}$ and t for some chosen lag $\bar{\tau}$. The knowledge IO linkage is given by

$$\kappa_{i,t}^{l,h} = \frac{\sum_{j\in\mathcal{I}} \sum_{\tau=0}^{\bar{\tau}} \sum_{s=0}^{t-\tau} Cites_{j,i,s,t-\tau}^{l,h}}{\sum_{k\in\mathcal{H}} \sum_{j\in\mathcal{I}} \sum_{\tau=0}^{\bar{\tau}} \sum_{s=0}^{t-\tau} Cites_{j,i,s,t-\tau}^{k,h}}. \tag{3}$$

In our analysis, we set the maximum lag used in the construction of the knowledge IO linkages to a ten-year window ($\bar{\tau}=9$) to allow for slow moving technological transitions.[22] The knowledge IO linkages capture the country-sector (i,h) citations made by patents filed over a ten-year window to all prior sector l patents from all countries as a share of total citations made by country-sector (i,h) patents filed over the ten-year window.

Similarly, we measure production IO linkages as the importance of goods produced in each input sector for each output sector. Because the availability of highly disaggregated data on inter-sectoral sales is comparatively limited, we focus on within-country transactions in the US. We define $\rho_{i,t}^{l,h}$, the analog to $\kappa_{i,t}^{l,h}$ for the production IO table, as

$$\rho_{i,t}^{l,h} = \frac{\text{Sales}_{j,i,t}^{l,h}}{\sum_{k\in\mathcal{H}} \text{Sales}_{i,t}^{k,h}}, \tag{4}$$

where $\text{Sales}_{i,t}^{l,h}$ is the total value of sector l goods sold to country-sector-year (i,h,t). Production IO linkages measure, for year t, the sales from sector l to sector h as a share of the total sales from all sectors to sector h.

The linkages are based on US data from the BEA Use tables as described in Section 2. Since the BEA Use tables are only available at five-year intervals, we use the production IO linkages constructed from the data in each table for multiple years. For consistency with the measurement of production IO linkages, we also use only knowledge IO linkages from the same years for which there is a BEA Use table. In addition, to allow sectoral variation in exposure to technology inputs to be determined in advance of exposure in a given year, we use IO linkages that are lagged relative to the years in which exposure is measured. This lag in exposure variation is applied to both knowledge and production IO linkages.[23]

4.2. Description of knowledge and production IO tables

The construction of the knowledge and production IO tables relies on different data. However, there is little point in examining the effects of embodied technology in knowledge and production inputs separately if the two IO tables are closely related to one another. We now turn to illustrating several stylized observations regarding the two IO tables to demonstrate that they reflect distinct measures of the inter-sectoral relevance of technology inputs.

Both knowledge and production IO linkages take on values between zero and one. Values closer to one indicate stronger relationships whereas values closer to zero indicate weaker relationships. In Fig. 1, we depict the knowledge and production IO tables for the US economy in 2002, with values of $\kappa_{US,2002}^{l,h}$ represented in the left panel and $\rho_{US,2002}^{l,h}$ in the right panel. In each table, rows correspond to input sector l and columns correspond to output sector h. The color of each cell depends on the size of the IO linkage between the input and output sectors. We plot only those IO linkages for which the input sector accounts for at least 1% of the inputs used by the output sector. We also sort sectors in the IO tables based on their relative importance as a source of production inputs across output sectors to visually highlight the differences in the IO tables.

An immediate insight from Fig. 1 is that there are clear differences in the patterns of knowledge and production IO linkages for many sectors. We formalize and build on this visual intuition through five descriptive observations that highlight the differences between the knowledge and production IO tables.[24]

Observation 1. The sources of knowledge and production inputs are not highly correlated for the average sector in the US.

Observation 2. The sources of production inputs are more highly concentrated than the sources of knowledge inputs for the average sector in the US.

Observation 3. The key input-supplying sectors are distinct in the US knowledge and production IO tables.

Observation 4. For the average sector in the US, the sources of knowledge inputs are more highly correlated across time than the sources of production inputs.

[22] For example, Berkes et al. (2022) and Ayerst (2022) find that ICT sectors have become more important sources of innovations over the period of our analysis, highlighting the need for dynamic IO linkages. Baslandze (2018) finds an overall increase in the interconnectedness of sectors over this period.

[23] To be more precise, we use $\kappa_{i,1992}^{l,h}$ and $\rho_{i,1992}^{l,h}$ for exposure measured between 1995 and 2000, $\kappa_{i,1997}^{l,h}$ and $\rho_{i,1997}^{l,h}$ when we examine exposure between 2001 and 2005, $\kappa_{i,2002}^{l,h}$ and $\rho_{i,2002}^{l,h}$ for years between 2006 and 2010, and $\kappa_{i,2007}^{l,h}$ and $\rho_{i,2007}^{l,h}$ between 2011 and 2015.

[24] One can also clearly see that own-sector IO linkages along the diagonal are, in general, large relative to off-diagonal IO linkages in both the knowledge and production IO tables. We discuss the importance of own-sector versus cross-sector (off-diagonal) linkages both for the presentation of these observations in Online Appendix A and for our empirical results in Section 6.

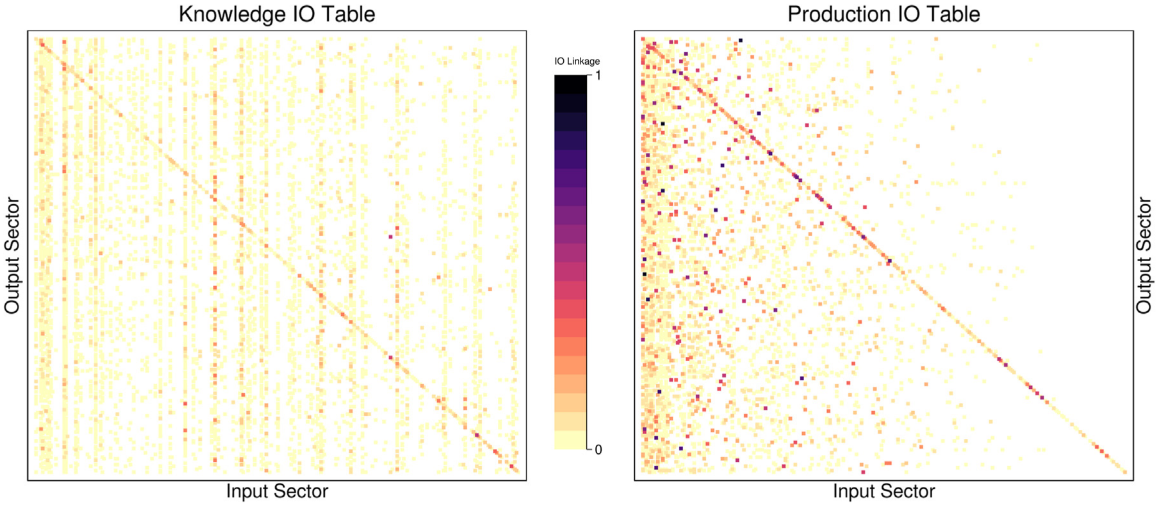

Fig. 1. Input-Output Tables. *Note*: This figure displays the knowledge and production IO tables where each point represents an IO linkage. The row position of each output sector and column position of each input sector are held constant across both IO tables to facilitate visual comparisons across the tables. Sectors are sorted based on their economy-wide importance as suppliers of production inputs by summing up the production IO linkages of each input sector over off-diagonal output sectors. The plots include the 292 2002 BEA sectors in agriculture, forestry, fishing and hunting, manufacturing, and mining with a non-zero sum of knowledge IO linkages across input sectors. Knowledge (production) IO linkages are defined in Eq. (3) (Eq. (4)). Knowledge IO linkages are based on backward citations of patents assigned to the US filed between 1993–2002 while production IO linkages are based on the 2002 BEA Use table. Both plots only display IO linkages that account for at least 1% of the inputs used by an output sector while all other IO linkages are visually suppressed.

