INTERNATIONAL MONETARY FUND

Knowledge Diffusion Through FDI: Worldwide Firm-Level Evidence

JaeBin Ahn, Chan Kim, Nan Li, and Andrea Manera

WP/24/152

2024
JUL

WORKING PAPER

 WP/24/152

IMF Working Paper
Research Department & Western Hemisphere Department

Knowledge Diffusion Through FDI:
Worldwide Firm-Level Evidence*

Prepared by JaeBin Ahn, Chan Kim, Nan Li, and Andrea Manera

Authorized for distribution by Petia Topalova and Koshy Mathai
July 2024

ABSTRACT: This paper examines the impact of Foreign Direct Investment (FDI) on knowledge diffusion by analyzing the effect of firm-level FDI activities on cross-border patent citations. We construct a novel firm-level panel dataset that combines worldwide utility patent and citations data with project-level greenfield FDI and cross-border mergers and acquisitions (M&A) data over the past two decades, covering firms across 60 countries. Applying a new local projection difference-indifferences methodology, our analysis reveals that FDI significantly enhances knowledge flows both from and to the investing firms. Citation flows between investing firms and host countries increase by up to around 10.6% to 13% in five years after the initial investment. These effects are stronger when host countries have higher innovation capacities or are technologically more similar to the investing firm. We also uncover knowledge spillovers beyond targeted firms and industries in host countries, which are particularly more pronounced for sectors closely connected in the technology space.

RECOMMENDED CITATION: Ahn, JaeBin, Chan Kim, Nan Li, and Andrea Manera, 2024, “Knowledge Diffusion Through FDI: Worldwide Firm-Level Evidence”. IMF Working Paper 24/152.

JEL Classification Numbers:	F21, F23, F63, O31, O33
Keywords:	Greenfield FDI; Brownfield FDI; cross-border M&A; Inward FDI; Outward FDI; Knowledge spillover; Patent citation; LP-DiD
Author’s E-Mail Address:	jahn@imf.org, chankim@umd.edu, nli@imf.org, amanera@imf.org

* We would like to thank Fuad Hasanov, Wolfgang Keller, Chris Papageorgiou, Petia Topalova and seminar participants at the IMF for helpful comments and discussions. The research is partly supported by the Macroeconomic Policy in Low-Income Countries program of the UK’s Foreign, Commonwealth and Development Office. The views expressed here are those of the authors and do not necessarily represent the views of the IMF, its Executive Board, IMF management, or the supporting partner.

WORKING PAPERS

Knowledge Diffusion Through FDI: Worldwide Firm-Level Evidence

Prepared by JaeBin Ahn, Chan Kim, Nan Li, and Andrea Manera

Contents

1 Introduction

Over recent decades, multinational corporations (MNCs) have played a leading role in a globalized world, accounting for around 90% of total exports and imports (Bernard et al., 2009) and, inherently, the totality of Foreign Direct Investment (FDI) flows. Over the same period, the world has witnessed a surge in cross-border patent citations (e.g., IMF, 2018; LaBelle et al., 2023), often used as a direct measure for knowledge flows. These concurrent trends have prompted a growing body of literature to propose that knowledge diffusion through FDI plays a pivotal role in promoting growth and facilitating technological catch-up across countries (Keller, 2021). However, robust causal evidence on this channel remains limited, mainly confined to specific contexts or instances. Meanwhile, quantifying the extent of FDI-induced knowledge spillovers across diverse countries and contexts has become increasingly relevant, particularly given the current rise in geopolitical tensions that threaten to disrupt cross-border investment flows (IMF, 2023; Gopinath et al., 2024).

Our paper empirically examines the extent to which both inward and outward FDI influence global knowledge diffusion. We construct a novel dataset that combines the universe of project-level data on greenfield investment and cross-border mergers and acquisitions (brownfield FDI) from fDi Markets and Refinitiv Eikon, respectively, with worldwide patent citation data from PATSTAT, through a careful name-matching procedure. This process yields a database that includes transactions from 12,656 firms and bi-directional patent citations involving 60 countries over the period 2003-2022.[1] Specifically, we study how citations to and from a specific firm evolve around its initial FDI in a new country. Applying the novel local projection difference-in-differences (LP-DiD) methodology developed by Dube et al. (2023), we find a significant increase in citations by the host country towards the investing firm—by around 7.8% to 10.6%, depending on the investment type—after the firm's first investment. In addition, the investing firms also increase their citations towards the host country by around 4.5% to 13%, suggesting knowledge flows from the host to investing firms also increase. These baseline results are robust to a range of alternative specifications and definitions.

