Centre for Economic Performance

Discussion Paper

ISSN 2042-2695

No. 1960
November 2023

Exporting ideas: Knowledge flows from expanding trade in goods

Philippe Aghion
Antonin Bergeaud
Timothee Gigout
Matthieu Lequien
Marc Melitz

THE LONDON SCHOOL OF ECONOMICS AND POLITICAL SCIENCE

Economic and Social Research Council

Abstract
We examine the effect of entry by French firms into a new export market on the dynamics of their patents' citations received from that destination. Applying a difference-in-differences identification strategy with a staggered treatment design, we show that: (i) entering a new foreign market has a significant impact on the long-run flow of citations; (ii) the impact is mostly driven by the extensive margin; (iii) inventors in destination countries patent mostly in products that do not directly compete with those of the exporting firm; (iv) the spillover intensity decreases with the technological distance between the exporting firm and the destination.

Keywords: international trade, spillover, innovation, patent
JEL Codes: O33; O34; O40; F10; F14

This paper was produced as part of the Centre's Growth Programme. The Centre for Economic Performance is financed by the Economic and Social Research Council.

We are grateful for useful comments and suggestions to Matthieu Bussière, Gregory Corcos, Jérome Héricourt, Isabelle Méjean, Andreas Moxnes, Jose De Sousa, Claudia Steinwender and seminar participants at INSEAD, Collège de France, CREST, Harvard, Banque de France, Banca d'Italia, Insee, the AFSE 2021 conference, the 2021 NBER Summer Institute on International Trade, the 2022 BSE Summer Forum and the University of Bologna 2022 Workshop on International Trade and Intellectual Property Rights. This work is supported by a public grant overseen by the French National Research Agency (ANR) as part of the "Investissements d'Avenir" program (reference: ANR-10-EQPX-17 - Centre d'accès securisé aux données – CASD)
Philippe Aghion, College de France, INSEAD, London School of Economics and Centre for Economic Performance at LSE. Antonin Bergeaud, HEC Paris and Centre for Economic Performance at LSE. Timothee Gigout, Banque de France and College de France. Matthieu Lequien, Insee. Marc Melitz, Harvard University and NBER.

Published by
Centre for Economic Performance
London School of Economic and Political Science
Houghton Street
London WC2A 2AE

1 Introduction

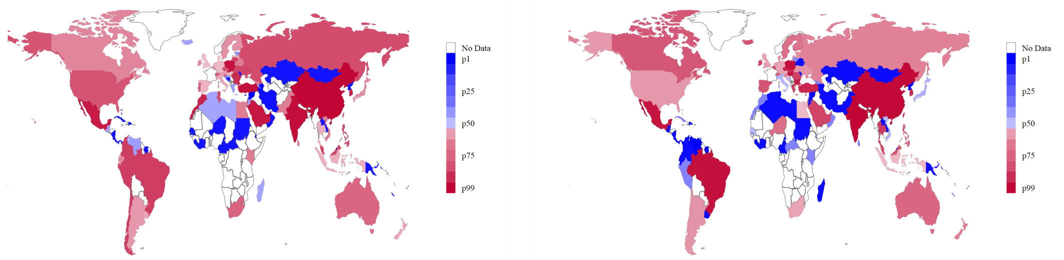
Modern growth theory predicts that international trade should enhance productivity growth for several reasons. First, trade allows potential innovators to sell to a larger market; and by *increasing market size*, trade increases the size of ex-post rents that accrue to successful innovators, thereby encouraging R&D investments. Second, trade *raises competition* in product markets, which in turn encourages innovation that aims to escape competition among the most advanced firms. Third, leaving aside these within-firm impacts on productivity, trade induces selection effects that favor more innovative firms. And finally, trade also induces *knowledge spillovers* that allow producers in recipient countries to catch up with the technological frontier (see Melitz and Redding, 2021 for a review). Using detailed trade and patent data, we focus on this last channel and show that firms export more than goods. They also export the ideas and technologies embedded in these goods, which then show up in the form of new innovations developed in the destination country.

To motivate our analysis, we document that the countries to which France exports are closely related to the countries whose patents cite prior French patents. In Figure 1a, we plot the long difference change in the number of French exporters from 1995 to 2012 (the difference between the number of French exporters in 2012 and the number in 1995) across countries. Each color corresponds to a decile in the long-difference distribution across countries. Dark red (dark blue) corresponds to countries with the largest (resp. smallest) increase in the number of exporters from 1995 to 2012. In Figure 1b, we plot the long difference change in the number of citations to French patents from 1995 to 2012 across countries; again the dark red (dark blue) color refers to countries in the highest (resp. lowest) decile in terms of long difference increases in citations. We see that countries experiencing the largest increase in the number of French exporters also experience the largest increase in patent citations to French innovations over the same time period. The regression coefficient between the two long differences is equal to 1.62 (with a standard error of 0.22).

To more directly measure the spillover from the French exporters to a destination to local innovation, we exploit comprehensive patent and French export data during the 1995-2012 period. For every year and potential export destination, we construct a citation count for each exporters' patents. These citations come from new patents introduced in that year by firms or inventors operating in the destination country. We then investigate how the pattern of a French firm's citation count in a destination changes over time whenever it starts exporting to that destination. Increases in a new exporter's citations represent new patents recorded in that destination subsequent to the exporter's entry in that destination. Those patents citing the French exporter represent a measure of its technological influence in that destination. We use the timing of the exporter's entry into a market relative to its citations in that market to infer a causal relationship between the

Figure 1: Evolution of Trade and Innovation Linkages

(a) Number of French Exporters (b) Patent Citations



Notes: Evolution in the number of French exporters in each country (left-hand side panel) and the number of citations received from each country (right-hand side panel) between 1995 and 2012. Colors correspond to different deciles in the corresponding quantity.

two.

More specifically, we use a difference-in-differences (DiD) strategy to analyze the dynamic response of patent citations to a French firm's export market entry in a particular year. We rely on a staggered treatment adoption design that exploits the fact that firms enter their export markets at different times. The identifying assumption is the existence of a common trend between entrants and non-entrants: inventors in a foreign country would have cited French entrants at the same rate as non entrants had the French firms not entered the country. To strengthen the plausibility of this assumption, we exploit the high dimensionality of our dataset and incorporate fixed effects for both firm-year and destination-technology field-year to capture common shocks to treated and not yet treated observations. We therefore isolate variations coming from within firm change in the flow of citations across all of its current and future export destinations after absorbing destination effects. Additionally, we address potential weighting issues introduced by conventional estimators in staggered designs by drawing on the latest improvements in DiD estimation methods from de Chaisemartin and d'Haultfoeuille (2020).

Our first finding is that exporting to a new foreign market increases the flow of triadic citations received by the exporter from firms in that market (i.e. citations from a patent belonging to a triadic family, see OECD, 2009). More precisely, entering a new foreign market has a significant and positive impact on its triadic patents' citations starting three years after export market entry and sharply increasing over time until 11 years after entry at which point citations start petering out. The underlying idea is that entry into this new market raises the visibility of the exporter's technology to domestic firms in the market. Those domestic firms can then more readily generate further innovations that build upon that technology, conditional on the host country's degree of absorptive capacity (Cohen and Levinthal, 1989). Our second finding is that entering a new foreign market has a greater impact on the extensive margin – the probability of obtaining at least one triadic citations – than the intensive margin – the probability of obtaining many citations.

These findings are robust to many different specification checks: 1) the choice of functional form for the left-hand side citation variable; 2) additional fixed effects; 3) dropping from the sample citations added by the patent office or by the patent examiner; and 4) restricting our sample to foreign destinations with no foreign direct investment ties with the exporting firm, thus controlling for citations generated by the exporter-owned affiliates toward their parent company.

Our third finding is that new patents in a destination country involve products that do not directly compete with those produced by the French exporter: the spillovers involving non-competing products are significantly higher than those associated with competing products. Finally, we examine how the magnitude of the knowledge spillovers vary with the technological distance between the French exporter and the destination country, as measured by differences in patent quality. We find that the spillover intensity is hump-shaped with a peak when the quality of the new patents in a destination lags slightly behind that of the French exporter. This is consistent with the view that firms in a destination country with substantially lower "absorptive capacity" are unable to take advantage of the exporter's more advanced technology. At the other end, destinations with the most advanced technologies may find the French exporter's technology less valuable relative to their own technologies. Conversely, firms situated in countries midway on the technological spectrum seem to gain the most, benefiting from an R&D-enhanced technology transfer from exporting firms, echoing the dual functions of R&D outlined by Griffith et al. (2004).

Our analysis relates to several other strands of literature. There is first the literature on spillovers and international trade. Seminal works by Coe and Helpman (1995) and Coe et al. (2009) demonstrate that a nation's productivity is intricately linked to the R&D and innovation levels of its trading partners, with the strength of these spillovers intensifying alongside the nation's trade openness. The use of patent data, especially citation networks as a means to track the international diffusion of innovation, was further advanced by Jaffe and Trajtenberg (1999).[1] Further evidence of the international dimension of such spillovers is provided by Branstetter (2001) and Keller (2004), suggesting a potential link to international trade and globalization. Keller and Yeaple (2009); Keller (2010) and Aitken and Harrison (1999) explore potential mechanisms and highlight the significant role played by firms operating across borders in driving the global distribution of technology and underscore the dual conduits of trade and multinational activity as vectors for these externalities. Our work adds to this discourse by pinpointing a causal effect of exports on the diffusion of innovation in the destination country that operates even in the absence of foreign direct investment.

More closely related to our analysis is MacGarvie (2006), which uses panel

[1] See also Eaton and Kortum (1996); Peri (2005); Cotterlaz and Guillouzouic (2020).

data on French manufacturing firms and their export links with eight developed countries to reveal that these firms' patents more frequently cite patents from countries they import from. The depth of our data, combined with our dynamic estimation method, allows us to thoroughly assess the long-run responses of foreign triadic citations well before and after a firm's initial export activity. This enables us to uncover the nuanced effects of trade on knowledge dissemination and to more accurately determine the timing and causality of these impacts.[2]

Our work contributes to the recent body of research on the nexus between trade and innovation. It aligns with studies examining the impact of imports on innovation (Bloom et al., 2016; Autor et al., 2020; Bombardini et al., 2017; Aghion et al., forthcoming) as well as those exploring the relationship between exports and innovation (Lileeva and Trefler, 2010; Aghion et al., 2022). While existing studies primarily focus on the effects of trade in terms of competition and market expansion, our research isolates and quantifies the technological spillovers arising from trade activities.

And lastly, our methodology share similarities with the literature on academia, scientists and citations. Azoulay et al. (2010) and Jaravel et al. (2018) analyze the impact of an inventor's death on the subsequent innovation and income patterns of the inventor's surviving coauthors. Waldinger (2011) analyzes the impact of the dismissal of Jewish scientists by the Nazi government in Germany in the 1930s. And Watzinger et al. (2017, 2018) analyze the impact of the mobility of scientists across German universities on local citations to their work. Our paper develops a way to adapt these experimental settings for firms in the context of the analysis of trade and technology spillovers.

The remaining parts of our paper are organized as follows. Section 2 presents the data and details of our empirical strategy. Section 3 presents our main results and performs robustness checks. Section 4 decomposes the spillover across products – separating out products that directly compete with those produced by the French exporter. Section 5 decomposes the spillover across characteristics of the exporter and the destination country, highlighting how the strength of the spillover varies with both the quality of the exporter's patents and the level of technological development in the destination country.

[2] Our analysis uncovers a short-term effect corresponding to a 23% increase from the sample mean, and a long-term effect of up to 107%, as opposed to the 71% short-term increase found in MacGarvie (2006).

2 Data and Methodology

2.1 Data

We build a database that covers all French firms and links export, production, and innovation/citation data from 1994 to 2012. Our database is built from three separate sources. First, detailed customs data provide information on French exports by firm, destination country, and product (covering over 10,000 different product categories consistent with 8-digit CN codes).[3] From this database, we extract the date of first entry into a foreign market for each firm (using exports in 1993 to infer entry in 1994).

Our second data source is the Insee-DGFiP administrative fiscal dataset (FICUS-FARE). This dataset is drawn from compulsory reporting to fiscal authorities in France, supplemented by further census data collected by Insee. We use this data to assign each firm to a 4-digit (NACE) sector and compute their labor productivity as the ratio of value added over employment in full time equivalent.

Our third data source is the Spring 2016 PATSTAT dataset from the European Patent Office. This dataset contains detailed information on all patent applications from most patent offices worldwide, including information on the network of patent linkages via citations. We use it to measure the French firms' patent stocks and the citations from foreign inventors across countries to French firms' patents and match them using the matching algorithm developed by Lequien et al. (2019) and used in e.g. Aghion et al. (2022), which generates matches between all French firms with more than ten employees and their patent applications. We restrict the set of citations, i.e. new foreign patents citing a French patent, to "triadic citations", that is, a citation from a foreign patent belonging to a family of triadic patents to a French patent. A family of triadic patents is a family of patents (patents related to the same invention but filed in different patent offices) that include publications in the European Patent Office (EPO), the Japan Patent Office (JPO) and the United States Patent and Trademark Office (USPTO). Filing applications in all three of those major patent offices represents a significant cost for the patenting firm. Those patents represent innovations that were deemed valuable enough to warrant the additional protection coverage (Lanjouw, 1998, Harhoff et al., 2003, Squicciarini et al., 2013).

Patent citations are commonly utilized as indicators of the knowledge flow between entities (Jaffe and Trajtenberg, 1999). However, not every citation necessarily signifies a knowledge spillover between the owner of the citing patent and the cited one. Instead, it may reflect the patent examiner's awareness of prior art that the inventor was not privy to (Alcacer and Gittelman, 2006). Within the PATSTAT patent database, citations fall into three primary categories (Figure A.1 in

[3] i.e. one extra level of disaggregation compared to HS6 UN Comtrade harmonized data. Customs data are from the *Direction Générale des Douanes et Droits Indirects*.

Appendix A): (i) those added by the applicant prior to filing (approximately 34%), (ii) those included during the patent office's initial search process (around 80%), and (iii) those inserted by the examiner (about 0.5%). Yet, even if a citation is not directly added by the applicant, the latter sometimes discloses to the patent office "the trade names and providers of any goods or services in competition with the goods" that inspired the invention. The patent office then finds the relevant patent and adds the citation. In this case, the patent added by the patent office during the search may still reflect knowledge spillovers between patent owners. Nevertheless, we check the robustness of our main result by estimating our baseline model on a sub-sample restricted to citations added by the applicant. We find similar effects that follow the same dynamics with the same order of magnitude.

The patent data also provide us with the main technological field of each patent as defined by the World Patent Intellectual Property Organisation. This classification assigns the finer International Patent Classification (IPC) codes to 35 large categories - "technological fields" - that balance both size and homogeneity. We use this patent level information to assign a technological field to each firm based on the most frequent field in its stock of patents.

Finally, we identify the French exporters' foreign affiliates using the Insee LIFI dataset. Specifically, we use the "fichier des entreprises" table to back out the date when a French parent company opens or acquires an affiliate in a foreign country.

We seek to measure the knowledge spillovers from new French exporters-innovators who enter an export destination during our 1995-2012 sample years. There are 6,753 such firms who enter one of 156 destinations during our sample years and own at least one patent. Those firms account for 125,700 patents that have generated 614,847 citations (new patents) in foreign destinations. Those entries represent 139,954 potential knowledge spillover links between a foreign patent and a French firm.