Observation 5. The sources of knowledge inputs are less persistent on average in non-US country-sectors than in US sectors. In part, this reflects a convergence on average between each non-US country-sector and the same sector in the US.

We relegate the elaboration of these observations to Online Appendix A as a comparison of the IO tables is tangential to our main objectives. Nevertheless, a key implication of the observations is that the knowledge and production IO structures of the economy capture different relationships between sectors and, consequently, may capture different potential sources of technology spillovers. Furthermore, although US knowledge IO linkages tend to be highly persistent, there is a considerable degree of variation over time in US production IO linkages (as well as in non-US knowledge IO linkages). In our baseline analysis we explore the diffusion of knowledge through imports of embodied technology weighted using both dynamic knowledge IO linkages and production IO linkages.

The approach used to measure knowledge IO linkages in this paper shares many similarities with the method developed concurrently in Liu and Ma (2023). Like them, we construct knowledge IO tables for many countries across a long period of time using data from Google Patents and provide a descriptive analysis of the IO tables. Although both papers use different sources of data to construct production IO tables, both document a low correlation between the knowledge and production IO tables across all IO linkages in the tables. The analysis in this paper goes one step further to show that this low correlation prevails across most output sectors and at different points in time.[25] Although Liu and Ma (2023) describe their global knowledge IO table as being stable across time and across countries, our more disaggregated analysis reveals that there are many non-US country-sectors for which the sources of knowledge inputs change substantially across our sample period. Similarly, while Liu and Ma (2023) find that the knowledge IO tables in most top patenting countries are highly correlated with the US knowledge IO table when using patents from all years, our analysis suggests that there is considerable variation across country-sectors and across time in the correlation of US and non-US knowledge IO linkages.

5. Empirical specification

We now describe the main empirical specification of our analysis and the construction of key variables. Our main regressions involve regressing innovation outcomes on measures of embodied technology imports. We start by specifying our outcome variables. We then use the knowledge and production IO tables to develop the explanatory variables. Finally, we outline the empirical analog of Eq. (2) and an instrumental variable (IV) approach that we use to identify the effects of spillovers from embodied technology imports.

The US as the technological frontier. Throughout the analysis, we focus on the effects of imports from the US, which we consider to be the technology frontier. We make this assumption for two main reasons. First, the US is both the most

[25] Although the comparable regression-based analysis in both papers is undertaken at the level of output sectors and not countries, Liu and Ma (2023) do not document patterns of variation across output sectors and across time in the knowledge IO tables other than for domestic citation shares.

Table 1
Description of Outcome Variables.

Variable	Description
$Patents^{h}_{i,t}$	The count of patent applications using the allocation rules described in Section 2.
$FwdCites^{h}_{i,t} = \sum_{l\in\mathcal{H}}\sum_{j\in\mathcal{I}}\sum_{s=t}^{t+5} Cites^{h,l}_{i,j,t,s}$	Measures total citation-weighted patenting activity. The measure only includes citations received in the five years following the calendar year of the patent's application to mitigate truncation issues that would arise in later periods of the sample if citations received in all years were used.
$FwdRate^{h}_{i,t} = \frac{FwdCites^{h}_{i,t}}{Patents^{h}_{i,t}}$	Measures the average quality of patent applications as the citations received by patents divided by total patents.
$USBackCites^{h}_{i,t} = \sum_{l\neq h}\sum_{s=0}^{t} Cites^{l,h}_{US,i,s,t}$	The number of cross-sector citations to US patents. This measure excludes own-sector backward citations to be consistent with the focus on cross-sector imports of embodied technology (discussed below).
$USBackRate^{h}_{i,t} = \frac{USBackCites^{h}_{i,t}}{Patents^{h}_{i,t}}$	Measures the rate at which patents cite cross-sector US patents. In this regard, the measure captures the intensive margin of technology diffusion.
$USBackShare^{h}_{i,t} = \frac{USBackCites^{h}_{i,t}}{\sum_{l\neq h}\sum_{j\neq i}\sum_{s=0}^{t} Cites^{l,h}_{j,i,s,t}}$	The share of cross-sector citations to US patents in total cross-sector citations to foreign patents. The measure is informative of the extent to which knowledge inputs are substituted towards US knowledge in response to larger embodied technology flows from the US.

Note: Variables are defined for a country-sector-year (i,h,t). For *FwdRate*, *USBackRate*, and *USBackShare*, we exclude (i,h,t) observations with a value of zero in the denominator.

Table 2
Summary Statistics.

	A: Outcome Variables					B: Technology Variables			
	N	Median	Mean	SD		N	Median	Mean	SD
$Patents^{h}_{i,t}$	478880	0.083	0.786	1.384	$EmbTechK^{h}_{i,t}$	478880	16.090	15.788	3.119
$FwdCites^{h}_{i,t}$	478880	0.171	1.203	1.858	$EmbTechP^{h}_{i,t}$	478880	13.129	12.604	4.088
$FwdRate^{h}_{i,t}$	368950	1.394	1.405	0.745	$EmbTechDiag^{h}_{i,t}$	478880	14.174	13.425	5.397
$USBackCites^{h}_{i,t}$	478880	0.403	1.585	2.151	$OwnTech^{h}_{i,t}$	478880	2.471	3.160	3.003
$USBackRate^{h}_{i,t}$	368950	2.198	2.129	0.988	$USTech^{h}_{t}$	478880	10.988	10.827	1.200
$USBackShare^{h}_{i,t}$	364724	0.498	0.486	0.194	$USTechDiag^{h}_{t}$	478880	9.435	8.998	2.859

Note: Outcome variables are defined in Table 1 and technology variables are defined in Eqs. (5) to (10). All outcome variables are averaged over the three-year window t to $t+2$. All statistics for outcome variables are calculated on the log of one plus the variable except for the statistics for *USBackShare*. All statistics for the technology variables are calculated on the log of the variable.

innovative country and the largest originator of cross-country citations over this time period (see Berkes et al., 2022, for evidence), which reflects the characteristics of the frontier economy. Second, setting the frontier economy to the US allows us to be consistent with our data measurement, since the US has detailed data available for multiple years to construct both the production and knowledge IO tables described in Section 4 that are necessary for our analysis.

Sample Description. The unit of observation in our analysis is a country-sector-year. We limit our final panel of data to the years 1995 to 2015. Years prior to 1995 lack trade data for many countries and including later years would give rise to truncation issues for patents and forward citations, which are the main data used for our innovation outcome variables. We also limit the set of countries in our final sample based on their population, triadic patenting activity, and export-to-GDP and import-to-GDP ratios. We do this to avoid including countries where patenting outcomes may be too noisy, or countries that act as intermediaries in the global trade network. Online Appendix B provides the details of our sample selection procedure.