Our extensive dataset and advanced methodological approach contribute to the literature of FDI and knowledge diffusion in three dimensions. First, we enhance causal identification of FDI's impact on knowledge spillovers by introducing the LP-DiD methodology to this topic for the first time. This method effectively addresses the issue of heterogeneous dynamic treatment effects that often complicate staggered difference-in-differences (DiD) analyses. In staggered DiD settings, treatment—such as FDI events—occurs at different times for different projects, potentially yielding

[1] As explained below, these countries are the top 60 countries by the number of granted patents in PATSTAT, collectively representing 99.9% of the granted patents in this dataset.

differential effects over time. Traditional DiD methods fall short in capturing these varying impacts, introducing bias to the results. The LP-DiD approach overcomes this by utilizing local projections to estimate the dynamic effects of an event, while employing a "clean control" condition to define appropriate treated and control groups. Our global firm-level dataset is particularly well suited for the LP-DiD application, as it allows us to account for time-varying factors at both the host-country(-industry) and firm levels that could influence citation trends. By incorporating an extensive set of fixed effects, we successfully control for variations in innovation rates, citation propensities, patent activities, and other unobserved shocks across countries, industries, and firms. This approach ensures our estimates more accurately reflect the causal effects of FDI on knowledge spillovers.

Second, by leveraging both greenfield and brownfield investment data at the same time, we can investigate previously under-explored aspects of knowledge spillovers—namely, the differential impacts of greenfield versus brownfield FDI through potentially distinct channels (e.g., Antràs and Yeaple 2014). Until now, empirical studies could hardly distinguish different types of FDI, while most of the theoretical studies on FDI spillovers have focused on greenfield FDI, neglecting brownfield FDI which may serve alternative objectives such as competitive reduction (e.g., Neary 2007; Cunningham et al. 2021). As such, we separately estimate the extent of knowledge spillovers from greenfield and brownfield FDI, representing one of the few attempts to compare spillovers from greenfield brownfield FDI in a unified empirical framework.

Third, the detailed granularity and extensive global coverage of our dataset provide insights into the heterogeneous impacts of FDI on knowledge diffusion across a diverse range of countries and industries. This includes assessing absorptive capacity along the technological dimension as well as exploring cross-industry spillovers. Moreover, it enables examination of bi-directional knowledge spillovers, allowing us to analyze not only the knowledge diffusion that host countries benefit from foreign investments but also the reciprocal knowledge gains experienced by the investing firms themselves.

Our main findings are summarized as follows. First, as highlighted above, after a firm's initial entry into a country, there is a significant and roughly equal increase in citations between the host country and the investment firm, suggesting a notable uptake in bi-directional knowledge exchange. Second, regarding types of FDI, patent citations tend to increase more following greenfield compared to brownfield investment, albeit the difference is modest. Third, knowledge spillovers extend beyond directly targeted firms and industries, affecting technologically related sectors, which sometimes benefit more—percentage-wise—than the sectors directly targeted by FDI. Finally, consistent with the literature on the role of "absorptive capacities", our analysis reveals substantial heterogeneity in the degree of knowledge spillovers across countries. Interactions with host countries that have larger pre-existing patent stocks result in spillovers that are two to ten times greater. Moreover, technological similarity between investing firms and host countries facilitates knowledge spillovers, with notable increases in citation effects when their technologies are alike. In fact, when there is low technological similarity, knowledge spillovers appear to be insignificant.

Related Literature This paper contributes to several strands of the literature. The conventional wisdom that inward FDI benefits a host country stems from theoretical literature. In addition to the job-creation and capital-formation effects that directly come with inward FDI, several sources of externalities have been postulated including the technology spillover effect, spillovers through backward and forward linkages, and the pro-competitive effect.[2] Subsequent empirical studies explored various kinds of datasets to identify the presence of positive FDI spillovers and the specific channels through which FDI spillovers take place, mostly using changes in productivity to indirectly infer knowledge spillovers.[3] Our paper directly traces knowledge flows using citations between patents *à la* Jaffe et al. (1993), which not only circumvents measurement concerns on productivity (Keller and Yeaple, 2009) but also helps uncover underlying mechanisms through which FDI improves productivity.[4] Furthermore, examining both patent citations made by the host country to the investing firm and reverse citations made by the investing firm to the hosting country enables an analysis of the direction of knowledge flows.[5]

In a sense, our study is closely related to prior works that associate FDI with cross-border patent citations with a particular focus on specific country cases. For example, Branstetter (2006) studies Japanese firms' innovative activity and their investment in the U.S., and finds that FDI

[2]Formal description of each channel has been proposed in Rodriguez-Clare (1996) for backward and forward linkages, Glass and Saggi (1998) for the technology spillover effect, and Navaretti and Venables (2004) for the pro-competitive effect.