2.2 Identification strategy

We want to measure how a French firm's entry into a new export market affects its flow of new citations (to its patents) from that destination. The most natural approach is to aggregate our data at the firm-destination level. More specifically, we estimate how a firm f's entry into a new export market j affects subsequent citations $Y_{f,j,t}$ in year t to its stock of existing patents received from firms located in that destination j. Our baseline specification is:

$$Y_{f,j,t} = \sum_{k>0} \beta_k \times 1_{E_{f,j}=t-k} + \chi(f,j,t) + \epsilon_{f,j,t}. \tag{1}$$

The generic control function $\chi(f,j,t)$ incorporates characteristics of the firm f and destination market j in year t. In this baseline specification, we use a set of firm-year and destination-technological field-year fixed effects. The main coefficients of interest β_k capture the additional citations that can be linked to firm f's export market entry k years before the citation date t (entry in year $t-k$).

We use a DiD estimation with a staggered treatment adoption design, which relies on variation in the year of export market entry for a firm. The identifying assumption is the existence of a common trend between entrants and non-entrants: inventors in a destination j would have cited French entrants at the same rate as non-entrants had the French firms not entered that destination.

Two reasons support this assumption: First, the sample is restricted to innovative active exporters (firms with a positive stock of patents that export during the sample period) and firm-destination pairs that are eventually "treated" by export market entry. Within this relatively more homogeneous sample, we only exploit differences in the *timing* of entry between two or more export destinations for a given firm. Second, there are no discernible pre-tends linking the entry decision to prior citations.

We implement the de Chaisemartin and d'Haultfoeuille (2020) DiD estimator. This estimation accounts for potential weighting issues that arise with the standard DiD estimator (see for instance Callaway and Sant'Anna, 2019 and Goodman-Bacon, 2018).[4] Thus we estimate β_k in equation (1) using:

$$\mathrm{DID}_k = \sum_{t=0}^{T} \frac{N_t^k}{N_{\mathrm{DID}_k}} \mathrm{DID}_{t,k}, \tag{2}$$

where

$$\mathrm{DID}_{t,k} = \underbrace{\sum_{f,j:E_{f,j}=t-k} \frac{1}{N_t^k} (\overbrace{\tilde{Y}_{f,j,t}}^{\text{Now}} - \overbrace{\tilde{Y}_{f,j,t-k-1}}^{\text{Before}})}_{\text{Treated}} - \underbrace{\sum_{f,j:E_{f,j}>t} \frac{1}{N_t^{nt}} (\overbrace{\tilde{Y}_{f,j,t}}^{\text{Now}} - \overbrace{\tilde{Y}_{f,j,t-k-1}}^{\text{Before}})}_{\text{Not yet Treated}}, \tag{3}$$

where N_t^k denotes the number of firm-destination links treated at date $t-k$ and $N_{\mathrm{DID}_k} = \sum_t N_t^k$. $\tilde{Y}$ is the residualized outcome over a set of fixed effects. In our baseline regression, we include firm-year and destination-technological field-year fixed effects. These capture global innovation shocks in a given market and the firm's innovation intensity. We compute a relatively large set of lags and leads in order to capture the full evolution of citations to firm f 's patents following its

[4] In particular de Chaisemartin and d'Haultfoeuille (2020) show that the coefficients identified by the canonical two-way fixed effect (TWFE) estimator capture a combination of the actual treatment effects and of "weights" effects. In the case of a staggered design, the TWFE mechanically computes negative weights for some periods and groups. In some cases this may result in negative estimated coefficients when the treatment effects are in fact positive. This problem is more acute in the presence of treatment effect heterogeneity, either across groups or across periods. The methodology developed in de Chaisemartin and d'Haultfoeuille (2020), which we follow here, is meant to avoid this problem.

entry into destination j. We cluster standard errors at the destination-sector level as we expect that entry by French firms is not randomly assigned across those markets. This allows for auto-correlation and correlation of the error terms across exporters in this cluster.

In our baseline, we estimate DID_k with the outcome variables $\tilde{Y}$ measured in levels as the number of triadic citations. Thus, $\Delta\tilde{Y}$ captures the average change for the treated destination-firm pairs relative to the untreated. As opposed to a measure in logs, we do not need to drop observations with zero citation flows. We expect a lower frequency of zeroes from the treated destination following entry. Dropping those observations would bias $DID_{t,k}$ toward zero. We show that these results are robust to functional form misspecification in Section 3.2 and to the use of various other transformations for the left-hand side variable in Section B.6.

3 Results

3.1 Baseline firm-level results

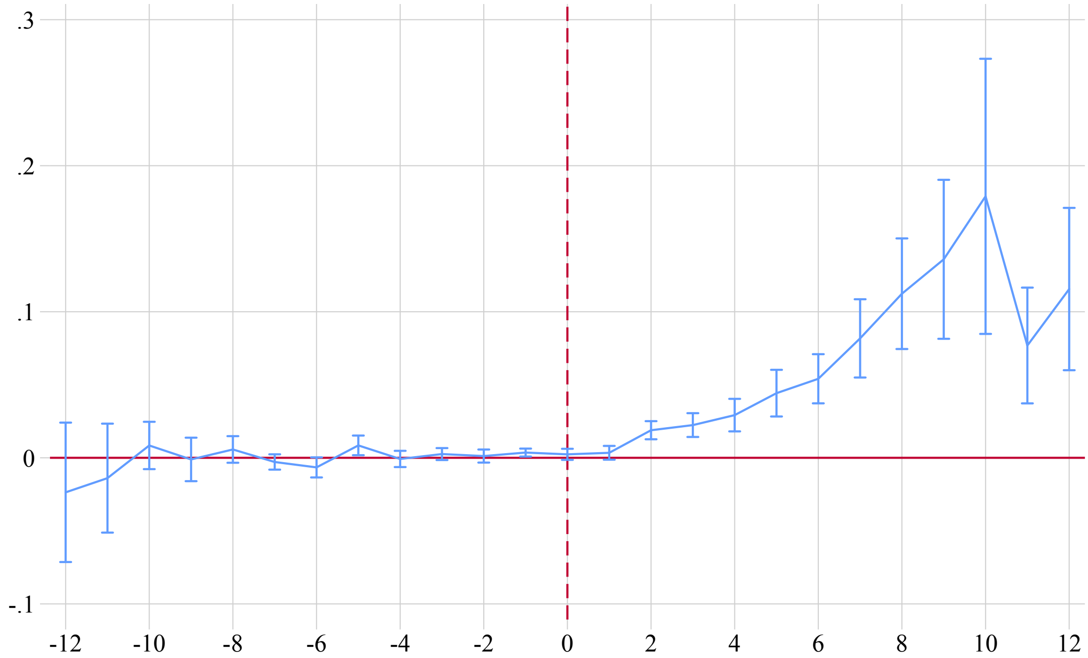
We estimate the parameters DID_k from Equation (1) and report their values $\widehat{\text{DID}}_k$ and associated 99% confidence intervals in Figure 2. Recall that $k = 0$ marks the first year a firm exports to country j. The outcome variable is the count of triadic citations to the French firm's patents. In the regression, we control for destination-technological field-year and for firm-year fixed effects. Standard errors are clustered at the destination-sector level.

As we see in Figure 2, the estimates for the pre-entry coefficients are not significantly different from zero: in other words, there are no pre-trends. Starting 2 years after export market entry, the estimates for the post-entry coefficients are significantly positive, reflecting the time lag for inventors in the destination to build upon the exporter's technology. The triadic citations then steadily and sharply increase over time until they eventually peter out 11 years after entry. This shows that the trade-induced technological spillovers lead to long-lasting bursts of innovation in the destination country.

Magnitude Quantitatively, new exporters receive on average an additional 0.067 citations for their patents from that destination during the 12 years after entry (compared to a destination that the firm has not yet entered at that time). This corresponds to a 99.1% increase from the 0.068 mean citation rate in our sample.

To assess the magnitude of the full treatment effect, we compute the sum of coefficients at different horizons, starting from the year of treatment. After 6 years, we find a cumulative coefficient of 0.175. Over this 6 year time window post-entry, a new exporter receives an average of 0.583 citations from that destination, whereas a non-exporter (to that destination) receives an average of 0.408 citations.

Figure 2: Baseline estimation of the effect of entry on citations



Notes: This figure presents estimates of the coefficient DID_k from Equation (2). The x-axis represents the value of k, $k = 0$ being the first year the firm exports to country j. 99% error bands, computed with robust standard errors clustered at the destination-sector level, are displayed as blue brackets.

This corresponds to a 71.1% higher level of citations from the destination country. After 12 years, the cumulative coefficient is 0.876. A new exporter receives an average of 1.692 citations from that destination, whereas the non-exporter receives an average of 0.816 citations. Those results and results for other time horizons are presented in Table B.1 in Appendix B.1.

3.2 Entry and the distribution of citation counts

The effect of export market entry on the number of citations to the new exporter may be driven either by an "extensive margin" effect – entry into a destination induces the exporter's patents to be cited for the first time in that destination – or it may be driven by an "intensive margin" effect – the exporter's citations in a destination increase with entry from a positive base level. To quantify those margins, we use the "distribution regression" method developed in Chernozhukov et al. (2013) to estimate the entire conditional distribution of the citation count variable Y. Importantly, this does not require the outcome to have a smooth conditional density as in quantile regressions.

In Figure 3a, each point is estimated using a linear probability model. This corresponds to the effect for the probability of the exporter accumulating more citations than the indicated value (horizontal axis) six years after entry. The fixed effects are the same as in the baseline regression (firm-year and destination-technological field-year) and we once again cluster the standard errors at the destination-sector level (the blue bands show the 99% confidence intervals). The first coefficient corresponds to a regression where the outcome variable is a dummy that is set to one when the citation rate is strictly greater than zero. This is the extensive margin. We find that entry increases the probability of being cited by 0.3 percentage points (p.p.) after 6 years relative to a base probability of 1.4%. The effect is similar for low (below 5)—but positive—citation counts, but then decreases for higher citation counts. Entry increases the probability of receiving more than 10 citations by only 0.12 p.p.; and it increases the probability of receiving more than 20 citations by just 0.05 p.p.. Yet, those lower effects remain significant over the entire distribution of citation counts. We report the full dynamic response for selected citation count thresholds in Appendix B.2. The extensive margin effect on the probability of receiving at least 1 citation increases with time and is the highest after 8 years (+ 0.7 p.p.). To summarize, export market entry has a much more substantial impact on the extensive margin of citations (the exporter's patent first being cited in the destination) than on the intensive margin (further increases in the exporter's citation counts).

In Figure 3b, we plot the sample cumulative distribution for citations in red. This is computed only on not yet treated observations. The counterfactual distribution is plotted in blue. It is obtained by taking the sample distribution and adding the coefficients from the top panel, thereby including the impact of entry.[5]

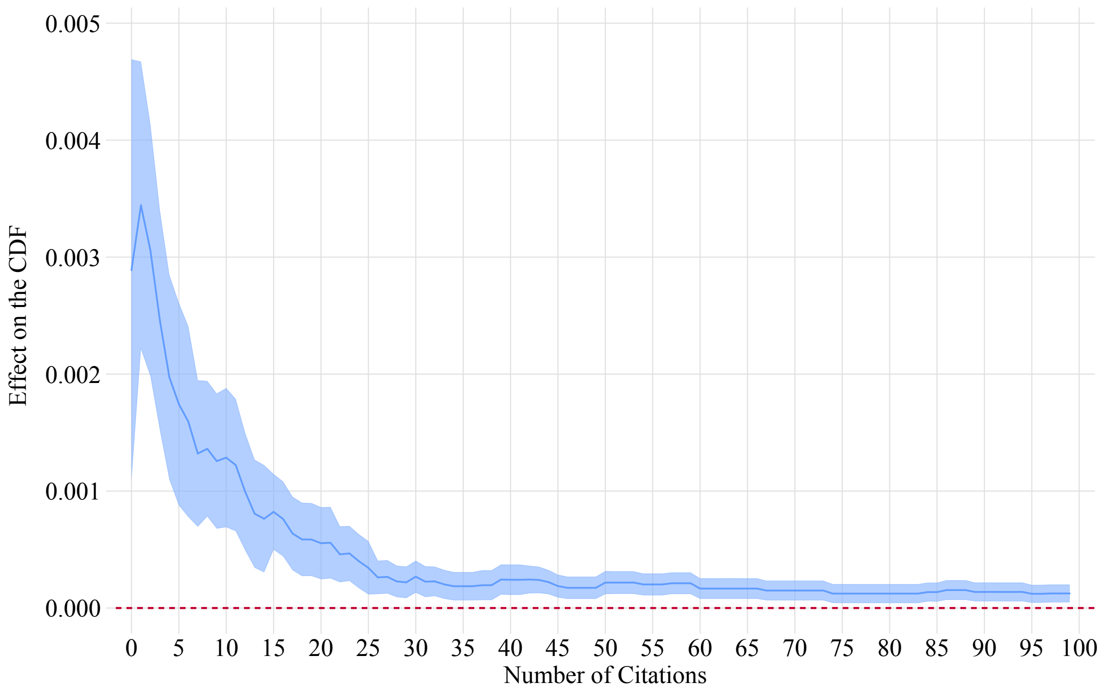
3.3 Robustness checks

Our findings are robust to the choice of functional form for the left-hand side variable (Figure B.5), to alternate fixed effects specifications (Figure B.7), to using the Borusyak et al. (2021)'s DiD estimator (Figure B.6), to dropping from the sample the citations added by either the patent office or the patent examiner (Figure B.2), and to keeping only the patents filed prior to entry (Figure B.3). We also show that while opening an affiliate in a given country is associated with an increase in citations to the parent company from this country (Figure B.8), this mechanism does not drive our results: they are robust to restricting our sample to destinations with no affiliate links with the exporter. This controls for citations by affiliates back to patents owned by their parent company (assuming the French exporter is ultimate parent/owner of the multinational network, see Figure B.4).

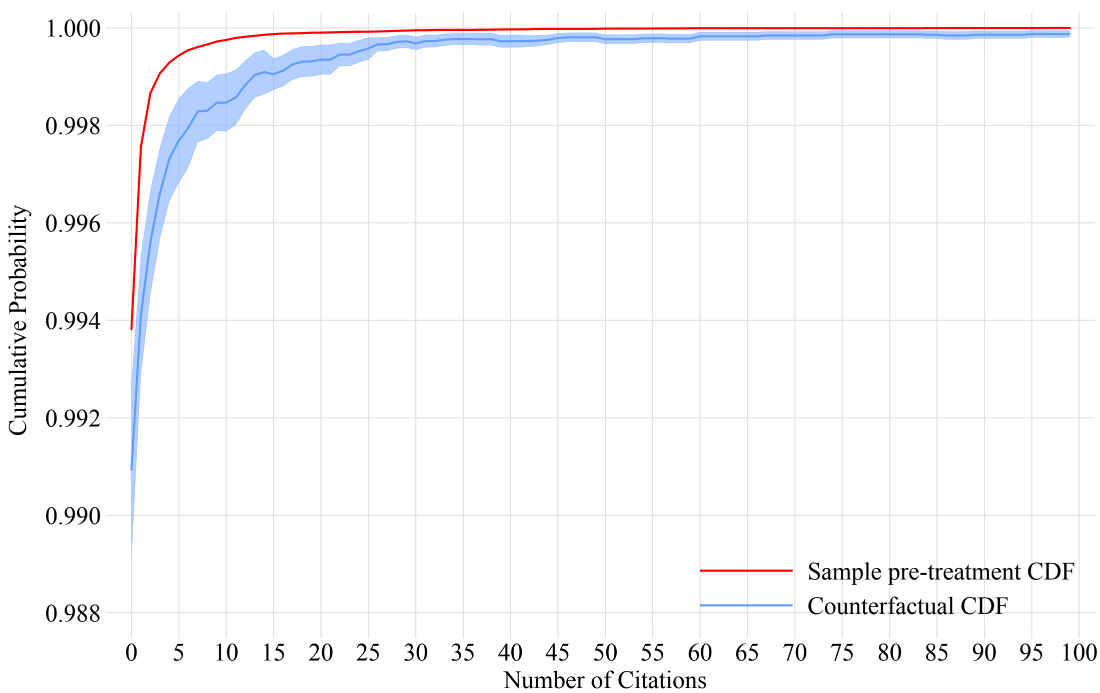
[5] As the sample cumulative distribution function shows, triadic citations are rare events. They become less rare after entry. While this is especially true for firms receiving a small number of triadic citations, it is striking that virtually no firm receives a large number of triadic citations from a country where it has not exported to yet.