5.1. Outcome variables

Table 1 summarizes, defines, and describes the main outcomes variables. The conceptual framework illustrates the relationship between frontier knowledge spillovers and innovation outcomes. The focus is on the rate of innovation as the main outcome variable, which we measure in the data using both the rate of patenting (***Patents***) and the citation-weighted rate of patenting (***FwdCites***). We also use the average quality of patents, measured by the number of forward citations per patent (***FwdRate***), to examine both the intensive and extensive margin effects of frontier knowledge spillovers.

We also examine evidence on the extent to which trade of embodied technology is a source of technology diffusion and leads to higher flows of knowledge from the US. Specifically, we use the backward citation information underlying the knowledge IO table as a measure of cross-country knowledge flows. We construct three outcome variables measuring cross-sector backward citations to US patents (***USBackCites***), the per-patent rate of backward citations to US patents (***USBackRate***), and the share of backward citations to US patents in total cross-sector backward citations to foreign patents (***USBackShare***).

In our baseline specification, we measure outcome variables using the average of the variables in the three-year window between year t and $t+2$. Panel A of Table 2 presents summary statistics of the main outcome variables. *FwdRate* and *USBackRate* are calculated only for observations with non-zero patent counts, while *USBackShare* is calculated only for observations with non-zero cross-sector citations to foreign patents. The summary statistics show that the distribution of

Patents and *FwdCites* are highly skewed with the median country-sector-year having values close to zero. The distribution of outcomes based on backward citations are similarly skewed.

5.2. Embodied technology imports and other controls

Our main variable of interest is the frontier knowledge spillovers through embodied technology imports ($EmbTech^h_{i,t}$ in the conceptual framework). Before turning to our main variables of interest, we discuss the construction of knowledge stocks $K^p_{i,t}$, which are used in the construction of the main variables.

We measure the technological content of a subsector's goods using patents data as

$$K^p_{i,t} = (1-\delta)K^p_{i,t-1} + FwdCites^p_{i,t},$$

where $FwdCites^p_{i,t}$ is the five-year forward citations in country-subsector-year (i, p, t) and δ is the depreciation rate of knowledge that we set to 5% to be consistent with commonly used values. For each country and subsector, we initialize the stock of knowledge in 1940 with the value $K^p_{i,1940} = FwdCites^p_{i,1940}/\delta$. The initial value has little influence on the knowledge stocks used in the period of our analysis because of the long time period.

Embodied technology imports. We set the linkages γ in the conceptual framework to be a combination of knowledge κ and production ρ IO linkages, $\gamma^{l,h}_{F,t} = \alpha\kappa^{l,h}_{F,t} + (1-\alpha)\rho^{l,h}_{F,t}$. We do not impose structure on whether knowledge flows between sectors are better captured by the knowledge or production IO linkages. For each country-sector-year (i, h, t), our measure of knowledge-weighted embodied technology imports is given by

$$EmbTechK^h_{i,t} = \prod\nolimits_{l\neq h}\left(1 + \sum\nolimits_{p\in\mathcal{P}^l}\left(w^l(p)K^p_{US,t}M^p_{US,i,t}\right)\right)^{\kappa^{l,h}_{US,t}} \tag{5}$$

and production-weighted embodied technology imports are given by

$$EmbTechP^h_{i,t} = \prod\nolimits_{l\neq h}\left(1 + \sum\nolimits_{p\in\mathcal{P}^l}\left(w^l(p)K^p_{US,t}M^p_{US,i,t}\right)\right)^{\rho^{l,h}_{US,t}}, \tag{6}$$

where $w^l(p)$ is the concordance weight discussed in Section 2.

The amount of imported embodied technology depends on the flow of knowledge into country i from every sector $l \neq h$ in the United States. This flow is increasing in the volume of imports and the stock of knowledge in sector l. Countries that spend more on sector l goods from the United States have a higher flow of knowledge into them from that sector. For example, a larger volume of imports could reflect more varieties of a sector's goods being imported. Our measure reflects the idea that as a country imports more, ideas upon which domestic innovators can build become more abundant and readily available. The effect of a flow of knowledge from a given sector l is weighted by the tendency of that sector's knowledge to be used in sector h in the US.

An important distinction from the conceptual framework is that we construct the measures of embodied technology from the levels of US imports ($M^p_{US,i,t}$), rather than trade scaled by output ($m^p_{US,i,t} = M^p_{US,i,t}/Y^p_{i,t}$), since output data is unavailable at the level of aggregation we examine. A potential issue with this construction is that higher imports could simply reflect that the importing country has a larger population or economy.[26] We include granular fixed effects as a best attempt to deal with this issue.[27]

We omit the own-sector component in the embodied technology spillover terms as within-sector imports and innovation outcomes can potentially be related to each other for multiple reasons. Within-sector demand shocks can lead to countries importing more foreign products to satiate demand while at the same time investing more in innovation activities in the sector due to increased returns. Moreover, own-sector imports can also affect innovation outcomes in a country through import competition effects, since firms may invest more in innovation in order to escape foreign competition. Finally, R&D productivity shocks and profitability shocks to a country-sector can lead to imports and innovations by the country-sector to be correlated. Given these concerns, we include a control for the own-sector component of embodied technology spillovers constructed as

$$EmbTechDiag^h_{i,t} = 1 + \sum\nolimits_{p\in\mathcal{P}^h} w^h(p)K^p_{US,t}M^p_{US,i,t}. \tag{7}$$

We do not scale the knowledge inputs by the IO weights, $\kappa^{l,l}_{US,t}$ or $\rho^{l,l}_{US,t}$, since we expect that this variable also captures factors not directly related to the effects of technology diffusion, such as import competition.[28]

[26] It is also worth noting that this is not an issue with the US technology stocks $K^p_{US,t}$ since they are common to all importers and capture the abundance of technology embodied within imports, meaning their levels are important for the interpretation of our results.

[27] An alternative would be to use US import shares, i.e., US imports to country i over all imports to country i. This construction leads to misleading conclusions because trends in trade—e.g., trade increases for most country sectors over this period—lead to declining US import shares for most countries.

[28] The evidence on the effect of import competition on innovation activity is mixed. For example, Bloom et al. (2016) find that increased trade with China between 2000 and 2005 led to an increase in patenting activity in European firms that were more exposed to that competition (which was also the case for increased exposure to trade from other low-wage countries). In contrast, Autor et al. (2020) find that import competition due to increased trade with China decreased patenting activity in publicly listed US firms and technology classes more exposed to that competition. In contrast, they find that changes in import penetration of high-wage countries like the US had no effect on patenting. Nevertheless, we include this variable to mitigate concerns that the estimated effects of imported embodied knowledge pick up these import competition effects.

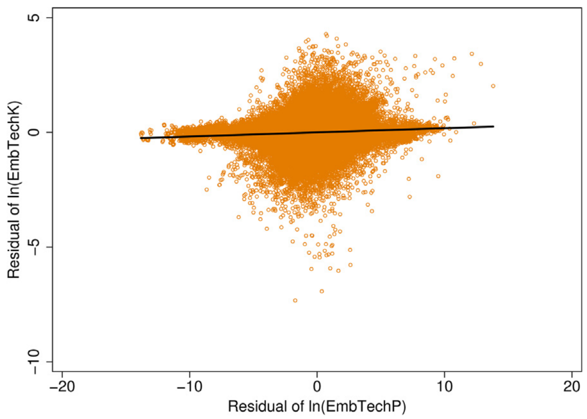

Fig. 2. Residual Variation in Measures of Embodied Technology Imports. *Note*: This figure plots residuals of ln($EmbTechK$) and ln($EmbTechP$), defined in Eq. (5) and (6), and the line of best fit from the regression of the former measure on the latter. Residuals are computed by regressing each measure on the set of fixed effects included in the baseline specifications discussed in Section 5.3.