[3]Seminal works in this line of research include, among others, Haskel et al. (2007) and Keller and Yeaple (2009) who report positive intra-industry spillover effects in the United Kingdom and United States, respectively. By contrast, Aitken and Harrison (1999) find a negative intra-industry spillover effect in Venezuela, which is attributed to the market-stealing effect caused by entering foreign firms. As for inter-industry spillover effects, most studies found positive effects, particularly for backward linkages (Harrison and Rodríguez-Clare 2010). For example, Javorcik (2004) explores Lithuanian firm-level data and find supporting evidence of the presence of backward linkages in both countries: positive productivity spillovers from FDI take place mostly through contacts between foreign firms and their local suppliers in upstream sectors. Blalock and Gertler (2008) confirm technological spillovers from FDI via backward linkages among Indonesian firms. Jiang et al. (2018) find both backward and intra-industry spillover effects from international joint ventures in China. A more comprehensive review of the recent literature is provided in Harrison and Rodríguez-Clare (2010) and Keller (2021). Alternatively, Javorcik et al. (2018) and Deng et al. (2024) report product innovation among Turkish and Chinese firms in upstream sectors, respectively, as evidence for FDI spillovers.

[4]Beyond spillovers from FDI, patent citation measure has been also increasingly used in the trade literature to assess trade channels of knowledge spillovers (e.g., Aghion et al., 2023;Maurseth and Verspagen, 2002; Jinji et al., 2015). Akcigit and Melitz (2022) provide theoretical discussions on trade-induced knowledge spillovers.)

[5]This is particularly relevant because, unlike spillovers in host countries, studies of knowledge diffusion to source countries are limited. Few exceptions providing country-specific evidence include Navaretti and Venables (2004), Griffith et al. (2006), Goldbach et al. (2019), Ni et al. (2021).

increases patent citations both from and to the Japanese investing firms. Globerman et al. (2000) show that Swedish outward FDI is a channel for the diffusion of foreign technology to home-country MNCs and SMEs. Most recently, Akcigit et al. (2024) find that the foreign corporations that invest in U.S. startups increase their own citations to those U.S. startups post-investment, suggesting benefits in the form of knowledge spillovers.[6] Unlike previous studies that focus on specific host country, source country, or country pairs, we expand the scope of the analysis to cover multiple host and source countries at the same, thereby exploiting rich variation across countries.

One conclusion that is broadly supported in the literature is the uneven effect of inward FDI across countries. This variation likely reflects levels of absorptive capacity, as domestic firms may not benefit from the positive externatlities provided by foreign-owned firms if the technology gap is too large (Keller, 1996). Indeed, Borensztein et al. (1998) show that countries with higher levels of human capital can benefit more from a given level of inward FDI compared to countries with lower levels of human capital. This suggests that FDI contributes to economic growth only when the host economy has sufficient absorptive capacity to utilize the advanced technologies. Similarly, Alfaro et al. (2004) find that countries with more developed financial system can exploit inward FDI more efficiently. Our study, in contrast, specifically explores the extent of absorptive capacity along the technological dimension.

While numerous studies explore the motivations behind greenfield investment compared to brownfield investment (Dikova and Brouthers, 2016), the literature provides scant evidence regarding how knowledge diffusion may differ between these two types of investment. Firms seeking to exploit proprietary technology may prefer greenfield investment, because it may enhance the prospect of maintaining firm-specific advantages (Chen and Zeng, 2004). On the other hand, while mergers and acquisitions (M&A) are often driven by the direct access to innovation or synergies from combining innovation capabilities (Bena and Li, 2014; Phillips and Zhdanov, 2013), there is also evidence that M&A can serve purposes unrelated to innovation, such as preempting future competition (Cunningham et al., 2021) or securing tax advantages (Belz et al., 2013). As such, it remains uncertain whether these differing motivations lead to distinct knowledge diffusion outcomes, let alone the channels through which they might influence knowledge spillovers.