Figure 3: Effect of Entry on the Distribution of Citations per Firms

(a) Distribution Regression Coefficients



(b) Cumulative Distribution of Citations per firm



Notes: Panel 3a provides estimates of the coefficients $DID_{k=6}$ associated with the initial entry into a foreign destination on the probability of having a citation rate greater than the amount on the x-axis. 99% error bands, computed with standard errors clustered at the destination-sector level, are displayed as a blue band. In Panel 3b, we plot the sample CDF in red and the estimated counterfactual CDF in blue. For details on distribution regressions see Chernozhukov et al. (2013).

4 The Channels for Knowledge Spillovers

The knowledge spillovers we have uncovered could be driven by horizontal competition forces, by vertical production incentives, or by other synergies that cannot be attributed to either of those channels. An example of the horizontal channel is a new patent related to a product that is similar to the one produced by the new French exporter. An example of the vertical channel is a new patent related to an upstream or downstream product relative to the one produced by the new French exporter. In this section, we evaluate the relative magnitude of these channels for knowledge spillovers. We also further separate out knowledge spillovers that are associated with multinational ties between the French exporter and the destination, regardless of whether the channel is horizontal or vertical.

4.1 Descriptive statistics

In order to measure the contribution of those distinct spillover channels, we first need to classify the cited and citing patents according to the products with which they are associated. We start with the most detailed technological IPC codes of each patent in our database and convert it into a 6-digit HS product code and then associate this product with an end-use category: consumption, intermediate processed, intermediate primary, and capital.[6] We attribute the spillover to the horizontal competition channel when the cited-citing patents share the same product code. We then attribute the spillover to the vertical channel when the patent pair have different end-use categories. This leaves a group of "non-competing" patent pairs that share the same end-use category but have different product codes. Finally, any citations from the previous categories where the citing patent destination contains affiliates of the French exporter is assigned to a fourth category, which thus captures within-firm knowledge flows in addition to pure spillovers.

Figure B.10a reports the frequency of the cited-citing pairs using all end-use categories, further separating out pairs related to the same product code and to the same multinational network. Figure B.10b reports the frequency across the aggregated groups that we defined previously: the bulk of the patent pairs (79%) are associated with non-competing products, split relatively evenly between patent pairs with the same end-use category (horizontal) and different end-use categories (vertical). 11% of the cited-citing patent pairs are associated with competing products. And the remaining 10% of the patent pairs involve French exporters that own an affiliate in the destination country, hence spillovers associated with these pairs are more likely to reflect intra-firm knowledge transfers.

[6] See Appendix A for details on this overall mapping procedure.

4.2 Regression results

We use our baseline identification strategy to measure the effect of export entry on the spillover channel towards the citing firms in the destination countries. We deviate from our baseline in one respect: we normalize the flow of citations for each type by the pre-entry average citation rate for that specific type. This allows us to compare spillover effects across the different groups of cited-citing patent pairs. We index the groups previously defined and used in Figure B.10b by c. The outcome variable $y^c_{f,j,t}$ is the raw citation counts to exporter f's patents coming from destination j in year t, divided by the average pre-entry citation rate for the corresponding type c. Thus, we estimate:

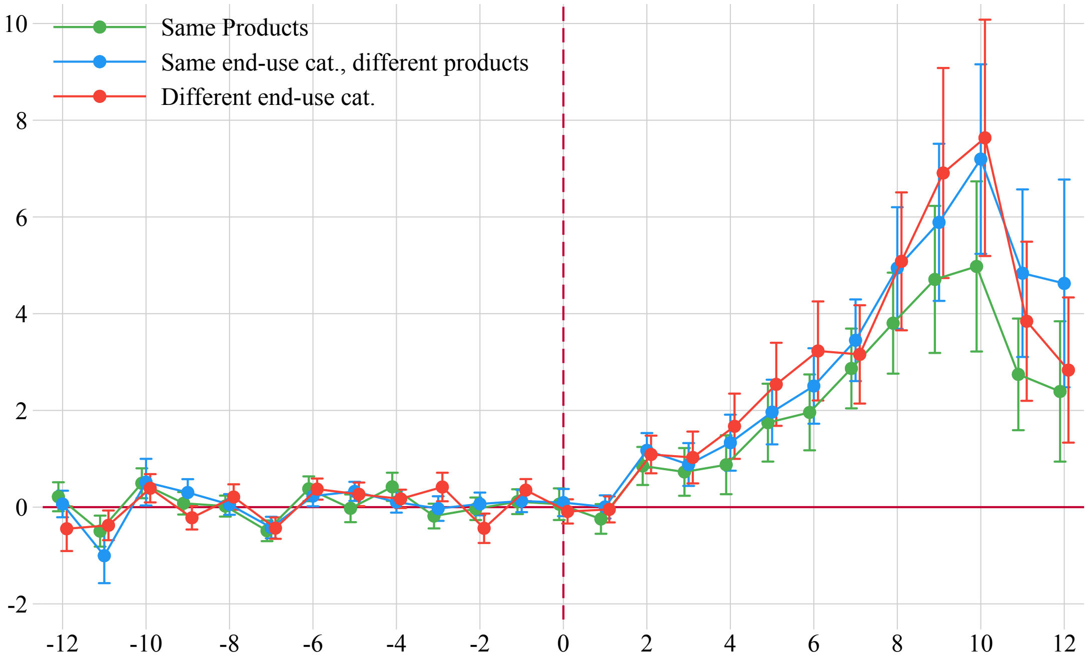
$$y^c_{f,j,t} = \sum_{k>0} \beta^c_k 1_{E_{f,j}=t-k} + \chi(f,j,t) + \epsilon_{f,j,t}. \tag{4}$$

Figure 4 summarizes the estimation results. It shows that export entry has a positive and significant impact on all three types of knowledge spillovers.[7] However, the spillovers from the two "non-competing" channels – both horizontal and vertical – are larger. Those differences between each of the two "non-competing" channels and the "competing" (same product) one are significant in most of the later years, four or more years after export market entry.[8]

[7]The impact is even larger for the fourth group corresponding with destinations with affiliates, as it cumulates pure spillovers effects with within-firm citations.

[8]See Appendix B.9 for details.

Figure 4: Knowledge spillovers by channel



Notes: Figure 4 presents estimates of the coefficient β_k^c from Equation (4) for each sub-sample of diffusion channels. 99% confidence interval are presented. Standard errors are clustered at the destination-sector level.

5 Knowledge spillovers and technology gaps between exporter and destination

In this section, we investigate how knowledge spillovers vary with the technology gap between the exporter and the corresponding sectors in the destination. We measure those gaps using the average difference in patent quality between France and the destination in each technological field.

Our conjecture is that it should be more difficult for innovators in destinations with technologies far behind the French ones to absorb and build upon them.[9] At the other extreme, innovators in destinations with technologies closer to the technological frontier than the French ones should find little or less value in innovating upon them. Accordingly, destinations with technologies closer to the French ones should exhibit the highest spillovers.

[9] In line with the theory of Aghion et al. (2005). See also Van Patten (2021) for a framework that incorporates heterogeneous learning abilities among importers.

5.1 Measuring the technology gaps

To test this conjecture, we follow the innovation literature and the OECD patent data manual to construct a measure of patent quality based on a patent's forward citations. However, it is well known that propensity to patent and cite vary strongly from one technology to another and from one country to another. Therefore, we proceed as follows. For each technological field and destination, we measure proximity to the technology frontier using the number of all citations received from USPTO patents and standardize this number by the number of citations during the same period to USPTO patents in the same technological field and coming from all USPTO patents (denoted $QUAL_{i,j}$).[10] We then measure the proximity to the technology frontier in the same way for the French firm using the French proximity to the frontier for each of the technological fields of the firm's patent portfolio (denoted $QUAL_{i,FR}$). We then construct the technology gap as the difference in those ratios, destination versus France, for the exporter's technological fields. We therefore use the US as a benchmark to standardize the count of citations in each technology and country. In Appendix C, we show that the citation share from the United States is a good proxy for the distance to the technological frontier for the French exporters: The correlation between that citation share and the exporters' labor productivity growth is positive and significant. We also construct a different measure of patent quality based on the number of patent offices protecting the firm's patents; and then show that we can replicate all of our main results using this alternate measure (see Appendix D).

In practice, we proceed in three steps: (i) We collect all the technological field codes for each cited and citing patent in our dataset. (ii) For each destination-technological field pairs, we compute the ratio of citations received from U.S. patents to all citations from U.S. patents to U.S. patents in the same technological field (we use 3-digit IPC codes) over the period 2000-2010; and repeat this exercise for France. Then, in order to aggregate this measure at the firm level, we consider a weighted average of the difference between $QUAL_{i,j}$ in the destination $j \neq France$ and $QUAL_{i,FR}$ across firm f's technological fields:[11]

$$\mathrm{GAP}_{j,f} = \sum_i w_{i,f} \left(\mathrm{QUAL}_{i,j} - \mathrm{QUAL}_{i,FR} \right).$$

[10]Formally, for any technology i and country j, we calculate the proximity to the frontier as:

$$QUAL_{i,j} = \frac{\sum_{p \in (US,.)} \sum_{p' \in (j,i)} Cit(p,p')}{\sum_{p \in (US,.)} \sum_{p' \in (US,i)} Cit(p,p')},$$

where $Cit(p,p')$ is an indicator equal to one if patent p (filed at the USPTO regardless of the technological field) cites patent p' (in country j and technological field i).

[11]We proxy firm f's technological quality by its exposure to the quality of French technological fields. The more a firm patents in technological fields where France gets a high share of U.S. citations, the higher its measure of quality is.

The firm-level technological field weight $w_{i,f}$ is computed using the count of that technological field's appearance in the portfolio of the firm's patents – expressed as a share of the total count. For example, if the car technological field appears twice and the GPS technological field appears once in the full firm's patent history, then the car and GPS technological fields are assigned weights of 2/3 and 1/3.

Figure C.2 in Appendix C shows the cumulative distribution of $GAP_{j,f}$. Approximately 90% of the destinations have negative technology gaps with the corresponding French exporter. For 50% of those destinations, the gap is greater than 30 citations per thousand US citations.

5.2 Knowledge spillover and absorptive capacity

As we previously motivated, we anticipate that destinations with technologies that are closer to those of the French exporter will exhibit the highest spillovers. In other words, that destination's absorptive capacity for the knowledge spillover is highest relative to destinations with technologies that are either far below or far above the French exporter's technology. To test this, we use our technology gap measure to estimate:

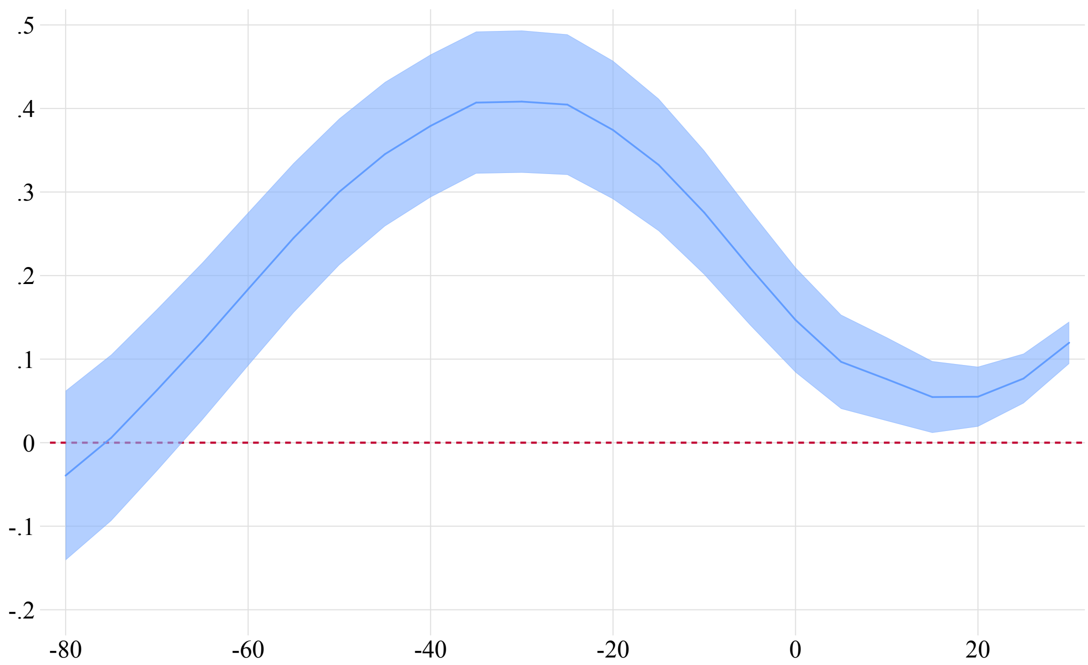
$$Y_{f,j,t} = \sum_{k>0} \left(\beta_k \times 1_{E_{f,j}=t-k} + \gamma_k^g \mathrm{GAP}_{f,j}^g \times 1_{E_{f,j}=t-k} \right) + \chi(f,j,t) + \epsilon_{f,j,t}, \tag{5}$$

where γ_k^g is a time varying coefficient that represents the effect of entry at different technology gap values g, $GAP_{f,j}^g$. To capture the possibility of a non-linear relationship between the technological gap and the trade-induced knowledge spillovers $Y_{f,j,t}$ from the exporter f to destination j, we combine the DiD estimator from de Chaisemartin and d'Haultfoeuille (2020) with a local linear estimator. In words, we estimate a sequence of DiD coefficients locally at various points in the distribution of $\mathrm{GAP}_{f,j}^g$.

The results are shown graphically in Figure 5. We find strong evidence that the strength of the spillovers is hump-shaped in the destination's absorptive capacity relative to the new French exporter. Theoretically, there is no reason to expect that peak spillovers occur when the technology gap between the destination and the French exporter is zero: there are forces pushing down the spillover for destinations with large positive or negative gaps, but no reason to expect that those forces are symmetric. Empirically, we find that they are not: the peak spillovers occur for destinations that lag behind by 30 US citations per thousand citations. This is the median technology gap in our sample. Figure 5 clearly highlights how the strength of those spillovers decreases for destinations with both higher and lower levels of technology gaps. For destinations with positive technology gaps, the spillover effect is nonetheless still positive and significant. For destinations with the lowest negative technology gaps below -70 , accounting for roughly 5%

of our sample, the spillover is no longer significantly different from zero.

Figure 5: Technology gap and the effect of entry on citations



Notes: This figure presents estimates of the coefficient $\gamma_{k=6}$ from Equation (5). This figure plots the effect of the initial entry into a foreign destination estimated locally at a given point of the distribution of the firm-destination technological distance. The dependant variable is the number of citations. We use Gaussian weights with a bandwidth set to 5. 99% confidence interval are presented. Standard errors are clustered at the destination-sector level.

6 Conclusion

Using French firm-level fiscal, customs, and patent citation data over the period 1995-2012, we estimate the impact of export market entry on the citations of the exporter's patents in the destination country. We find a positive and significant effect of entry on those citations. Moreover: (i) entering a new foreign market has a greater impact on the extensive margin – the probability of obtaining at least one or a few triadic citations – than the intensive margin – the probability of obtaining many citations; (ii) inventors in destination countries patent mostly in products that do not directly compete with those of the exporting firm; (iii) the spillover intensity is hump-shaped with respect to the technological distance between the exporting firm and the destination country. Overall, our results validate the notion that trade induces technological spillovers (in line with Coe and Helpman, 1995). And the results are also consistent with Cohen and Levinthal (1989)'s view that spillovers occur conditionally upon the recipient country exhibiting sufficient *absorptive capacity*.