Estimating the separate effects of knowledge- and production-weighted imports of embodied technology requires that there is sufficient variation across observations in our sample in these two measures. To assess this, we regress the logs of both *EmbTechK* and *EmbTechP* on the set of fixed effects included in our baseline specifications. Fig. 2 plots the fitted residuals from these regressions and a line of best fit from a regression of one set of residuals on the other. The figure highlights that there is considerable orthogonal variation in the two residualized variables.

Other variables. We also include measures of the country-sector domestic and US stocks of knowledge to control for alternative channels of diffusion. We construct the measure of the domestic knowledge stock as

$$OwnTech_{i,t}^{h} = \prod_{l} \left(1 + \sum_{p\in\mathcal{P}^l} w^l(p) K_{i,t}^{p}\right)^{\kappa_{i,t}^{l,h}}. \tag{8}$$

Higher values of *OwnTech* reflect higher stocks of country-sector-year (i, l, t) knowledge that tend to be cited by sector h. We use the domestic knowledge IO linkages $\kappa_{i,t}^{l,h}$ in the construction because this is the best measure of the relevance of sector l knowledge for producing innovations in country-sector-year (i, h, t).[29]

Analogous to *EmbTechK* and *EmbTechDiag*, we also control for US knowledge stocks relevant to a sector constructed as:

$$USTech_{t}^{h} = \prod_{l\neq h} \left(1 + \sum_{p\in\mathcal{P}^l} w^l(p) K_{US,t}^{p}\right)^{\kappa_{US,t}^{l,h}} \tag{9}$$

$$USTechDiag_{t}^{h} = 1 + \sum_{p\in\mathcal{P}^h} w^h(p) K_{US,t}^{p}\ . \tag{10}$$

These controls serve two purposes. First, they control for the effect of the growth in frontier knowledge on each sector's innovation activity that is common to all countries irrespective of each country's trade integration with the frontier economy. Since our choice of instrumental variables (described in Section 5.4) precludes us from including strict sector-year-level fixed effects, these controls help in capturing part of sector-year trends relevant for knowledge production that are common to all countries. Second, we expect that knowledge diffuses across countries through mechanisms other than trade, and controlling for *USTech* captures some of these aspects of knowledge diffusion. Diffusion can occur through direct technology licensing, FDI, immigration, and other channels (Keller, 2004; 2010; 2021). Including *USTech* allows us to separately identify the effect of traded embodied technology from broader technology diffusion across countries. Panel B of Table 2 provides summary statistics for all the technology-related variables discussed in this section.

5.3. Estimation equation

We now present the empirical counterpart of Eq. (2) in terms of our constructed variables that serves as our baseline specification:

$$\begin{aligned}\ln(1 + Outcome_{i,t}^{h}) = {} & \theta_1 \ln EmbTechK_{i,t-1}^{h} + \theta_2 \ln EmbTechP_{i,t-1}^{h} + \theta_3 \ln OwnTech_{i,t-1}^{h} \\ & + \theta_4 \ln EmbTechDiag_{i,t-1}^{h} + \theta_5 \ln USTech_{t-1}^{h} + \theta_6 \ln USTechDiag_{t-1}^{h}\end{aligned}$$

[29] We do not have highly detailed and dynamic measures of production IO linkages for most countries, which prevents us from constructing a similar measure with $\rho_{i,t}^{l,h}$.

$$+V_{i,t-1}^{h}\beta + f_{i,t} + f_{t}^{n(h)} + f_{i}^{h} + \epsilon_{i,t}^{h}, \tag{11}$$

where $V_{i,t}^{h}$ is a vector of controls that includes own-sector imports from the world and exports to the world and $f_{i,t}$, $f_{t}^{n(h)}$ and f_{i}^{h} are country-year, (summary) sector-year, and country-sector fixed effects. In the baseline regressions, we lag all our regressors as well as average outcomes over a three-year window from t to $t+2$ to reduce noise and to allow for a more gradual diffusion of knowledge. With the exception of *USBackShare*, we also transform the outcome variable as $\ln(1+Outcome_{i,t}^{h})$ to keep observations that have zero-valued innovation/diffusion outcomes.

In all regressions, we allow for the possibility that the residuals are correlated across years within a country-sector pair (due to serial correlation) and across countries in each year within a sector (since much of the variation in our variables of interest is at the sector-year level). To do so, we estimate multi-way clustered standard errors at the country-sector and sector-year levels (Cameron et al., 2011).

The coefficients of interest are related to model parameters as $\theta_1 = \eta_S\alpha, \theta_2 = \eta_S(1-\alpha)$, and $\theta_3 = \eta_Z$, where α is the contribution of knowledge linkages towards cross-sectoral spillovers of embodied technology imports. The estimates of θ_1 and θ_2 should be positive since our hypothesis is that spillovers from embodied technology imports should improve innovation outcomes. We also expect that the estimates for θ_1 will in general be larger than for θ_2 since the knowledge weights reflect a more direct measure of the relevance of embodied technology imports for patenting. Similarly, the estimate of θ_3 should be positive for the rate of innovating (*Patents* and *FwdCites*) but may be ambiguous for the quality of patents if higher knowledge stocks correspond to higher rates of low-quality innovations. For diffusion outcomes, the estimate of θ_3 should be positive for overall US citations and near zero for the other outcomes.

5.4. Endogeneity concerns

The fixed effects in Eq. (11) control for time-invariant characteristics of country-sector pairs, factors that vary at the country level over time, and sector-year shocks that are common to sectors within a summary sector. Despite the inclusion of these fixed effects, there remain potential endogeneity concerns with our regressors of interest.

One possibility is that variation across country-sector-years in the amount of relevant technology inputs embodied in a country's imports in prior years could reflect demand shocks for those inputs that also directly affect patenting outcomes. For example, shocks to expected profits, captured by $u_{i,t}^{h}$ in the conceptual framework,[30] would both increase R&D investment but also the imports of intermediate inputs used in the production of goods in (i, h, t).[31] Similarly, shocks to domestic R&D productivity, captured by $\nu_{i,t}^{h}$ in the conceptual framework, can give rise to endogeneity concerns. If domestic R&D and embodied technology in cross-sector imports are substitutes in the production of new innovations, then shocks to the domestic R&D productivity will reduce demand for embodied technology imports and ordinary least squares estimates of the effects of embodied technology imports would suffer from a negative bias. On the other hand, if imported embodied technology and R&D are complements in the production of new innovations, then there would be a positive bias affecting the OLS estimates. If these shocks were serially correlated, there would be a spurious correlation between innovation output and imports of embodied technology in past years arising from the shocks. Since there is no data available on R&D expenditures at the level of sectoral disaggregation used in our analysis, we cannot control for these innovation inputs which may cause an omitted variable bias to affect our estimates.