The rest of the paper is structured as follows. In Section 2, we describe our data sources, the name-matching procedure to assemble the data, and stylized facts that highlight the positive correlation between FDI and citation flows. Section 3 describes the application of the LP-DiD to our context, its suitability and advantages relative to other methods and all our choices of treatment variables, outcomes and empirical specification. Section 4 reports our results. After describing the

[6]In contrast, Chen et al. (2022) report FDI affects innovation in the host country mostly due to an increase in competition, rather than by knowledge spillovers, using matched firm-level patent data of Chinese firms.

baseline effects of brownfield and greenfield FDI on citations, we move to industry-level results and later analyze the heterogeneity of our findings depending on countries' absorptive capacity. We close this section with a discussion of numerous robustness exercises. Section 5 concludes.

2 Data Sources and Descriptions

2.1 FDI data

Greenfield FDI We obtain project-level greenfield FDI data from fDi Markets, a service provided by fDi Intelligence, a part of the Financial Times Group. This dataset tracks announcements of new physical projects or expansion of existing investment projects which create jobs and capital investment. It serves as a primary source for global greenfield FDI reported in the World Investment Report by UNCTAD. The data are collected primarily from publicly available sources (e.g., media sources, industry organizations, investment promotion agencies news wires) and cover investment-level information for over 300,000 FDI deals between January 2003 and December 2022. For each project, the database provides information on the parent company name, the source and destination countries, the industrial sector the project belongs to, the activity type (e.g., business services, sales, R&D), investment category (new investment or expansion), as well as value of investments and the estimated number of jobs created.[7]

Brownfield FDI We also employ transaction-level brownfield FDI data from Refinitiv Eikon, formerly known as SDC Platinum by Thomson Reuters. The database provides detailed information on cross-border M&A transactions that represent the acquisition of at least a 5% stake or of a 3% stake with a deal value of at least USD$ 1 million, covering more than 1.45 million deals in the world since the 1970s. This dataset is also the primary source for cross-border M&As patterns reported in World Investment Report by UNCTAD, and has been studied extensively in recent academic research such as Erel et al. (2022) and Bergant et al. (2023).[8] For each cross-border M&A deal, the database offers information on the acquirer and target firm names and locations, the sector associated with the target firm, and its purchase value (in USD).

[7] The reliability of these data has been confirmed with official statistics by, for instance, aggregating FDI values at the destination country-year level and comparing them with gross FDI inflows data from official sources (e.g., Toews and Vézina, 2022; Aiyar et al., 2023).

[8] According to Bollaert and Delanghe (2015), most empirical papers analyzing M&As and published in top journals use this database to construct their sample of observations: for instance, in the top four finance journals from 2000 to 2012, more than 75% of papers employed this database as sole data source or in addition to other sources.

2.2 Citation Data

We obtain citation data from the Spring 2022 release of PATSTAT, maintained by the European Patent Office (EPO). This dataset contains the most comprehensive bibliographic details on patents granted by or applied to 90 patent-issuing authorities, which the EPO compiles using data provided by individual patent including non-European countries and developing economies. We consider only citations between firms, meaning companies as opposed to patents owned by individual inventors. We determine the country associated with each patent based on the address reported by the firm to the patent authorities at the time of application. In what follows, we provide details on the features of this dataset and describe our method for assigning patents to countries.

First, we shall describe how the EPO assigns identifiers to companies. Patent-issuing authorities transmit patent applications to the EPO, which then extracts information pertaining to the applicants (both inventors and owners) based on the data that these applicants self-report to the patent office. The EPO then analyzes the applicant names to assign unique identifiers, consolidating various versions of a company's name that are assessed to refer to the same entity (EPO IDs). This procedure, common in patent offices, is called "name disambiguation", and helps correct issues like misspelling of company names. Through this procedure, the EPO also labels applicants to indicate if they are a "company". For the purposes of our analysis, we focus on applicants classified as companies, excluding those who are inventors.[9]

The disambiguation procedure identifies which self-reported applicant names correspond to the same company, but it sometimes leaves us with multiple self-reported addresses associated to a single EPO ID, and these addresses are sometimes in different countries.[10] In these cases, we attribute the country that is most frequently associated with each given EPO ID. In some instances, we do not have self-reported addresses associated with EPO ID's. For these firms, we attribute the country corresponding to the patent-granting authority in which the EPO ID occurs most frequently.

We also map International Patent Classification (IPC) into International Standard Industrial Classification (ISIC) via a probabilistic crosswalk (Lybbert and Zolas, 2014). In addition, in some analysis we use triadic patents—patents registered in the US Patent Office (USPTO), Japanese Patent Office, and EPO—as high-quality indication (de Rassenfosse et al., 2014).