Our main finding that trade induces knowledge spillovers is in line with the notion that trade is a source of cross-country convergence. In addition, fostering development in the destination country increases the country's ability to build upon the innovations brought by foreign exporters, thus inducing a reinforcing loop for productivity growth.

References

Aghion, Philippe, Antonin Bergeaud, Matthieu Lequien, and Marc J. Melitz, "The Heterogeneous Impact of Market Size on Innovation: Evidence from French Firm-Level Exports," *The Review of Economics and Statistics*, 05 2022, pp. 1–56.

_ , _ , _ , _ , and Thomas Zuber, "Opposing firm-level responses to the China shock: Output competition versus input supply," *American Economic Journal: Economic Policy*, forthcoming.

_ , Nicholas Bloom, Richard Blundell, Rachel Griffith, and Peter Howitt, "Competition and Innovation: an Inverted-U Relationship," *The Quarterly Journal of Economics*, 2005, *120* (2), 701–728.

Aitken, Brian J and Ann E Harrison, "Do domestic firms benefit from direct foreign investment? Evidence from Venezuela," *American economic review*, 1999, *89* (3), 605–618.

Alcacer, Juan and Michelle Gittelman, "Patent citations as a measure of knowledge flows: The influence of examiner citations," *The Review of Economics and Statistics*, 2006, *88* (4), 774–779.

Amiti, Mary, Cédric Duprez, Jozef Konings, and John Van Reenen, "FDI and superstar spillovers: Evidence from firm-to-firm transactions," Technical Report, National Bureau of Economic Research 2023.

Autor, David, David Dorn, Gordon H Hanson, Gary Pisano, and Pian Shu, "Foreign competition and domestic innovation: Evidence from US patents," *American Economic Review: Insights*, 2020, *2* (3), 357–374.

Azoulay, Pierre, Joshua S Graff Zivin, and Jialan Wang, "Superstar extinction," *The Quarterly Journal of Economics*, 2010, *125* (2), 549–589.

Bloom, Nicholas, Mirko Draca, and John Van Reenen, "Trade induced technical change? The impact of Chinese imports on innovation, IT and productivity," *The Review of Economic Studies*, 2016, *83* (1), 87–117.

Bombardini, Matilde, Bingjing Li, and Ruoying Wang, "Import Competition and Innovation: Evidence from China," Technical Report 2017.

Borusyak, Kirill, Xavier Jaravel, and Jann Spiess, "Revisiting event study designs: Robust and efficient estimation," *Unpublished working paper, version May*, 2021, *19*, 2021.

Branstetter, Lee G, "Are knowledge spillovers international or intranational in scope?: Microeconometric evidence from the US and Japan," *Journal of international Economics*, 2001, *53* (1), 53–79.

Callaway, Brantly and Pedro HC Sant'Anna, "Difference-in-differences with multiple time periods," *Available at SSRN 3148250*, 2019.

Chen, Jiafeng and Jonathan Roth, "Log-like? ATEs defined with zero outcomes are (arbitrarily) scale-dependent," *arXiv preprint arXiv:2212.06080*, 2022.

Chernozhukov, Victor, Iván Fernández-Val, and Blaise Melly, "Inference on counterfactual distributions," *Econometrica*, 2013, *81* (6), 2205–2268.

Coe, David T. and Elhanan Helpman, "International R&D spillovers," *European Economic Review*, 1995, *39* (5), 859–887.

Coe, David T, Elhanan Helpman, and Alexander W Hoffmaister, "International R&D spillovers and institutions," *European Economic Review*, 2009, *53* (7), 723–741.

Cohen, Wesley M and Daniel A Levinthal, "Innovation and learning: the two faces of R & D," *The economic journal*, 1989, *99* (397), 569–596.

Cotterlaz, Pierre and Arthur Guillouzouic, "The Percolation of Knowledge across Space," *Available at SSRN 3010092*, 2020.

de Chaisemartin, Clément and Xavier d'Haultfoeuille, "Two-way fixed effects estimators with heterogeneous treatment effects," *American Economic Review*, 2020, *110* (9), 2964–96.

Eaton, Jonathan and Samuel Kortum, "Trade in ideas Patenting and productivity in the OECD," *Journal of international Economics*, 1996, *40* (3-4), 251–278.

Goodman-Bacon, Andrew, "Difference-in-differences with variation in treatment timing," Technical Report, National Bureau of Economic Research 2018.

Griffith, Rachel, Stephen Redding, and John Van Reenen, "Mapping the two faces of R&D: Productivity growth in a panel of OECD industries," *Review of economics and statistics*, 2004, *86* (4), 883–895.

Harhoff, Dietmar, Frederic M Scherer, and Katrin Vopel, "Citations, family size, opposition and the value of patent rights," *Research policy*, 2003, *32* (8), 1343–1363.

Jaffe, Adam B and Manuel Trajtenberg, "International knowledge flows: Evidence from patent citations," *Economics of innovation and new technology*, 1999, *8* (1-2), 105–136.

Jaravel, Xavier, Neviana Petkova, and Alex Bell, "Team-specific capital and innovation," *American Economic Review*, 2018, *108* (4-5), 1034–73.

Keller, Wolfgang, "International technology diffusion," *Journal of economic literature*, 2004, *42* (3), 752–782.

_ , "International trade, foreign direct investment, and technology spillovers," in "Handbook of the Economics of Innovation," Vol. 2, Elsevier, 2010, pp. 793–829.

_ **and Stephen R Yeaple**, "Multinational enterprises, international trade, and productivity growth: firm-level evidence from the United States," *The Review of Economics and Statistics*, 2009, *91* (4), 821–831.

Lanjouw, Jean Olson, "Patent protection in the shadow of infringement: Simulation estimations of patent value," *The Review of Economic Studies*, 1998, *65* (4), 671–710.

Lequien, Matthieu, Martin Mugnier, Loriane Py, and Paul Trichelair, "Linking patents to firms: insights with French firms," 2019.

Lileeva, Alla and Daniel Trefler, "Improved Access to Foreign Markets Raises Plant-Level Productivity… for Some Plants," *Quarterly Journal of Economics*, 2010, *125* (3), 1051–1099.

Lybbert, Travis J. and Nikolas J. Zolas, "Getting patents and economic data to speak to each other: An 'Algorithmic Links with Probabilities' approach for joint analyses of patenting and economic activity," *Research Policy*, 2014, *43* (3), 530–542.

MacGarvie, Megan, "Do firms learn from international trade?," *Review of Economics and Statistics*, 2006, *88* (1), 46–60.

Melitz, Marc J. and Stephen J Redding, "Trade and innovation," Technical Report w28945, National bureau of economic research 2021.

OECD, *OECD Patent Statistics Manual*, OECD publishing, 2009.

Patten, Diana Van, "International Diffusion of Technology: Accounting for Heterogeneous Learning Abilities," 2021. Mimeo.

Peri, Giovanni, "Determinants of knowledge flows and their effect on innovation," *Review of economics and Statistics*, 2005, *87* (2), 308–322.

Squicciarini, Mariagrazia, Hélène Dernis, and Chiara Criscuolo, "Measuring patent quality: Indicators of technological and economic value," 2013.

Waldinger, Fabian, "Peer effects in science: Evidence from the dismissal of scientists in Nazi Germany," *The Review of Economic Studies*, 2011, *79* (2), 838–861.

Watzinger, Martin, Lukas Treber, and Monika Schnitzer, "Universities and Science-Based Innovation in the Private Sector," 2018. mimeo LMU Munich.

_ , Markus Nagler, Lukas Treber, and Monika Schnitzer, "Mobility of Scientists and the Spread of Ideas," Discussion paper series in econompics and management 17-22, German Economic Association of Business Administration 2017.

APPENDIX

A Patent Data

Our patent data source is the Spring 2016 PATSTAT dataset from the European Patent Office. This dataset contains detailed information on all patent applications from most patent offices worldwide, including information on the network of patent linkages via citations and the origin of these citations (see Figure A.1 and discussion in Section 2). We use it to measure French firms' patent stocks and the citations from foreign inventors to French firms' patents. PATSTAT is a set of tables, each containing different types of information, that can be combined together in many different ways. Table A.1 describes specifically which tables were used in this paper.

Table A.1: PATSTAT

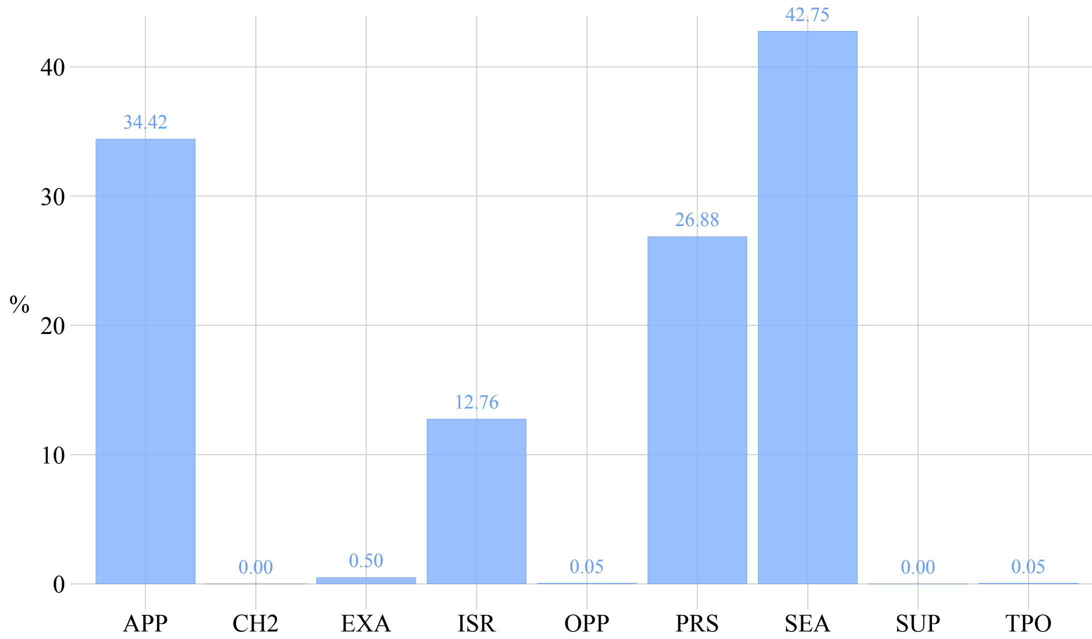
Table	Content & Purpose
TLS201	Application filling date of each patent & application authority of each patent to determine which patent families are "triadic" (patents related to the same invention but filed in at least the European Patent Office (EPO), the Japan Patent Office (JPO) and the United States Patent and Trademark Office (USPTO).
TLS207	List of applicants of each patent
TLS906	Country of residence of each applicant
TLS212	Citation linkage between each pairs of citing-cited patents & nature of the citation (added by the applicant, the examiner, etc.)
TLS230	Aggregate technological field for each patent that we use to compute the fixed effects.
TLS209	IPC technological codes associated with each patent. Used to compute technological field to technological field citations

Notes: This table describes which PATSTAT tables were used in this paper.

Additionally, we convert IPC (International Patent Classification) technological code to HS (Harmonized System) product code using the matching developed by Lybbert and Zolas, 2014. Their probabilistic matching is based on the lexicographic proximity between industry and patent description. The HS codes are then converted to BEC (Broad Economic Categories) end-use categories using the appropriate UNSTAT concordance table (https://unstats.un.org/unsd/classifications/Econ).

Finally, each French firm has a unique identifying number (Siren) across all French administrative databases while patent offices identify firms using only their name. We use the matching algorithm developed by Lequien et al. (2019) to link each patent application with the Siren numbers of the corresponding French firms.

Figure A.1: Citation Origins



Notes: This figure shows who added the citation to a given patent. APP: Applicant added citations; ISR/PRS/SEA: citations added during the 1st phase of the research for prior art by the patent office; EXA: Examiner added citations. A citation may belong to several categories.

B Robustness

B.1 Cumulative results at different horizons

Table B.1: Results Summary

	$k = 3$	$k = 6$	$k = 9$	$k = 12$
$\sum_k DID_k$	0.047	0.175	0.505	0.876
% Increase	23.09	42.82	82.45	107.34
$E(C) \times k$	0.204	0.408	0.612	0.816

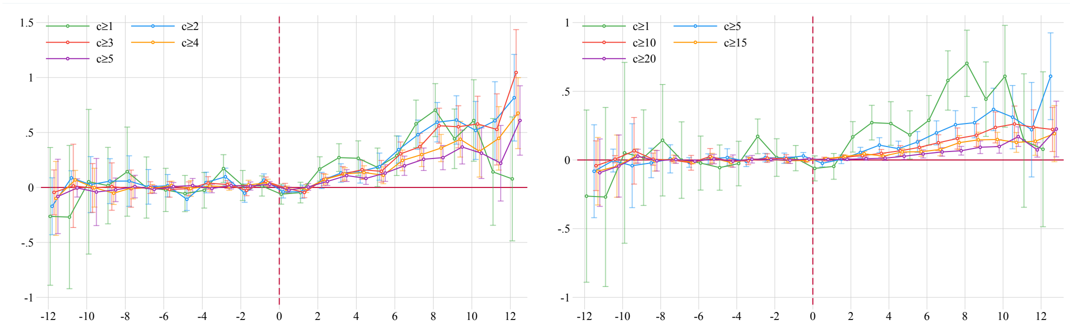
NOTE: This table presents estimates of the cumulutive effect $\sum_{k>0} DID_k$ at different horizons and its percentage increase relative to the sample mean.

B.2 Distribution regressions: dynamic results

We reports the results from estimating the same specification as in Section 3.2 for all pre- and post-entry coefficients for selected thresholds of the triadic citations variable in Figure B.1.

Figure B.1: Fully dynamic distribution regressions

(a) Small values of C (b) High values of C



Notes: This figure presents the results from estimating Equation (2) where the outcome variable is a dummy variable equals to 1 when the French firm is cited more than a certain threshold of citations by the destination country that year. Everything else in unchanged.

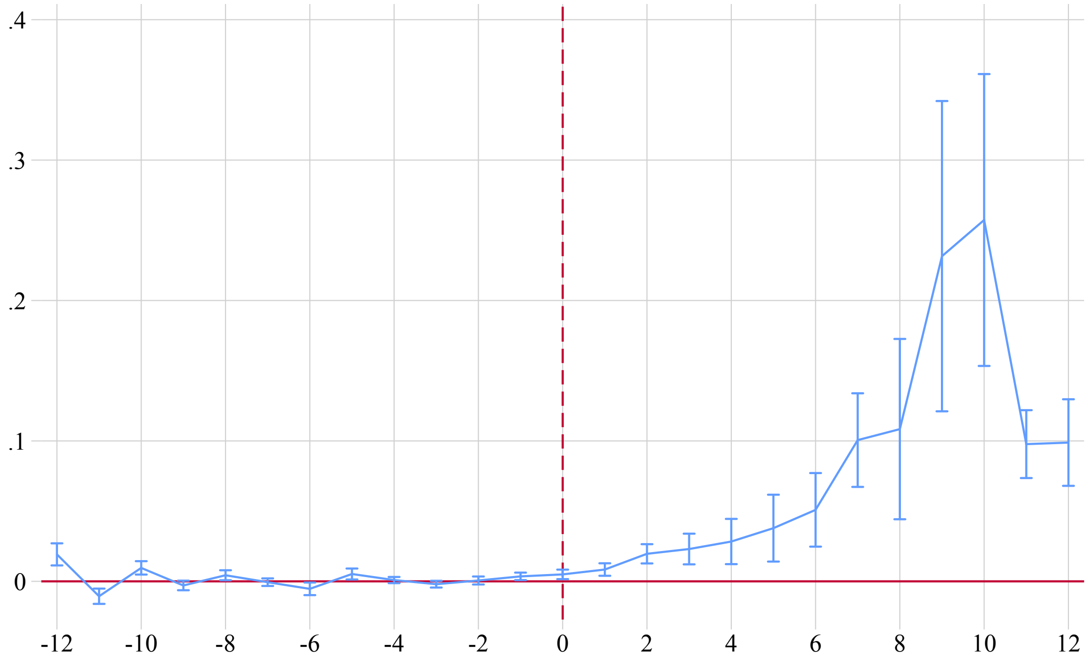
B.3 Citations added by the applicant

Here we estimate our baseline model, but we restrict our attention to citations added by the foreign patent applicant, thus we ignore the citations added during the initial search process by the patent office, and we also ignore the citations added by the patent examiner. Fixed effects and the clustering of standard errors remain the same as in the baseline regression. In Figure B.2 the outcome variable is the raw count of triadic citations Y as in the baseline regression. Entry entails a similar impact as in the baseline regression. We can therefore exclude that the knowledge spillover we uncover only touches the patent examiner.