To address these concerns, we use an instrumental variable strategy that focuses on variation in US imports that is a function of supply shocks to US exports. Specifically, we instrument each regressor that includes US imports with an instrument that constructs the variable using US exports to all countries outside of a country-specific cluster (discussed below). For our main regressors, we construct the instrumental variables as

$$IVEmbTechK_{i,t}^{h} = \prod\nolimits_{l\neq h}\left(1+\sum\nolimits_{p\in\mathcal{P}^{l}}(w^{l}(p)K_{US,t}^{p}X_{US,i,t}^{p})\right)^{\kappa_{US,t}^{l,h}}, \tag{12}$$

and

$$IVEmbTechP_{i,t}^{h} = \prod\nolimits_{l\neq h}\left(1+\sum\nolimits_{p\in\mathcal{P}^{l}}(w^{l}(p)K_{US,t}^{p}X_{US,i,t}^{p})\right)^{\rho_{US,t}^{l,h}}, \tag{13}$$

where $X_{US,i,t}^{p} = \sum_{j\notin\mathcal{C}_i} M_{US,j,t}^{p}$ and $\mathcal{C}_i$ is a cluster of countries with similar characteristics to country i. We construct the cluster $\mathcal{C}_i$ as the countries that fall into both the same quintiles of the distribution of GDP-per-capita (in 1995) and the total trade (imports plus exports) to GDP ratio as country i.[32] We include the first variable to capture similarities in technological development across countries and the second variable to capture similarities in trade patterns across countries.

Our IV strategy isolates trade flows to the domestic country that stem from supply shocks in the US. A standard leave-one-out instrument would exclude the domestic economy to discount changes in trade that result from domestic demand

[30] For example, demand shocks or policy driven shocks in a sector can change expected future profits in that sector.

[31] We do not explicitly model demand for production inputs from different sectors and instead implicitly subsume them into the expected profit function.

[32] We use the entire sample of countries that we have data for, not just those included in the final sample.

shocks. We extend this intuition by not only excluding the domestic economy but also countries that share similar characteristics and, consequently, may face demand shocks that are correlated with those facing the domestic country.

Although the instrumental variables correct for demand shocks to imports that vary at the country-sector-year level, we would be remiss not to mention the principal endogeneity concern that may still be present. If there are demand shocks to imports at the sector-year level for sectors within a BEA summary sector that are common across all countries, our instruments may reflect these demand shocks rather than US export supply shocks. Since the instruments are constructed at the sector-year level, we are unable to include sector-year fixed effects at the same level of disaggregation and so cannot preclude that these common demand shocks affect our estimates.[33]

6. Results

We now discuss the results from estimating the effects of knowledge-weighted and production-weighted embodied technology imports on the innovation and diffusion outcomes. We report the estimates using the baseline specifications described in Eq. (11) and use the empirical model to quantify the magnitudes of the results. We close this section by discussing the robustness of the results to alternative specifications.

6.1. Baseline results

Table 3 reports the results for the innovation and diffusion outcomes. We interpret the innovation results as corresponding to $X_{i,t}^{h}$ in the context of the conceptual framework. All regressions include country-sector, summary sector-year, and country-year fixed effects as described in the model and standard errors are clustered at the country-sector and sector-year levels. The IV estimates are larger than the OLS estimates (reported in Online Appendix D) for knowledge-weighted embodied technology imports *EmbTechK* and are similar for production-weighted embodied technology imports *EmbTechP*.

The innovation outcomes, reported in columns (1) to (3), point to a positive impact from knowledge-weighted and production-weighted embodied technology imports *EmbTechK* and *EmbTechP* on the innovation rate, as measured by *Patents* and *FwdCites*. The effect appears to be primarily through the extensive margin (i.e., more patenting) rather than the intensive margin (i.e., more citations). The elasticity of innovation outcomes with respect to shocks to *EmbTechP* is substantially lower than for *EmbTechK*, by multiple orders of magnitude, for both *Patents* and *FwdCites*, suggesting that technology spillovers through trade operate primarily through knowledge linkages. We show later that the quantitative comparison holds after accounting for the relative variation in *EmbTechK* and *EmbTechP*.

The diffusion outcomes, reported in columns (4) to (6), point to positive, albeit more mixed, impacts of knowledge-weighted and production-weighted technology imports *EmbTechK* and *EmbTechP*. The coefficient estimates for *USBackCites* are positive for both *EmbTechK* and *EmbTechP*. To some extent, this follows from the previous result since more frequent patenting implies more backward citations to the US, all else equal.

For *USBackRate* and *USBackShare* the point estimates are positive, in line with our expectations, but the estimates are statistically insignificant. The results suggest that while knowledge-weighted and production-weighted embodied technology imports improve innovation outcomes, knowledge may diffuse through channels that do not result in citations. That said, the estimates of *EmbTechK* for *USBackRate* are more mixed than when looking at alternative specifications (see Online Appendix D). This is consistent with the point estimate of *USBackCites* relative to that of *Patents*, with the former being almost twice as large as the latter in Table 3. We expect that for both *USBackRate* and *USBackShare* the estimated coefficient is biased downwards since our measure of backward citations is noisier than for *Patents* and *FwdCites*. Specifically, our measure includes citations to any previous US patent, many of which may include innovations that are already well-understood in the domestic economy. Additionally, many countries may not require or enforce that domestic patents cite foreign patents making backward citations a noisy measure of diffusion.[34] Another possibility is that innovators learn both about US patents but also other foreign innovations through US patents. For example, an innovator may learn about a French innovation that is built on by a US product and cite that patent instead. In this regard, we think of both *USBackRate* and *USBackShare* as being relatively strict measures of diffusion.

The coefficient for *EmbTechDiag* is near zero reflecting that this variable captures several channels with potentially opposing impacts on innovation, such as, for example, competition from imports and knowledge spillovers. We do not focus on the spillovers through this channel due to the aforementioned difficulties with their interpretation. The coefficients for *OwnTech* are positive for *Patents* and *FwdCites* suggesting that the stock of domestic knowledge relevant to a country-sector contributes positively to its innovation rate.

The coefficient on *USTech*, which measures relevant upstream knowledge produced by the US, is negative in our baseline results. While this measure captures broader technology diffusion from the US to other countries, it is also correlated with our measure of *EmbTechK*. A change in *USTech* will also result in a change in *EmbTechK* since both measures have the same

[33] While this may be the case, because our sample includes many countries at a wide range of stages of development and with different patterns of trade with the US, it is unlikely that there are sector-year demand shocks common to all the countries in our sample.

[34] Our choice to allocate patents based on the location of the innovator would add noise to both measures. For example, a citation to a patent with 50% US innovators and 50% French innovators would not increase *USBackShare* even if local innovators learned about the patent through interactions with the US innovators US employer.

Table 3
Main Results.

	(1) *Patents*	(2) *FwdCites*	(3) *FwdRate*	(4) *USBackCites*	(5) *USBackRate*	(6) *USBackShare*
ln *EmbTechK*	0.666***	0.545***	0.052	1.023***	0.076	0.002
	(0.048)	(0.054)	(0.048)	(0.082)	(0.055)	(0.011)
ln *EmbTechP*	0.003**	0.002*	-0.001	0.005*	0.000	0.000
	(0.001)	(0.001)	(0.001)	(0.002)	(0.002)	(0.000)
ln *EmbTechDiag*	0.019	0.001	-0.013	0.023	0.012	0.004
	(0.012)	(0.012)	(0.012)	(0.021)	(0.015)	(0.003)
ln *OwnTech*	0.005**	0.006**	-0.047***	0.019***	-0.051***	-0.005***
	(0.002)	(0.002)	(0.004)	(0.004)	(0.005)	(0.001)
ln *USTech*	-0.761***	-0.596***	-0.086	-1.130***	-0.128	-0.007
	(0.064)	(0.072)	(0.070)	(0.110)	(0.079)	(0.017)
ln *USTechDiag*	0.097***	0.165***	-0.003	0.261***	0.109***	0.009
	(0.015)	(0.016)	(0.022)	(0.026)	(0.026)	(0.006)
Observations	478,880	478,880	368,950	478,880	368,950	364,724
F-Stat *EmbTechK*	412.6	412.6	609.6	412.6	609.6	686.7
F-Stat *EmbTechP*	12806.4	12806.4	17535.9	12806.4	17535.9	18557.1
F-Stat *EmbTechDiag*	322.1	322.1	349.5	322.1	349.5	350.0

Note: Each column reports IV estimates of the regression specification in Eq. (11). All dependent variables, which are defined in Table 1, are first averaged over the three-year window t to $t+2$ and then transformed using $\ln(1 + Outcome)$, where *Outcome* is the variable in the column title. Explanatory variables are defined in Eqs. (5) to (10). The instrumental variable approach is described in Section 5.4. Other controls include one-year lags of the log of total exports from the country-sector to other countries and the log of total own-sector imports from other countries. All regressions include country-sector, (summary) sector-year, and country-year fixed effects. Standard errors clustered at both the country-sector and sector-year levels are in parentheses.

Table 4
Quantitative Significance.

Outcome	RSDO	Coefficient Estimate		Relative Implied RSD (%)	
		EmbTechK	*EmbTechP*	*EmbTechK* (RSDE = 0.037)	*EmbTechP* (RSDE = 1.197)
Patents	0.185	0.666	0.003	13.2	1.8
FwdCites	0.308	0.545	0.002	6.5	0.9
FwdRate	0.478	0.059	-0.001	0.5	-0.3
USBackCites	0.427	1.023	0.005	8.8	1.3
USBackRate	0.616	0.084	0.000	0.5	0.1
USBackShare	0.139	0.002	0.000	0.0	0.2

Note: RSDO is calculated as the standard deviation of the outcome variable after controlling for the fixed effects used in the baseline specification, Eq. (11). RSDE is calculated as the standard deviation of the explanatory variable after controlling for the other regressors and fixed effects used in that specification. For each of *EmbTechK* and *EmbTechP*, defined in Eqs. (5) and (6), relative implied RSD refers to the product of the coefficient estimate and the ratio of the RSDE to the RSDO. Outcome variables are defined in Table 1. Coefficient estimates are taken from Table 3.

underlying IO linkages and US technology levels by construction. Therefore, the negative estimate cannot be interpreted as the effect of *USTech* holding all else fixed. In order to understand the effect of *USTech*, we first residualize *EmbTechK* by regressing it on *USTech* and taking the residual. We then run our baseline regressions by replacing *EmbTechK* with its residualized version and keeping all other regressors the same. By the Frisch-Waugh-Lowell theorem, the estimates on all regressors except *USTech* will remain the same. The new estimate on *USTech* is positive, and can now be interpreted as the effect of *USTech* that is orthogonal to the trade-weighted US technology measure *EmbTech*. The details of this exercise are presented in Online Appendix C.1.

6.2. Quantitative significance

Our conceptual framework is used to derive Eq. (2), where the innovation rate $X^h_{i,t+1}$ depends on embodied technology imports *EmbTech* through an elasticity η_S and the IO linkages $\gamma^{l,h}_{F,t}$, which equals the knowledge $\kappa^{l,h}_{US,t}$ and production $\rho^{l,h}_{US,t}$ weights for the US scaled by α and $1-\alpha$, respectively. Taking the estimates for *Patents* and *FwdCites* implies that $\eta_S \alpha \in [0.54, 0.67]$ and $\eta_S(1-\alpha) = [0.002, 0.003]$.

To further understand the quantitative magnitude of the results, Table 4 compares the relative magnitude of the variation in the outcomes and the implied variation of outcomes attributable to the model variables. We focus on the residualized standard deviation of the explanatory variables, RSDE, which is calculated as the standard deviation of the variable after removing the estimated effects of all other regressors and fixed effects. The residual variation in the outcome variables, RSDO, is the standard deviation of the outcome variable in question after removing the estimated effects of the fixed effects. We do this to remove both variable trends as well as cross-country and cross-summary-sector variation in the variables. These differences are important for both the outcomes (e.g., increases in patenting over time) and embodied technology imports (e.g., increases in trade over time). However, these trends are not important for understanding the economic significance of

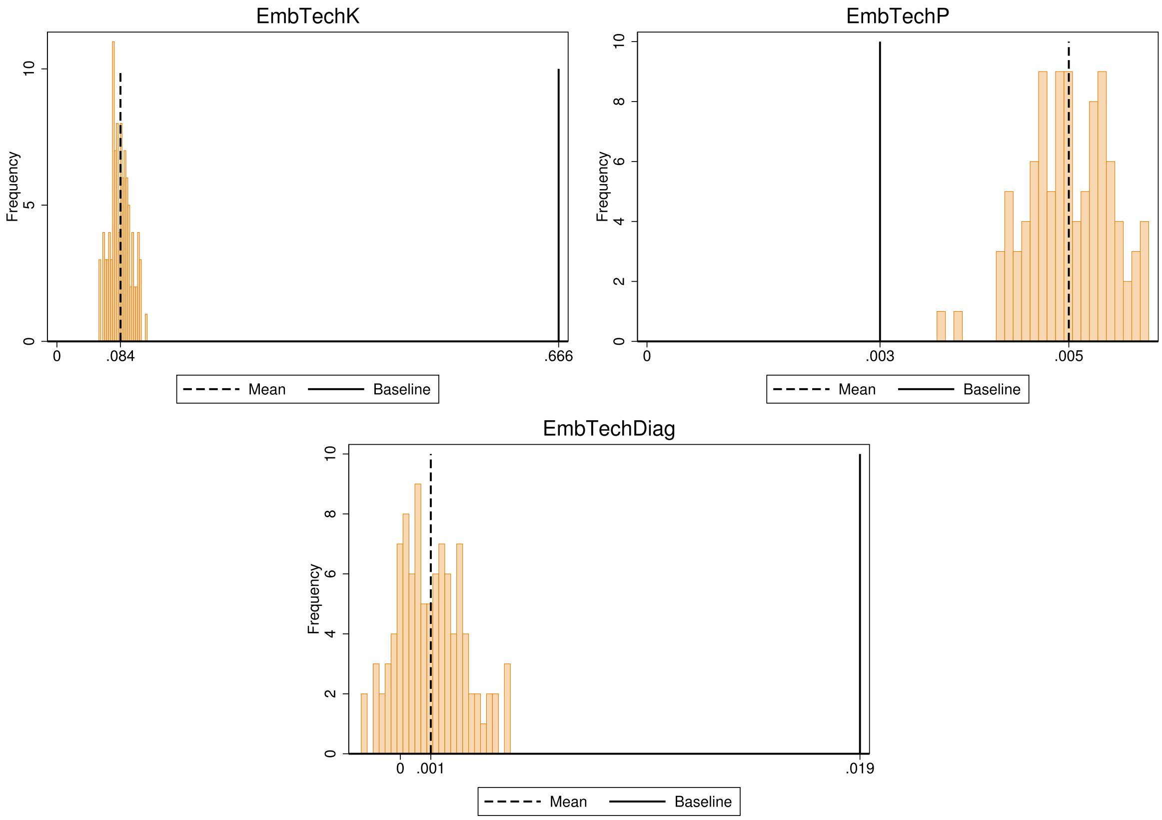

Fig. 3. Falsification Test Using Randomized Trade Flows. *Note*: This figure displays histograms of the coefficients θ_1^r, θ_2^r, and θ_4^r from the regression specified in Eq. (14) for 100 draws of randomized trade flows. The dependent variable in each regression is *Patents*. For each histogram, the dashed line indicates the mean of the distribution while the solid line represents the corresponding coefficient of interest from the baseline regression specified in Eq. (11) that is reported in Table 3.

the coefficient estimates.[35] We estimate the effect of the RSDE of the explanatory variables implied by the model and scale them by the RSDO of the outcome variables.