B.4 Citations to the pre-entry stock of patents

Keeping everything else unchanged, we now estimate the effect of entry on the citations to the pre-entry stock of patents of the exporting firm. The export entry generated exposure and citations to existing patents, as the positive coefficients presented in Figure B.3 show.

Figure B.2: Firm level regression - Applicant added Citations



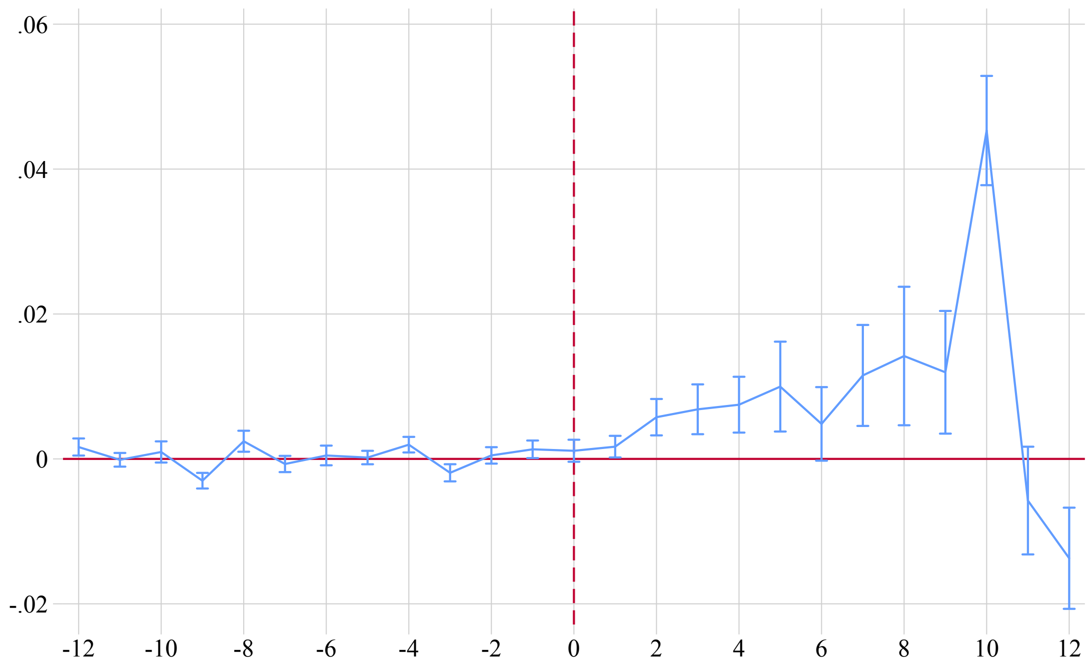
Notes: Panel B.2 presents the results from estimating Equation 2 after restricting citations to applicant added citations only. Everything else in unchanged.

B.5 Trade without affiliates

The affiliates of a French parent company may conduct R&D activity and would naturally be inclined to cite intellectual property developed and owned by their parent company. Furthermore with a foreign affiliate in a country, foreign firms in this country may cite the French firm's patents because they observe the production process or the products sold directly and locally by the affiliate rather than because they observe the exported product by the parent company. To account for these mechanisms, we reproduce our baseline results on a sample restricted to foreign destinations where the firm never opens an affiliate. As Figure B.4 shows the dynamics pattern of citations remains the same as in the baseline specification, even though the smaller coefficients indicate that part of the effect captured in the baseline estimates might be due to affiliates innovating in the destination country – which also constitutes a technological transfer.[12]

[12]Excluding countries with affiliates, 0.11 extra citations are recorded in the 10^{th} year after entry, which corresponds to an increase of 239% over the sample average of 0.046 citations. The corresponding increase in the baseline regression is 410%.

Figure B.3: Effect of entry on pre-entry patents



Notes: This figure presents estimates of the coefficient DID_k from Equation (2). The sample excludes citations to patents that did not exist at the time of entry. The x-axis represents the value of k, $k = 0$ being the first year the firm exports to country j. 99% error bands, computed with robust standard errors clustered at the destination-sector level, are displayed as blue brackets.

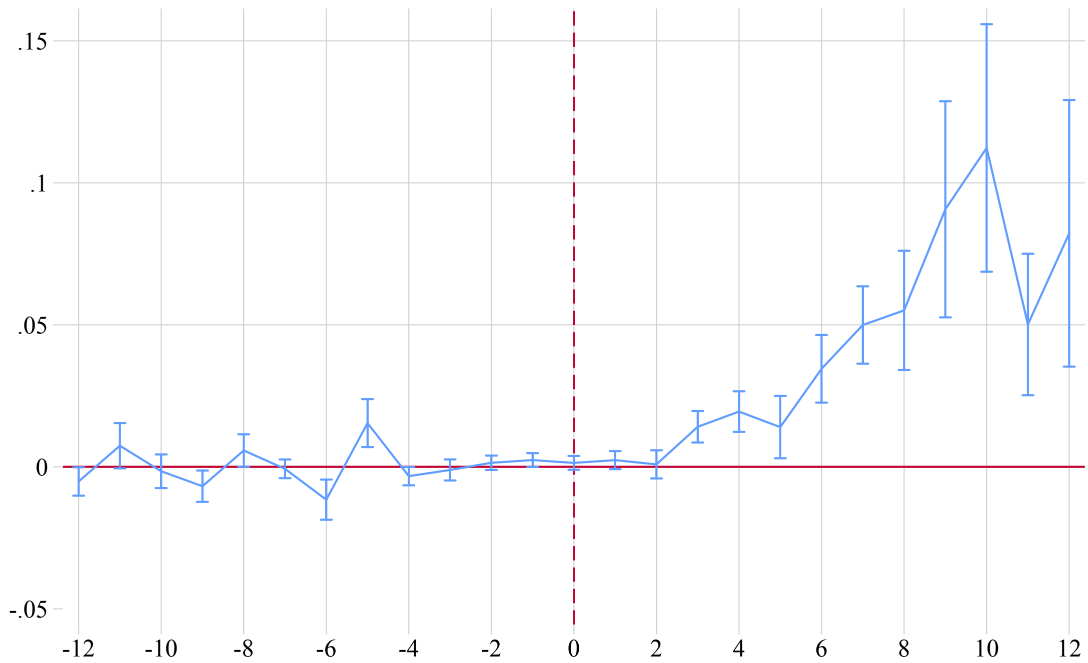
B.6 Different functional forms

We estimate Equation (2) replacing the triadic citation count Y by its inverse hyperbolic sinus $\mathcal{H}(C)$, by its cubic root and by $\log(1+Y)$ (Figure B.5). Each of those transformations assigns slightly different weights to the intensive and extensive margins (Chen and Roth, 2022). The results remain qualitatively the same as in the baseline regression: flat pre-trends, a small initial increase in the outcome variable, and a sharp increase after 2 years, which fades away after a decade.

B.7 Fixed Effects

In the main section, we showed results for the baseline specification where we use fixed effects to account for firm dynamics and destination dynamics by introducing both firm-year and country-technological field-year fixed effects. In Figure B.7, we contrast our baseline results (in blue) with several other plausible specifications. The results remain qualitatively the same: flat pre-trends, an effect that increases over time, then partially fades away after a decade. When controlling for a firm-year-region fixed effect, the effect on citations appears much more

Figure B.4: Excluding countries with affiliates



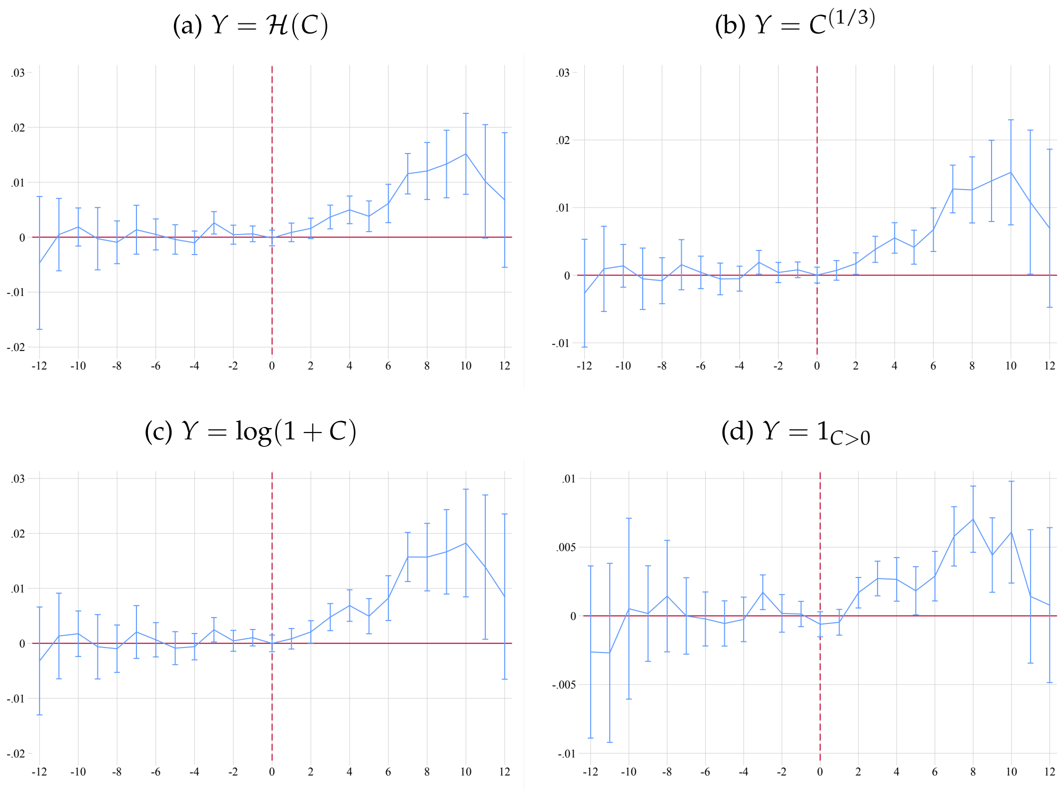
Notes: This figure presents estimates of the coefficient DID_k from Equation (2). The sample excludes the destination-firm pairs for destinations where the firm owns an affiliate. The x-axis represents the value of k, $k = 0$ being the first year the firm exports to country j. 99% error bands, computed with robust standard errors clustered at the destination-sector level, are displayed as blue brackets.

persistent (red curve).

B.8 Affiliate induced spillovers

While it is not the main purpose of this paper, a by-product of our data collection effort is that we can look at what happens to citation linkages after a French parent company opens or acquire an affiliate in a foreign country. Using the same identification strategy as the one we use to isolate the effect of trade on knowledge diffusion, we provide some evidence that opening an affiliate in a destination country also generates an increase in citations to the patents of the parent company. This complements the results from Amiti et al. (2023) that shows that domestic firms that start selling to foreign affiliates experience higher Total Factor Productivity which the authors interpret as evidence of spillovers from Foreign Direct Investment.

Figure B.5: Other LHS functional forms



Notes: Panel B.5a presents the results from estimating Equation (2) after transforming the outcome variable with the Inverse Hyperbolic Sine. In Panel B.5b we use the cubic root of citations, in Panel B.5c the log of 1+ the number of citations and in Panel B.5d a dummy variable equals to 1 when the French firm is cited by the destination country that year. Everything else in unchanged.

B.9 The Channels of knowledge spillovers

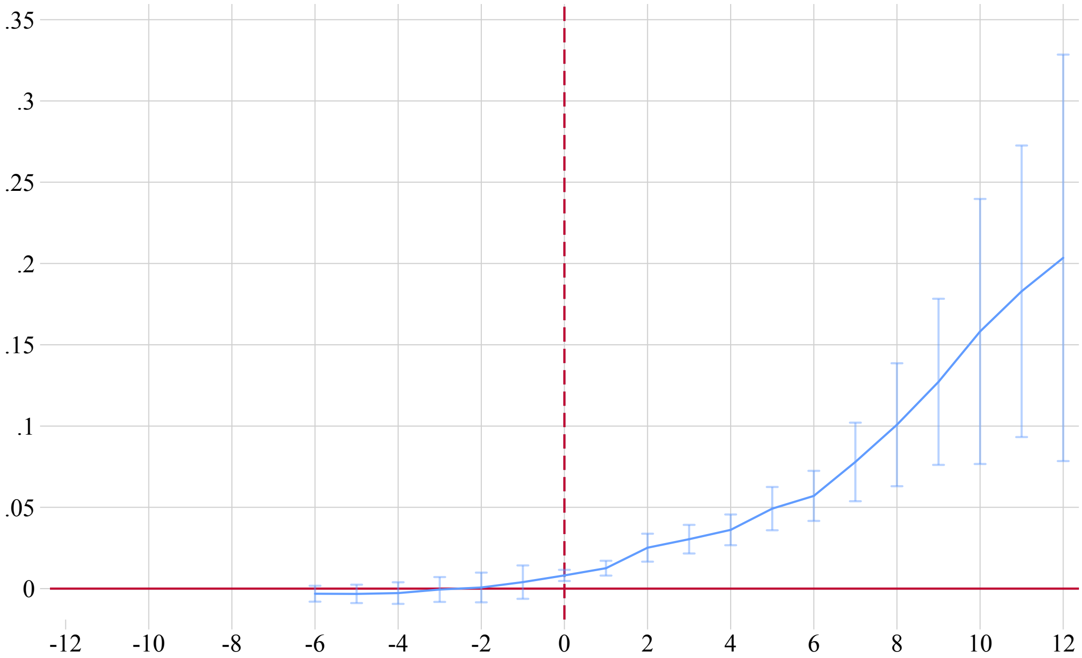
We estimate the following equation to formally test the statistical difference between the two "non-competing" products channels against the "competing" products channel, excluding the group with potential multinational connections:

$$y_{f,j,t}^{\text{non-competing}} - y_{f,j,t}^{\text{competing}} = \sum_{k>0} \beta_k^c 1_{E_{f,j}=t-k} + \chi(f,j,t) + \epsilon_{f,j,t}. \tag{6}$$

In this case, the coefficient β_k^c captures the increased citation rate of each non-competing channel (horizontal and vertical) relative to the competing channel. As in our baseline model, we use the generic control function $\chi(f,j,t)$ for the firm and destination controls: firm-year and destination-technological field-year fixed effects.