The table shows that *EmbTechK* explains around 13% of the residualized variation in *Patents*, 6.5% of the variation in *FwdCites*, and around 8.8% for *USBackCites*. Consistent with the earlier summary statistics, the table also shows that there is more residualized variation in *EmbTechP*, which increases its relative quantitative importance, but this gap is not large enough to offset the differences in coefficient estimates found in Table 3. That is, the overall impact of production-weighted embodied technology imports remains marginal compared with knowledge-weighted embodied technology imports.

6.3. *The role of trade flows in technology diffusion*

As discussed above, we do not expect technology to diffuse across borders only through trade linkages. We now consider a falsification exercise to help understand the role trade plays in this diffusion process. A first step towards identifying the role of trade, *separate* from broader diffusion channels, is by controlling for sector-relevant upstream US technology, *USTech*, as in our baseline specification. Our baseline results suggest that trade plays an important role in the form of diffusing embodied technology even after controlling for the same upstream technology that is not augmented by trade flows.

To further address this concern, we conduct a falsification test by randomizing the trade flows data used in the construction of our embodied technology variables. Specifically, we take the true imports from the US data for all countries that the data is available for, and randomize their allocation across years, countries, and sub sectors.[36] We construct the randomized counterparts of the embodied technology variables, which we label as *EmbTechK*(r), *EmbTeckP*(r), and *EmbTechDiag*(r) where r stands for the r^{th} randomization of trade flows. We then re-estimate our baseline specification using the randomized variables for 100 draws of randomized variables:

[35] For example, the inability of the empirical model to explain a secular trend in patenting over time is not informative to understanding the importance of embodied technology imports.

[36] See Online Appendix C.2 for details of the randomization procedure.

$$\ln(1+Outcome^{h}_{i,t}) = \theta^{r}_{1}\ln EmbTechK^{h}_{i,t-1}(r) + \theta^{r}_{2}\ln EmbTechP^{h}_{i,t-1}(r) + \theta^{r}_{3}\ln OwnTech^{h}_{i,t-1} + \theta^{r}_{4}\ln EmbTechDiag^{h}_{i,t-1}(r) + \theta^{r}_{5}\ln USTech^{h}_{t-1} + \theta^{r}_{6}\ln USTechDiag^{h}_{t-1} + V^{h}_{i,t-1}\beta^{r} + f^{r}_{i,t} + f^{rn(h)}_{t} + f^{rh}_{i} + \epsilon^{h}_{i,t}. \quad (14)$$

Fig. 3 shows the distribution of estimates θ^r_1, θ^r_2, and θ^r_4 from the regression specification in Eq. (14) for *Patents* as the outcome variable. The mean of the randomized model coefficient estimates is plotted as a dashed line and the baseline estimates are plotted as a solid line. The falsification test confirms that the trade weights used in the *EmbTechK* measure are important, with the baseline estimate (0.666) being eight times higher than the mean of the coefficients from the randomized data (0.084). However the baseline estimate on *EmbTechP* and its counterpart from the randomized data are relatively smaller and close to each other (0.003 and 0.005). This could indicate that diffusion of embodied technology through production linkages might not be a robust feature of the data. Finally, the coefficients on the own-sector embodied technology, *EmbTechDiag*, from the randomized data (mean of 0.001) are also an order of magnitude smaller compared to the larger, but statistically insignificant, baseline estimate (0.019). We conclude that the randomized data is unable to replicate the coefficient estimates for knowledge-weighted embodied technology imports, highlighting the importance of trade.[37]

6.4. Robustness checks and additional results

In Online Appendix D we demonstrate the robustness of the main results to alternative specifications and controls as well as present additional related results. The results are quantitatively similar when alternative lag structures are used instead of the baseline specification. Our baseline results construct *EmbTechK* and *EmbTechP* using upstream linkages. We show that the results hold when the corresponding variables constructed with downstream linkages are included. The results become stronger when the country sample is restricted to the top 40 countries based on patenting, consistent with our view that patenting activity measures are noisier in countries that patent less. The results hold using alternative IV strategies, including the traditional leave-one-out IV. The results hold using an alternative transformation of the outcome variables that avoids the log of one plus the outcome variable transformation. Finally, the results hold for alternative outcomes, based on triadic patents, and constructions of knowledge stocks.

7. Conclusion

Innovation activities are highly concentrated in a small number of countries, but new technology eventually diffuses to other countries. A potentially important channel through which technology diffuses across borders is international trade of goods, since importers can learn about the technology embodied in those goods. This paper assesses the extent to which knowledge and production inputs in traded goods contribute to the diffusion of technology and to the amount and quality of innovations developed in importing country-sector pairs.

To do this, we construct knowledge and production IO tables using data on inter-sectoral patent citations and sales. We combine these measures with data on sector-level trade flows between countries to construct measures of the knowledge-weighted and production-weighted technology imports. We show that increases in both measures of embodied technology lead to higher rates of innovation in an importing country-sector pair.

Our results point to important directions for future research, including towards better understanding the mechanisms underlying the trade channel of technology diffusion. For example, since knowledge linkages are a more important source of diffusion than production linkages and the sources of knowledge linkages are distinct from production linkages, diffusion through the trade of goods may not primarily occur between firm relationships that underpin our sector-level import data. Instead, spillovers may primarily be to other firms in importing countries. Future work using firm-level data can investigate the presence of these spillovers through knowledge IO linkages. The estimated elasticities in this paper could also be used to discipline a quantitative model to evaluate the aggregate growth and welfare implications of knowledge diffusion through trade and the effects of trade policy on innovation.

Supplementary material

Supplementary material associated with this article can be found, in the online version, at doi:10.1016/j.jmoneco.2023.05.002.

References

Acemoglu, D., Akcigit, U., Kerr, W.R., 2016. Innovation network. Proc. Natl. Acad. Sci. 113 (41), 11483–11488. doi:10.1073/pnas.1613559113.
Acharya, R.C., Keller, W., 2009. Technology transfer through imports. Canad. J. Econ. 42 (4), 1411–1448. doi:10.1111/j.1540-5982.2009.01550.x.
Aghion, P., Bergeaud, A., Gigout, T., Lequien, M., Melitz, M., 2021. Exporting ideas: knowledge flows from expanding trade in goods. Working Paper. https://conference.nber.org/conf_papers/f154069.pdf.

[37] We present the results of the falsification exercises for other outcome variables in Online Appendix C.2.