The differences between each of the two "non-competing" channels and the "competing" (same product) one are significant in most of the later years, four or

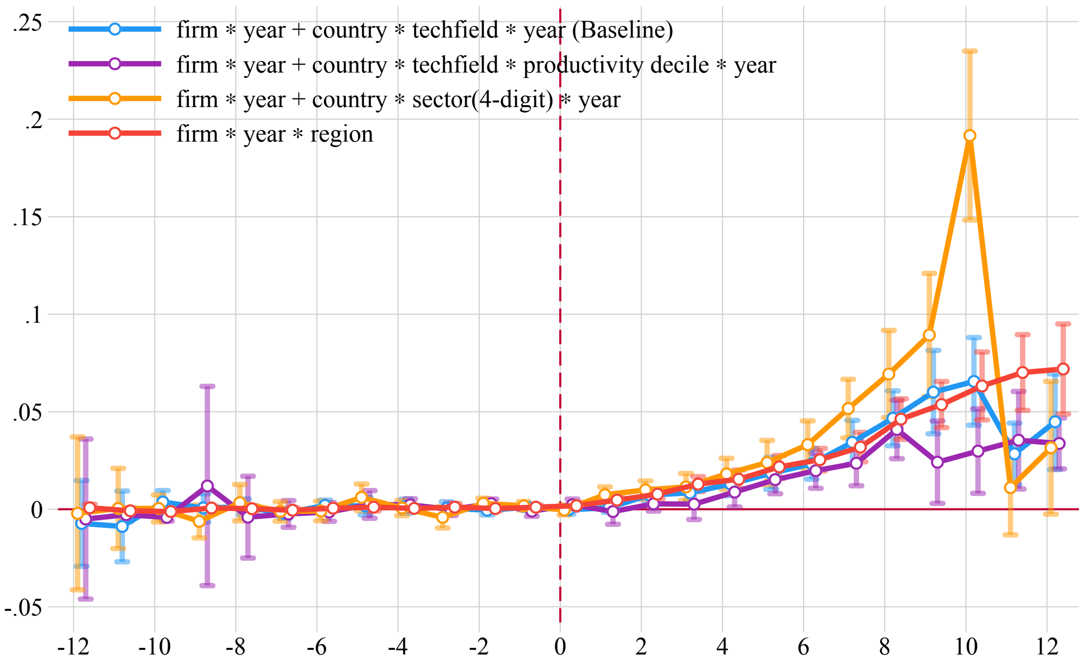
Figure B.6: Baseline estimation of the effect of entry on citations



Notes: This figure presents estimates obtained using the Borusyak et al. (2021) DiD estimator. The x-axis represents the value of k, $k = 0$ being the first year the firm exports to country j. 99% error bands, computed with robust standard errors clustered at the destination-sector level, are displayed as blue brackets.

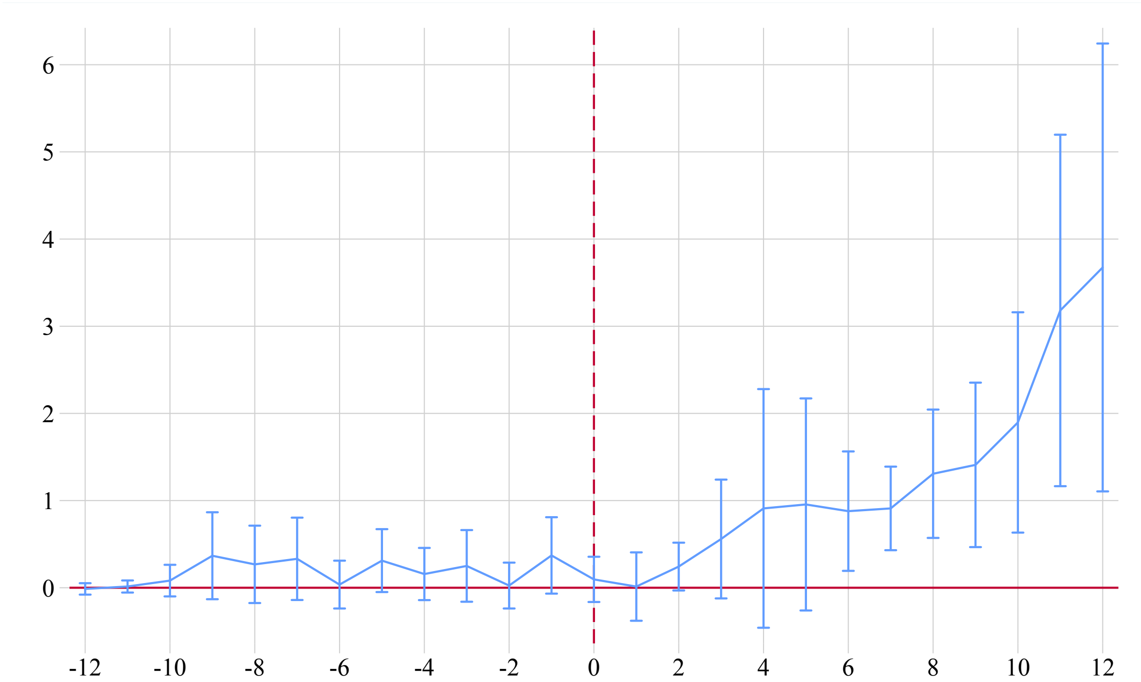
more years after export market entry.

Figure B.7: Controlling for different fixed effects structures



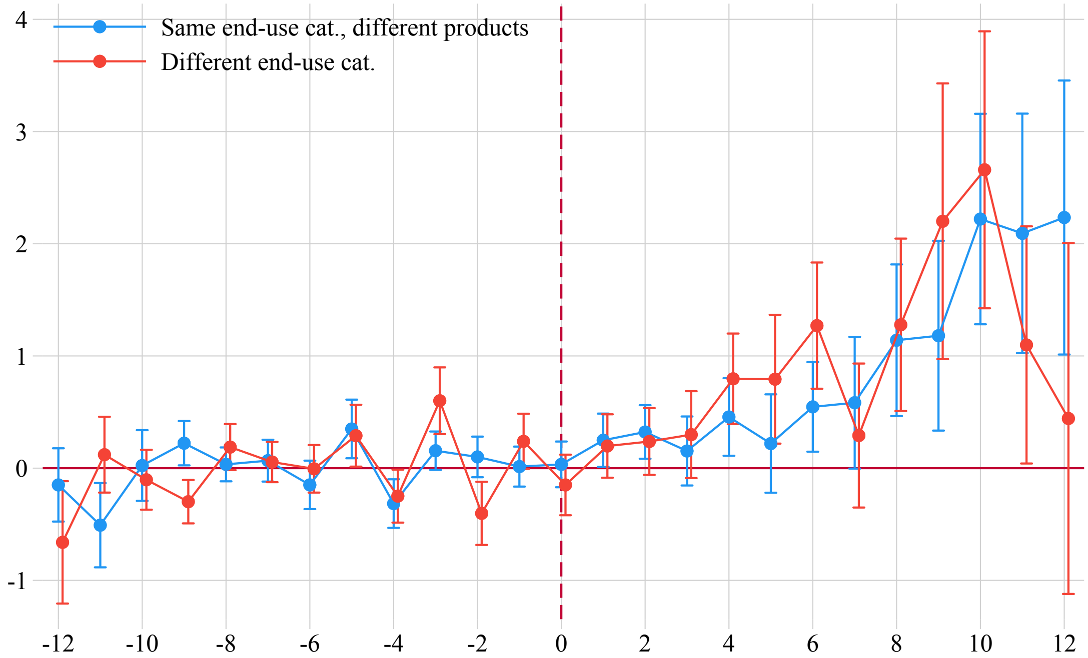
Notes: This figure presents estimates of the coefficients DID_k from Equation (2). The x-axis represents the value of k, $k = 0$ being the first year the firm exports to country j. 99% error bands, computed with robust standard errors clustered at the destination-sector level, are displayed as blue brackets. This specification replaces the set of fixed effects with various combination: firm-year + country-technological field-productivity decile-year in purple; firm-year + country-sector-year in orange ; firm-region-year fixed effects in red. Region groups correspond to wide geographical regions such North America, South America, Western Europe, etc. technological field groups are built based on the technological field that a firm most frequently patents in.

Figure B.8: Effect of opening an affiliate on citations



Notes: This figure presents estimates of the coefficient DID_k from Equation (2). The x-axis represents the value of k, $k = 0$ being the first year the firm declares owning an affiliate in country j. 99% error bands, computed with robust standard errors clustered at the destination-sector level, are displayed as blue brackets. This specification includes firm-year and destination-technological field-year fixed effects as well as a control for the log(1+exports) to the destination country.

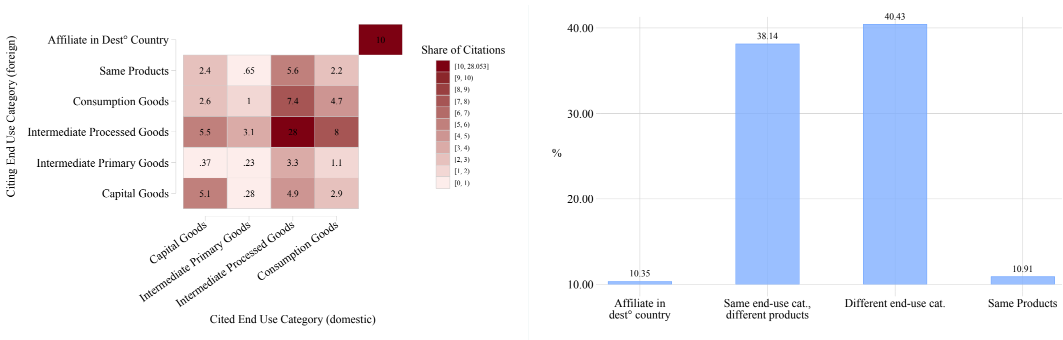
Figure B.9: Knowledge spillovers by channel



Notes: This Figure presents estimates of the coefficient β_k^c from Equation (6) where the dependant variable is the difference between the two categories presented and the reference category - "competing" (same product).

Figure B.10: Innovation Networks

(a) Citations by category of technologies (b) Citations by channel of diffusion

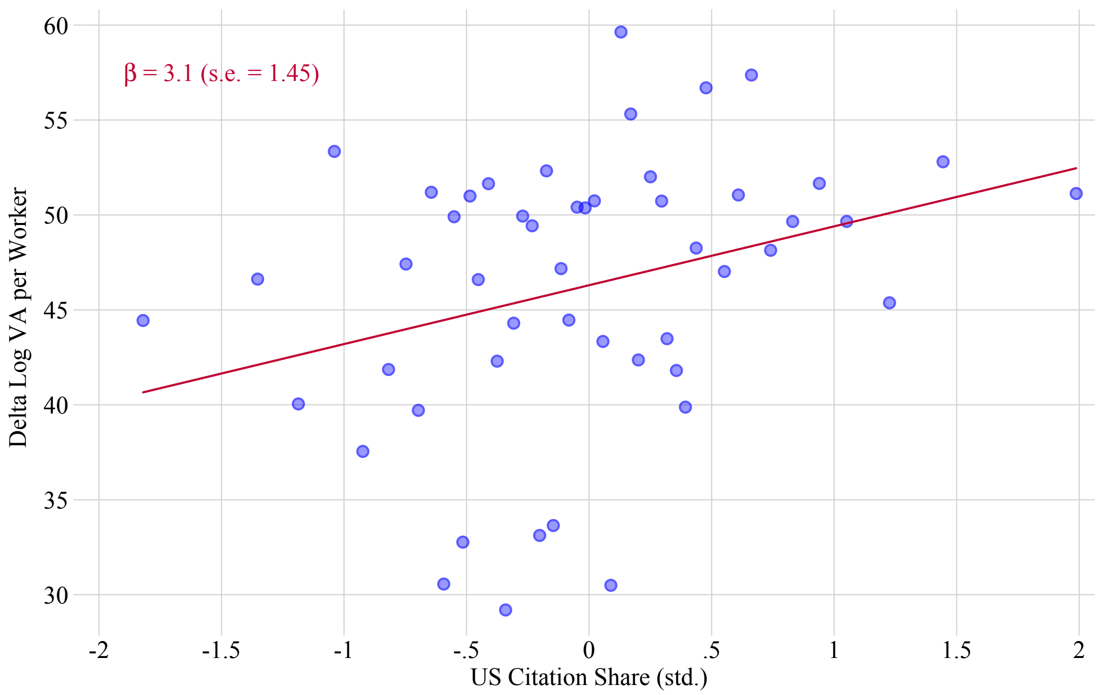


Notes: Panel B.10a presents the distribution of citations by the end-use categories of the associated technologies. Panel B.10b presents the distribution of citations by end-use for more aggregate categories (same products; different products, same end-use categories; different end-use categories).

C Citations from the US and productivity growth

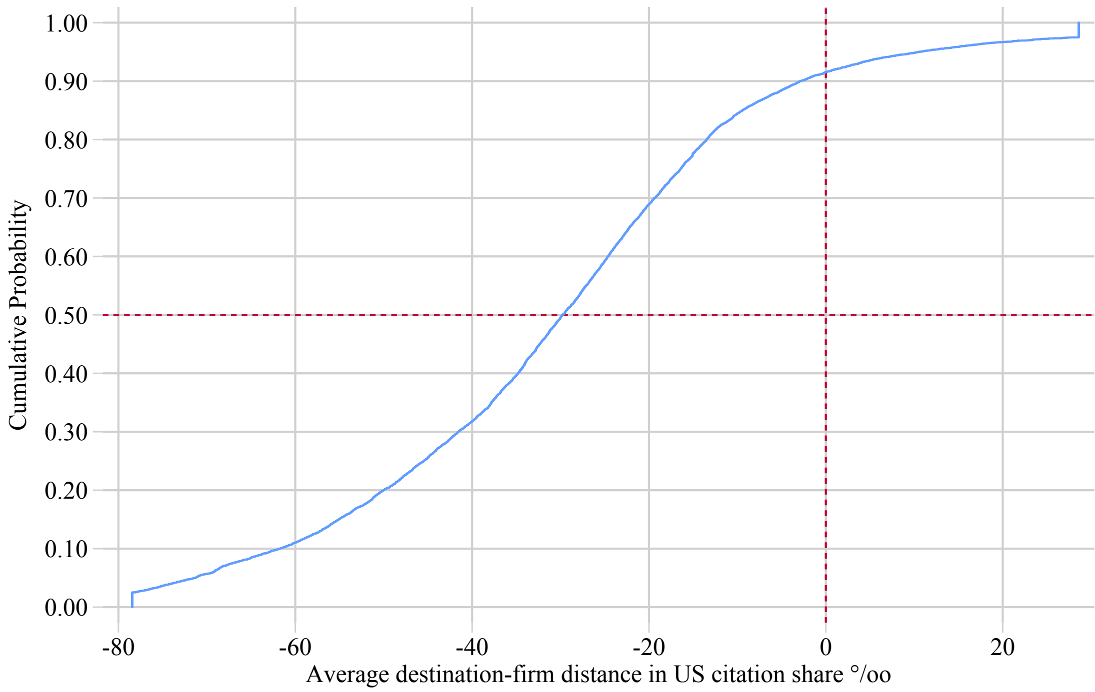
Figure C.1 suggests that a high citation share from the US is a good proxy of the distance to the technological frontier. Firms patenting in fields with a higher share of US citations display higher labor productivity growth. A one standard deviation higher citation rate from the US (e.g. going from the citation rate of the chemical industry to that of the pharmaceutical industry) is associated with a 3.1 log percentage points higher growth rate of value-added per worker.

Figure C.1: Labor productivity growth and citations from the US



Notes: This figure presents a binscatter of the standardized citation share from the US to French firms ($WQUAL_f$ over 2000-2010) versus the observed labor productivity growth (100× log difference of value added per worker) of those firms over 2000-2010.

Figure C.2: Distribution of technological distances



Notes: This figure presents the CDF of the firm-destination measure of technological distance. The measure is winsorized at the 2.5-97.5 level.

D Alternative way of measuring patent quality

D.1 Impact of the exporting firm's patent quality

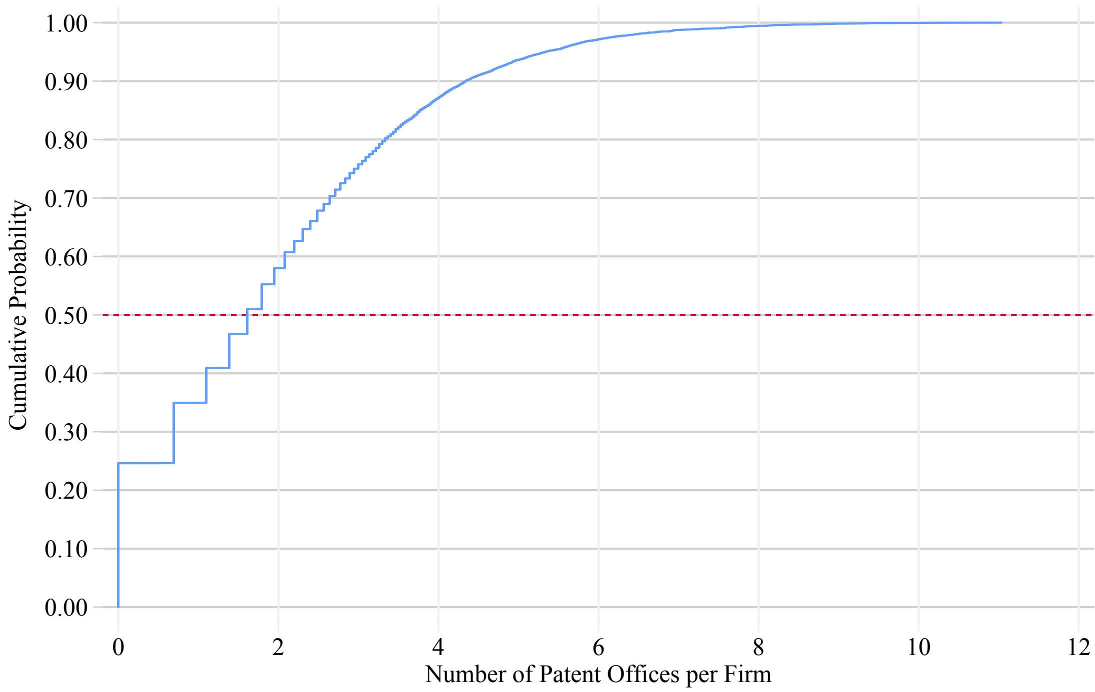
Measuring patent quality

We follow the OECD patent data manual to construct an alternative measure of patent quality for the French exporting firm. We use the number of patent offices the patent was submitted to by the firm. Submitting a patent to a higher number of patent offices amounts to extending the geographical coverage of intellectual property protection for the patent. Doing so is costly for the firm, hence the firm will decide to incur the additional cost of applying to more patent offices only if it anticipates the patent to be sufficiently successful, which we take as reflecting the patent's quality.

Figure D.1 shows the distribution of the number of patent office submissions (over all patents of the firm and in log) for the French exporting firms. We see that 25% of firms submit to only one patent office, 50% of firms submit to less than 6 patent offices. Only 10% of firms submit to more than 66 offices.

We test whether the number of patent office submissions is truly related to more common measures of firm quality. We first describe the relationship be-

Figure D.1: Distribution of the number of patent offices used by firms



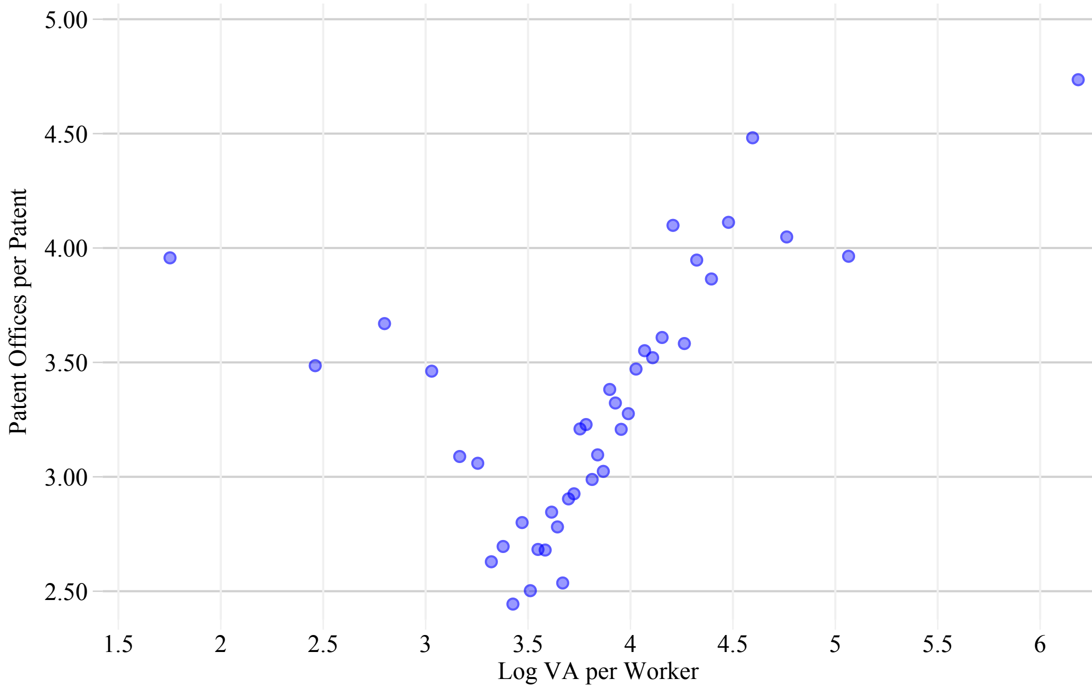
Notes: This figure presents the cumulative distribution of the log of the number of patent offices where firms apply for protection (summed over all patents).

tween the number of patent office submissions and firm level productivity measured by the value added divided by the number of employees. Figure D.2 depicts a binscatter of the log of the number of patent offices against the log of the firm's value added per worker. Each dot corresponds to a bin that contains 2.5 percent of the overall distribution of firm productivities. The figure shows that beyond the 15th percentile of firm-level productivities, patent quality measured by the number of patent offices is clearly positively correlated with labor productivity.

Next, one can look at the extent to which our measure of patent quality relates to the firm's ability and/or propensity to export to destinations that are further away from France. In Figure D.3 we group firms according to the average distance between France and the firms' export markets. Each color represents a different quartile on the scale of export distances. The figure shows a positive correlation between our measure of patent quality and the average export distance of a firm, which is consistent with the view that French firms with higher quality patents are willing to pay a higher trade cost to reach more remote export markets.

To address the concern that our measure of patent quality could reflect firm level characteristics other than patent quality per se, in Figure D.4 we regress the number of patent offices on a set of fixed effects. The figure shows that around 60% of the variance in the number of patent offices is unexplained by aggregate and firm level factors. Aggregate factors such as the year of application, techno-

Figure D.2: Relationship between labor productivity and the firm-level average number of patent offices per patent



Notes: This figure presents a binscatter of the relationship between labor productivity and the average number of patent offices per patent used by firms. Bin size is set at 2.5 percentiles.

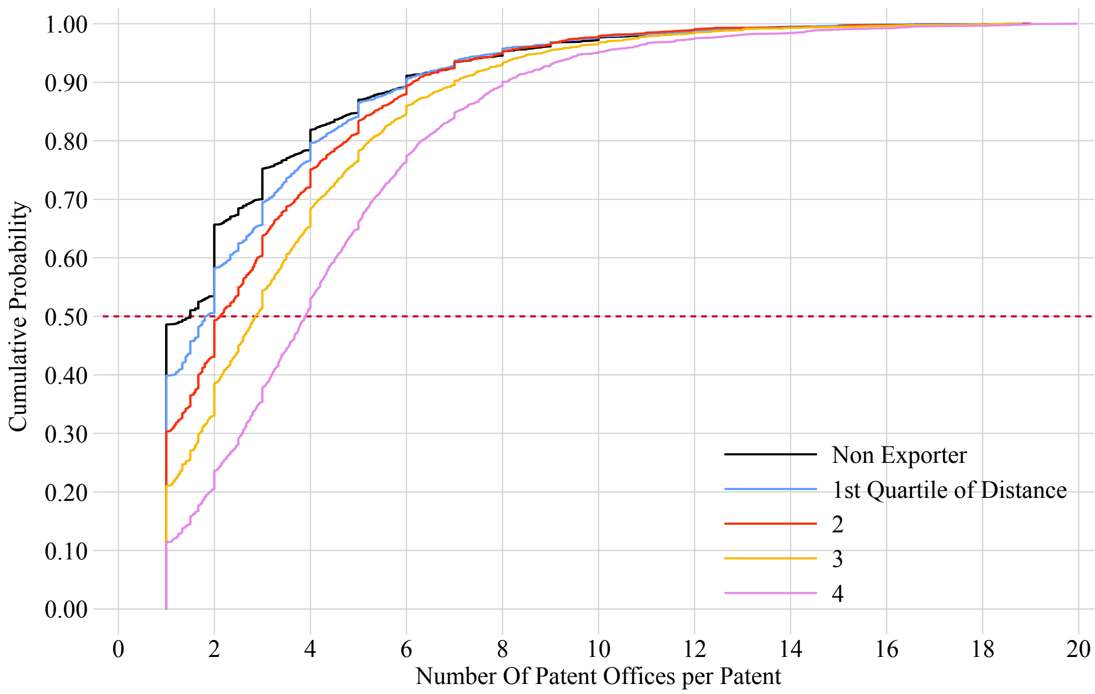
logical fields and the interaction between the two, only explains about 17% of the variance in the number of patent offices. A firm fixed effect only explains about 35% of the variance. The residual (60%), after controlling for all possible combinations of fixed effects, is by construction related to differences across patents within firms.

Figure D.5 shows the relationship between the number of patent offices used for a given patent and its total number of citations five years later. We see that the more patent offices a patent is applied to, the more citations it collects.

Heterogeneous effect of entry with respect to patent quality

Overall, Figures D.2, D.3 and D.4 provide support to using the number of patent offices per patent as a proxy for patent quality. We now look at the impact of a firm's average patent quality on how entry by firm affects subsequent citations to its patents by firms in destination countries. Presumably, firms with higher quality patents are more likely to induce follow-up innovations by firms in the destination countries, and therefore to generate more citations to their patents. To see this, we adapt the baseline regression to allow for varying coefficients with respect to the quality of the exporting firm's patents. More specifically, in Figure

Figure D.3: Relationship between the average distance to their export market and the firm average patent offices per patent



Notes: cumulative distribution of the number of patent offices per patent where firms apply for protection after grouping firms according to the average distance of their export market.

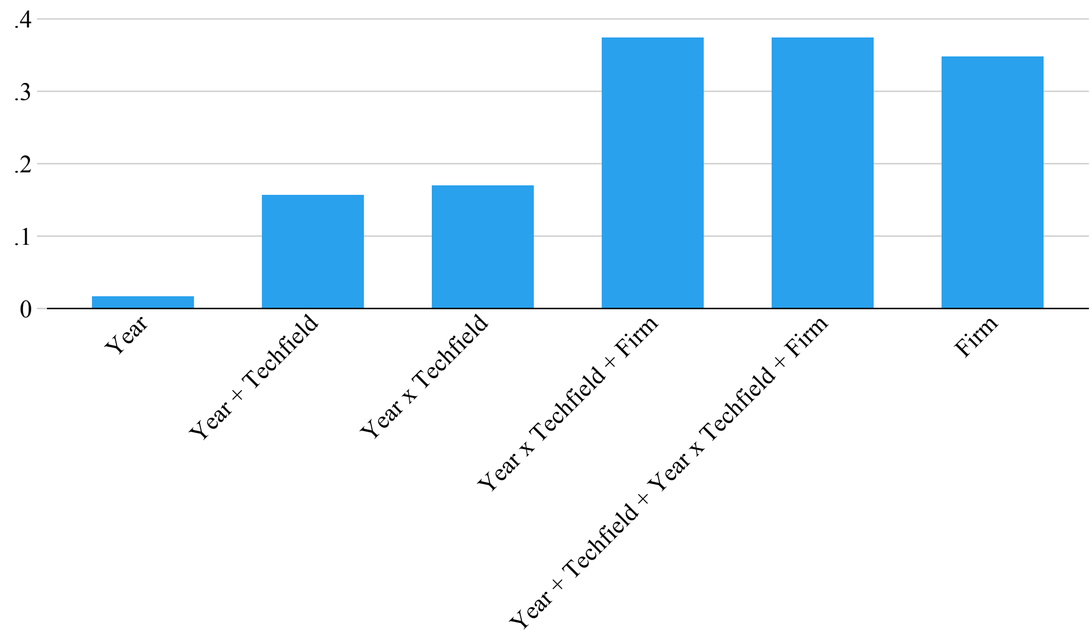
D.6 we combine the de Chaisemartin and d'Haultfoeuille (2020)'s estimator with a local linear estimation and use a kernel re-weighting scheme across levels of average patent quality. This kernel approach allows us to flexibly estimate the functional form of the marginal effect of entry on subsequent patent citations by firms in the destination countries, across level in average firm-level patent quality. The kernel is estimated with a bandwidth of x.

Each point on the blue line in Figure D.6 corresponds to the effect of entry on subsequent citations estimated at a given level of firm's average patent quality. We see that the higher the average quality of a French firm's patents, the stronger the effect of entry by that firm on subsequent citations by firms in the destination countries.

D.2 Knowledge spillover and absorptive capacity

The transfer of knowledge from a French exporter to firms in the export destination is likely to depend upon the destination's technological development relative to the French exporter. If firms in the destination country lag far behind the French firm, then presumably these firms are not adequately equipped to build on the French firm's innovation, and therefore the French firm's entry should

Figure D.4: Variance decomposition of patent quality



Notes: R^2 of regressions where the LHS variable is the number of patent offices where a patent is filed on various sets of fixed effects.

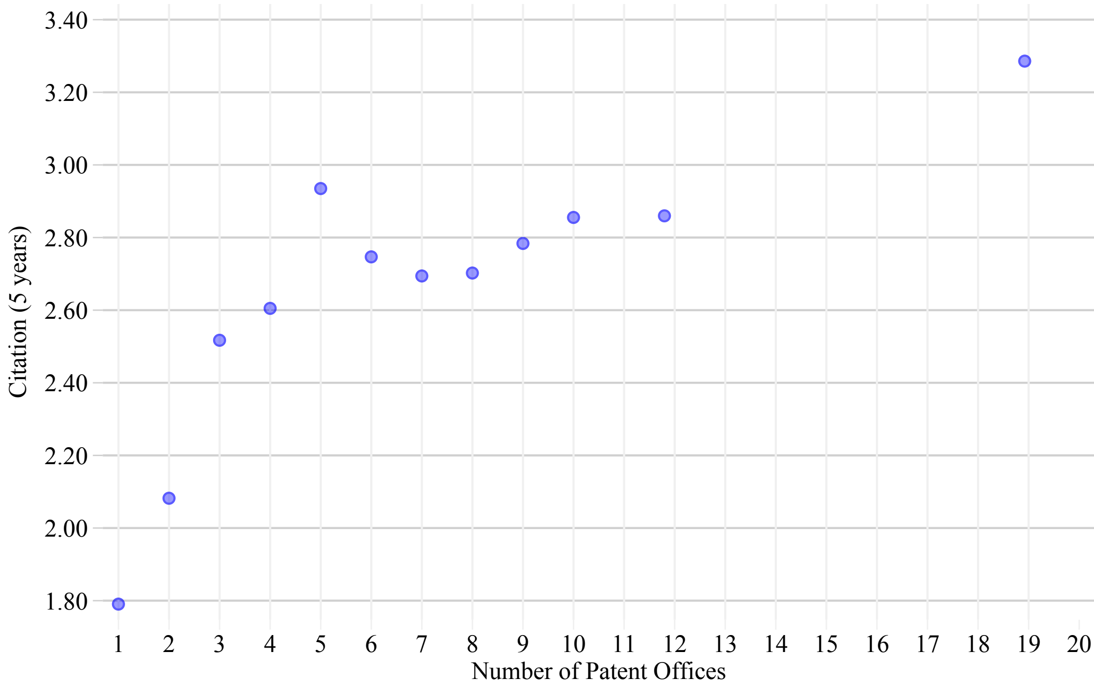
have a limited impact on innovation in the destination country. The French firm might even deter such innovation in the destination country due to the increased competition it induces for potential innovators in that country [13]: as a result, the impact of the French firm's entry on citations by firms in the destination country may even turn negative. On the other hand, if firms in the destination country are neck-and-neck with the French firm, then these firms can easily build upon the French firm's technology to generate new innovations: in that case entry by the French firm should increase citations by the destination country of the firm's innovations. Finally, if firms in the destination country are far ahead of the French firm's technology, then these firms will often not find it useful to develop further the French innovation as they already enjoy a better technology: entry by the French firm would then have little to no impact on its citations by firms in the destination country.[14]

To test for a differential impact of entry on citations varying with a destination's development level, we run a similar version of our baseline specification

[13]See Aghion et al., 2005. See also Van Patten (2021) for a framework that incorporates heterogeneous learning abilities among importers

[14]All these developments should have different consequences for the destination firms' products as well, but the lack of data on those products prevents us from assessing such impacts. They also bring about different consequences for the French exporter's products, which we plan to investigate in future work.