Aghion, P., Bergeaud, A., Lequien, M., Melitz, M.J., 2022. The heterogeneous impact of market size on innovation: evidence from French firm-level exports. Rev. Econ. Stat. doi:10.1162/rest_a_01199. (forthcoming)

Alvarez, F.E., Buera, F.J., Lucas Jr., R.E., 2013. Idea flows, economic growth, and trade. NBER Working Paper 19667. National Bureau of Economic Research doi:10.3386/w19667.

Amiti, M., Konings, J., 2007. Trade liberalization, intermediate inputs, and productivity: evidence from Indonesia. Am. Econ. Rev. 97 (5), 1611–1638. doi:10.1257/aer.97.5.1611.

Autor, D., Dorn, D., Hanson, G.H., Pisano, G., Shu, P., 2020. Foreign competition and domestic innovation: evidence from US patents. Am. Econ. Rev.: Insight. 2 (3), 357–374. doi:10.1257/aeri.20180481.

Ayerst, S., 2022. The diffusion of new general purpose technologies. Working Paper. http://www.stephenayerst.com/s/Adoption_v82-25a2.pdf.

Baslandze, S., 2018. The role of the IT revolution in knowledge diffusion, innovation and reallocation. Working Paper. https://www.dropbox.com/s/k5qcvclwsp7q7vj/Baslandze_IT_Innovation.pdf.

Berkes, E., Manysheva, K., Mestieri, M., 2022. Global innovation spillovers and productivity: evidence from 100 years of world patent data. Working Paper 2022-15. Federal Reserve Bank of Chicago doi:10.21033/wp-2022-15.

Bloom, N., Draca, M., Van Reenen, J., 2016. Trade induced technical change? The impact of Chinese imports on innovation, IT and productivity. Rev. Econ. Stud. 83 (1), 87–117. doi:10.1093/restud/rdv039.

Bloom, N., Schankerman, M., Van Reenen, J., 2013. Identifying technology spillovers and product market rivalry. Econometrica 81 (4), 1347–1393. doi:10.3982/ECTA9466.

Bombardini, M., Li, B., Wang, R., 2017. Import competition and innovation: evidence from China. Working Paper. https://www7.econ.hit-u.ac.jp/cces/trade_conference_2017/paper/matilde_bombardini.pdf.

Buera, F.J., Oberfield, E., 2020. The global diffusion of ideas. Econometrica 88 (1), 83–114. doi:10.3982/ECTA14044.

Cai, J., Li, N., 2019. Growth through inter-sectoral knowledge linkages. Rev. Econ. Stud. 86 (5), 1827–1866. doi:10.1093/restud/rdy062.

Cai, J., Li, N., Santacreu, A.M., 2022. Knowledge diffusion, trade and innovation across countries and sectors. Am. Econ. J.: Macroecon. 14 (1), 104–145. doi:10.1257/mac.20200084.

Cameron, A.C., Gelbach, J.B., Miller, D.L., 2011. Robust inference with multiway clustering. J. Bus. Econ. Stat. 29 (2), 238–249. doi:10.1198/jbes.2010.07136.

Coe, D.T., Helpman, E., 1995. International R&D spillovers. Eur. Econ. Rev. 39 (5), 859–887. doi:10.1016/0014-2921(94)00100-E.

Coelli, F., Moxnes, A., Ulltveit-Moe, K.H., 2022. Better, faster, stronger: global innovation and trade liberalization. Rev. Econ. Stat. 104 (2), 205–216. doi:10.1162/rest_a_00951.

De Loecker, J., Goldberg, P.K., Khandelwal, A.K., Pavcnik, N., 2016. Prices, markups, and trade reform. Econometrica 84 (2), 445–510. doi:10.3982/ECTA11042.

Dernis, H., Khan, M., 2004. Triadic patent families methodology. Science, Technology and Industry Working Paper 2004/02. Organization for Economic Co-operation and Development doi:10.1787/443844125004.

Fons-Rosen, C., Kalemli-Özcan, c., Sørensen, B.E., Villegas-Sanchez, C., Volosovych, V., 2019. Foreign investment and domestic productivity: identifying knowledge spillovers and competition effects. Working Paper. http://www.pseweb.eu/ydepot/seance/513324_spillovers_oct10.pdf.

Goldberg, P.K., Khandelwal, A.K., Pavcnik, N., Topalova, P., 2010. Imported intermediate inputs and domestic product growth: evidence from India. Q. J. Econ. 125 (4), 1727–1767. doi:10.1162/qjec.2010.125.4.1727.

Goldschlag, N., Lybbert, T.J., Zolas, N.J., 2020. Tracking the technological composition of industries with algorithmic patent concordances. Econ. Innovat. N. Technol. 29 (6), 582–602. doi:10.1080/10438599.2019.1648014.

Grossman, G.M., Helpman, E., 1991. Trade, knowledge spillovers, and growth. Eur. Econ. Rev. 35 (2–3), 517–526. doi:10.1016/0014-2921(91)90153-A.

Hötte, K., 2023. Demand-pull, technology-push, and the direction of technological change. Res. Policy 52 (5), 104740. doi:10.1016/j.respol.2023.104740.

Keller, W., 2004. International technology diffusion. J. Econ. Lit. 42 (3), 752–782. doi:10.1257/0022051042177685.

Keller, W., 2010. International trade, foreign direct investment, and technology spillovers. In: Hall, B.H., Rosenberg, N. (Eds.), Handbook of the Economics of Innovation, Volume 2. Elsevier, pp. 793–829. doi:10.1016/S0169-7218(10)02003-4.

Keller, W., 2021. Knowledge spillovers, trade, and FDI. NBER Working Paper 28739. National Bureau of Economic Research doi:10.3386/w28739.

Kukharskyy, B., 2020. A tale of two property rights: knowledge, physical assets, and multinational firm boundaries. J. Int. Econ. 122, 103262. doi:10.1016/j.jinteco.2019.103262.

Liu, E., Ma, S., 2023. Innovation networks and R&D allocation. NBER Working Paper 29607. National Bureau of Economic Research doi:10.3386/w29607.

Lucas Jr., R.E., Moll, B., 2014. Knowledge growth and the allocation of time. Journal of Political Economy 122 (1), 1–51. doi:10.1086/674363.

Lybbert, T.J., Zolas, N.J., 2014. Getting patents and economic data to speak to each other: an algorithmic links with probabilities approach for joint analyses of patenting and economic activity. Res. Policy 43 (3), 530–542. doi:10.1016/j.respol.2013.09.001.

MacGarvie, M., 2006. Do firms learn from international trade? Rev. Econ. Stat. 88 (1), 46–60. doi:10.1162/rest.2006.88.1.46.

Perla, J., Tonetti, C., 2014. Equilibrium imitation and growth. J. Polit. Econ. 122 (1), 52–76. doi:10.1086/674362.

Sampson, T., 2023. Technology gaps, trade, and income. Am. Econ. Rev. 113 (2), 472–513. doi:10.1257/aer.20201940.

Shu, P., Steinwender, C., 2019. The impact of trade liberalization on firm productivity and innovation. In: Lerner, J., Stern, S. (Eds.), Innovation Policy and the Economy, Volume 19. The University of Chicago Press, pp. 39–68. doi:10.1086/699932.

Topalova, P., Khandelwal, A., 2011. Trade liberalization and firm productivity: the case of India. Rev. Econ. Stat. 93 (3), 995–1009. doi:10.1162/REST_a_00095.