Figure D.5: Patent Citations and Number of Patent offices



Notes: Binscatter of the number of patent offices where each patent was applied to and the number of citations each patent got 5 years after the application date.

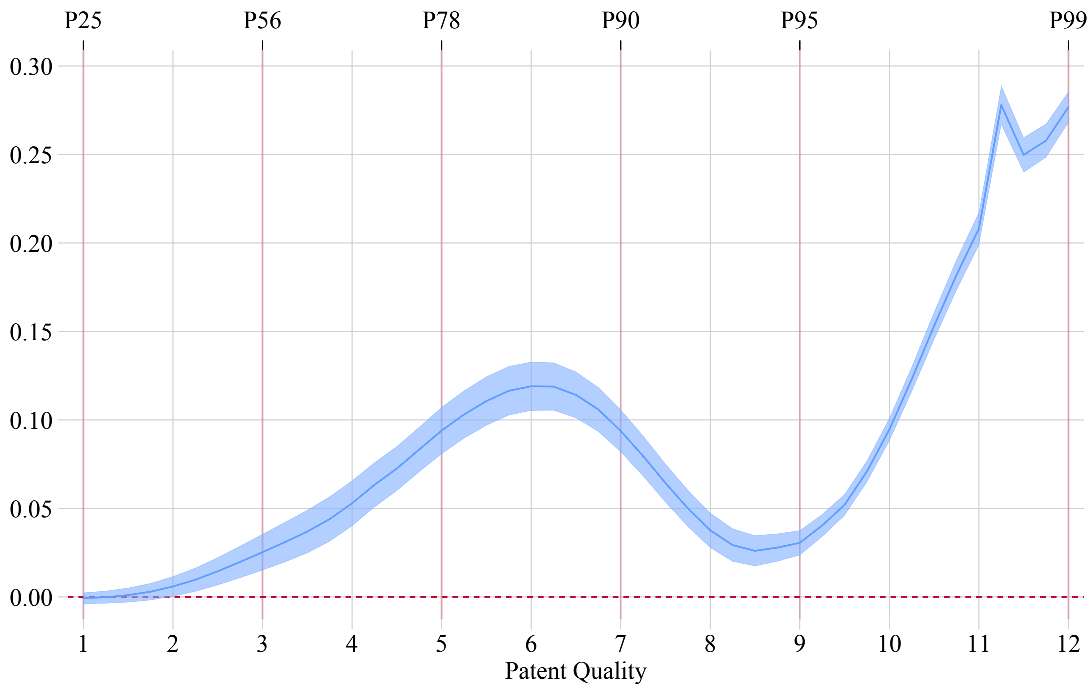
described above. But we now allow for our coefficient of interest to vary across the distribution of the destinations' distance to the firm in terms of patent quality. When the destination is far behind the firm, entry generates few citations (Figure D.7). When the destination and the firm are closely matched, entry increases citation the most. The effect decreases as technological distance increases, though it remains significantly positive.

D.3 Effect of entry on patent quality

Next, we look at the effect of the French firm's average patent quality on the quality of subsequent citations to the firm's patents by firms in the destination countries. For that purpose, we build a sample of all citations by firms in destination countries, not just the triadic citations used in our analysis so far. To measure the quality of a citing patent, we again use the number of patent offices the patent is submitted to: for example a citing patent submitted to one patent office counts for 1, whereas a citing patent submitted to three patent offices counts for 3. This gives us a quality weighted measure of citations.

We first estimate our baseline model with firm-year and destination-technological field-year fixed effects, but with quality-weighted citing patents as left-hand side variable. Figure D.8 shows that this quality-adjusted measure of subsequent cita-

Figure D.6: Relationship between patent quality and the effect of entry on citations



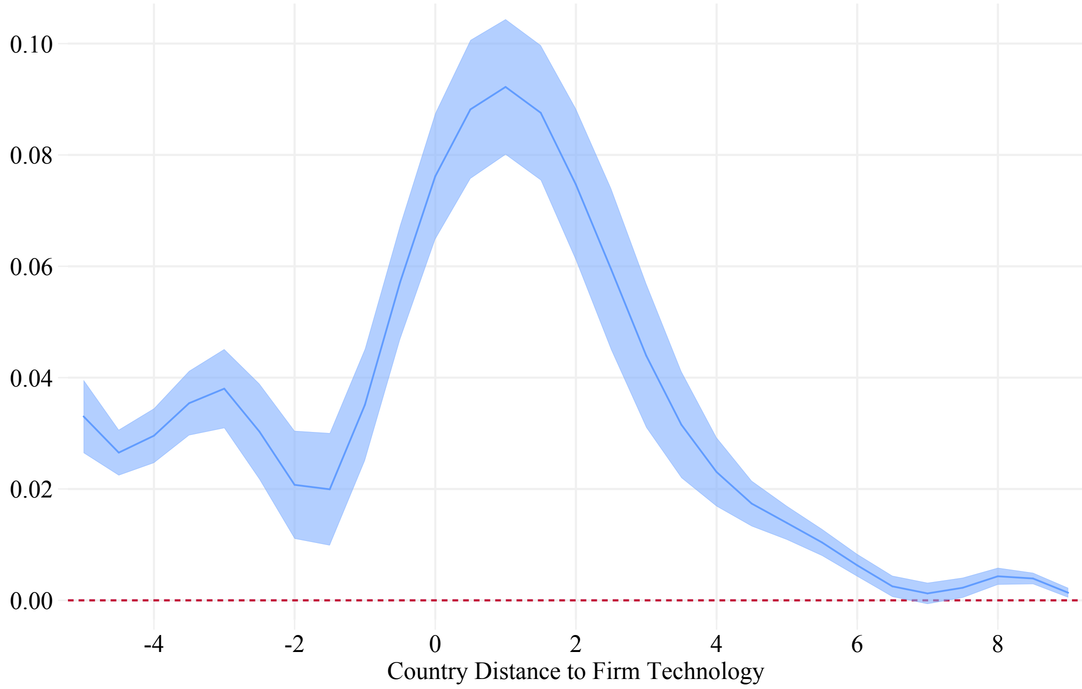
Notes: This figure presents estimates of the coefficient DID_k from Equation (2). The x-axis represents the value of the moderating variable. The coefficients are estimated locally along the support of the distribution of the firm level proxy for patent quality. 99% error bands, computed with robust standard errors clustered at the destination-sector level, are displayed as blue brackets.

tions evolves over time pretty much as the raw triadic citations to a French firm's patent in our baseline regression, namely: (i) the pre-trends are flat and precisely estimated close to zero; (ii) the increase in quality-weighted citations is moderate during the first few years after entry by the French firm into the new foreign market; (iii) the quality-adjusted citations then increase sharply until year 10 after entry.

We further decompose the impact of the French firm entry on the quality-adjusted citations received from the destination countries into a "quantity component", the number of citations by firms in destination countries, and a "quality component", the average number of patent offices per citation in destination countries (Figure D.9). Most of the increase in quality-adjusted citations after en entry comes from the quantity of citations. The quality contribution is positive but barely significant. The magnitudes of the effects uncovered are very large: after ten years, the citation rate is multiplied by 5.7, and given an increase in quality of 4.3%, the quality weighted citations are multiplied by 5.8.

Overall, the analysis in this subsection has several interesting implications. First it helps rule out that subsequent innovations post-entry would be purely

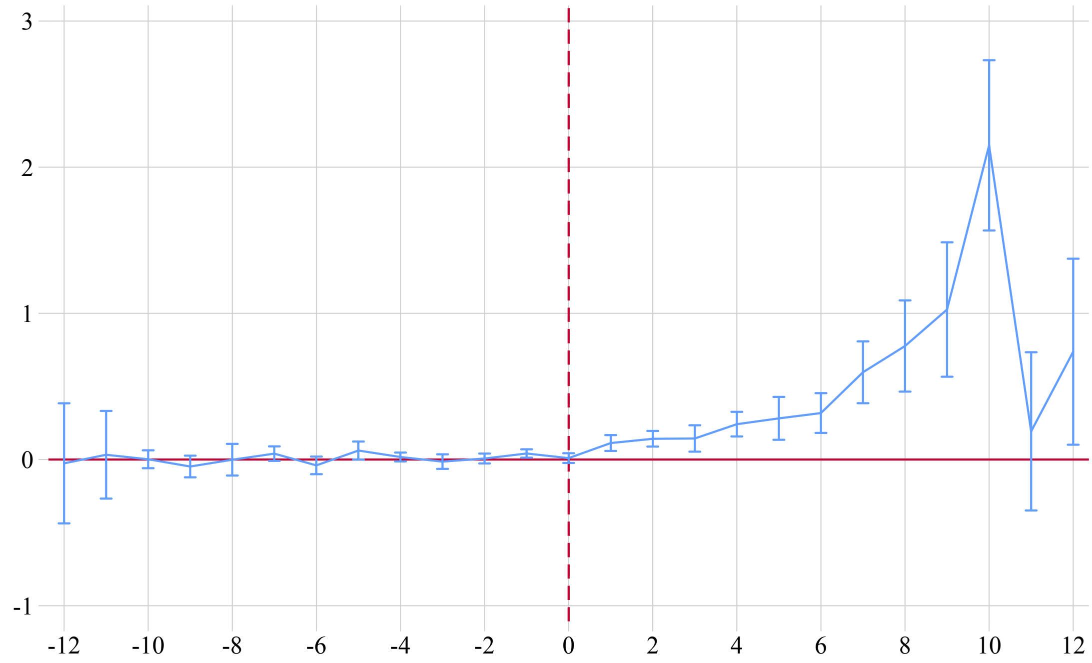
Figure D.7: Relationship between technological distance and the effect of entry on citations



Notes: This figure plots the effect of the initial entry into a foreign destination estimated locally at a given point of the ex-ante distribution of the firm-destination technological distance. The dependant variable is the number of citations. We use Gaussian weights with a bandwidth set to 0.5. 99% confidence interval are presented. Standard errors are clustered at the sector-country level.

defensive or marginal improvements with low potential: if anything the quality of the citations increase with entry. Second, the higher effect of entry on citations for firms with higher quality technology has important implication in terms of "gains from trade": not only do higher quality firms export to more destinations, but they also generate higher spillovers on inventors in destination countries. Additionally, such spillovers translate into both a higher number of citing patents and a higher quality of citing patents by firms in destination countries.

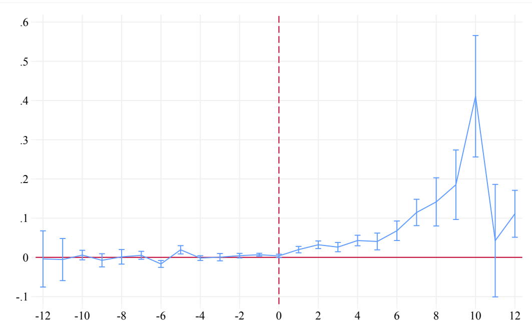
Figure D.8: Effect of Entry on quality weighted citations



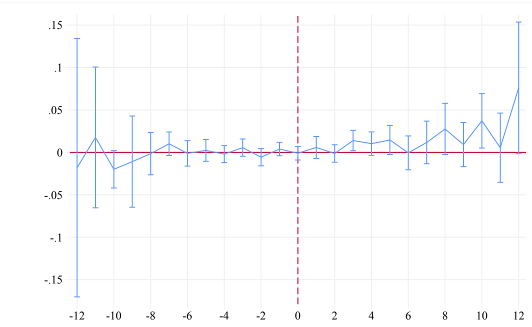
Notes: This figure presents estimates of the coefficient DID_k from Equation (2). The x-axis represents the value of k, $k = 0$ being the first year the firm exports to country j. 99% error bands, computed with robust standard errors clustered at the destination-sector level, are displayed as blue brackets.

Figure D.9: Quantity versus Quality

(a) Quantity



(b) Quality



Notes: Panel D.9a presents the results from estimating Equation 2 with the count of citations as outcome variable. Panel D.9b presents the results from the same estimation but with the citation quality proxy as outcome variable.

CENTRE FOR ECONOMIC PERFORMANCE
Recent Discussion Papers

1959	Gabriel M. Ahlfeldt Nathaniel Baum-Snow Remi Jedwab	The skyscraper revolution: Global economic development and land savings
1958	David Hémous Simon Lepot Thomas Sampson Julian Schärer	Trade, innovation and optimal patent protection
1957	Maria Cotofan Konstantinos Matakos	Adapting or compounding? The effects of recurring labour shocks on stated and revealed preferences for redistribution
1956	Javad Shamsi	Understanding multi-layered sanctions: A firm-level analysis
1955	Bridget Kauma Giordano Mion	Regional productivity differences in the UK and France: From the micro to the macro
1954	Tim Obermeier	Individual welfare analysis: A tale of consumption, time use and preference heterogeneity
1953	Kirill Borusyak Xavier Jaravel	The distributional effects of trade: Theory and evidence from the United States
1952	Ariela Caglio Sebastien Laffitte Donato Masciandaro Gianmarco Ottaviano	Has financial fair play changed European football?
1951	Eugenie Dugoua Todd D. Gerarden	Induced innovation, inventors and the energy transition
1950	Christian Krekel Johannes Rode Alexander Roth	Do wind turbines have adverse health impacts?

1949	Réka Juhász Claudia Steinwender	Industrial policy and the great divergence
1948	Virginia Minni	Making the invisible hand visible: Managers and the allocation of workers to jobs
1947	Eugenie Dugoua	Induced innovation and international environmental agreements: Evidence from the ozone regime
1946	Marco Bertoni Gabriel Heller-Sahlgren Olmo Silva	Free to improve? The impact of free school attendance in England
1945	Matthias Mertens Bernardo Mottironi	Do larger firms exert more market power? Markups and markdowns along the size of distribution
1944	Ria Ivandić Anne Sophie Lassen	Gender gaps from labor market shocks
1943	Ihsaan Bassier Alan Manning Barbara Petrongolo	Vacancy duration and wages
1942	Cheng Keat Tang Stephen Gibbons	Are friends electric? Valuing the social costs of power lines using house prices
1941	Luis Bauluz Pawel Bukowski Mark Fransham Annie Seong Lee Margarita López Forero Filip Novokmet Sébastien Breau Neil Lee Clément Malgouyres Moritz Schularick Gregory Verdugo	Spatial wage inequality in North America and Western Europe: Changes between and within local labour markets 1975-2019

The Centre for Economic Performance Publications Unit
Tel: +44 (0)20 7955 7673 Email info@cep.lse.ac.uk
Website: http://cep.lse.ac.uk Twitter: @CEP_LSE