We then count citations between each firm and several destination countries. As common in the literature, we measure citations at the level of a "patent family" in order to avoid issues of double

[9]Inventors can only be physical persons, so if an applicant is not an inventor, it means that they are listed as applicants only to indicate their ownership of the patent. Unlike the "company" designation, patent authorities communicate to the EPO whether applicants are inventors or not.

[10]We think that these different country designation may be due to, e.g., local subsidiaries of a company reporting their address in the country rather than their headquarters.

counting.[11] Additionally,to measure knowledge spillover as accurately as possible, we only include citations that were made by the patent applicants, ensuring our analysis reflects the knowledge base of inventors at the time of invention. This excludes any citations that patent examiners may have added at a later stage, as these may reflect technological relevance but not direct knowledge spillover.

When counting patents and citations, we employ fractional counting, dividing each patent family across the different sectors and applicants proportionally. In particular, for a patent applicable to N IPC codes and belonging to M companies, we assign $1/(NM)$ patents to each firm-IPC pair. We obtain counts at the firm-ISIC code level applying the probabilistic crosswalk of Lybbert and Zolas (2014) to firm-IPC counts. When working at higher levels of aggregations, we sum the relevant fractional counts. For example, the country-level count of a patent registered in multiple countries will be given by the sum of each country's assignees over the total number of assignees.

2.3 Linking Citations and FDI Information

To analyze citations related to specific firms, we need to align names from FDI datasets with those in PATSTAT, facing a challenge due to over 107,000 and 541,000 unique names in each dataset, respectively. To manage this, we standardize names following established procedures described by Arora et al. (2021b), and then restrict the PATSTAT sample to include only firms with five or more patents, significantly reducing the unique name count to 100,000.

We then apply string similarity metrics, specifically cosine similarity with TF-IDF (term-frequency-inverse document frequency) weighting at the tri-gram character level and a matching algorithm developed by the Dutch central bank to identify potential matches between the datasets.[12] Lastly, among firm-name pairs with a similarity score over the threshold of 0.7 (out of 1) from either metric, we filter out mismatches by manually checking the firm names. We refer to these matches as "World sample", signifying their comprehensive coverage of corporate entities globally across the datasets.

While global name matching offers the advantage of wide coverage, it may raise concerns about accuracy since it relies solely on string similarities among names. For instance, some patents initially

[11] A family collects patent applications referring to the same invention, which can appear at different times in different patent offices. Counting patent families is therefore the preferred approach to measure inventions in the patent literature (see, e.g., de Rassenfosse et al., 2014).

[12] The main advantage of using cosine similarity is that it allows for faster computation when comparing a large number of names, by utilizing matrix operations instead of pairwise comparisons. The tri-gram character approach is based on the intuition that less frequent tri-grams, such as 'hyu' (as an example of the first three characters (tri-gram characters) of a company 'Hyundai'), will be assigned higher weights by TF-IDF compared to common words like 'com' from a word like 'company'. As a result, a pair of words that commonly contain 'hyu' will receive a higher similarity score compared to a pair of words that commonly contain 'com'. For more information on the string matching Python package developed by the Dutch central bank, please refer to https://github.com/DeNederlandscheBank/name_matching.

invented by a firm might be assigned to its subsidiaries or parent companies, and a change in name can lead to mismatches. To address this concern, we created a separate dataset specifically for U.S. investing firms. This allows us to focus on a smaller sample size and use well-established matches from previous studies. This subset builds on the firm name matches between PATSTAT and U.S. Compustat firms developed by Arora et al. (2021b). By assigning patents to their ultimate owners and accommodating changes in ownership, covering patents from 1980 to 2015, this crosswalk provides better ownership data for U.S. firms' patents.[13]

Next, we match this set of U.S. firms with FDI events from fDi Markets and Refinitiv. In addition to the two string similarity metrics mentioned earlier, we also use online search similarity to match names, following Autor et al. (2020). We further refine the matches by ensuring that each pair of firm names has at least one common webpage among the first ten results from Bing's search API. As a result, the dataset achieves higher accuracy for both patent and FDI information for U.S. firms. We refer to this as "US sample."

Using these matched datasets, we compute firm-level citations to and from host countries by year. Firm *citations* from country c in year t are the sum of fractional citations received by all patents belonging to a firm f in that year. Conversely, *reverse citations* are the sum of fractional citations made by all (granted) patents applied in year t by the investing firm to patents in country c. These two measures are proxies for knowledge flows from investing firms toward destination countries and vice versa. In our specification, we focus on the *stock* of citations as our primary outcome variable because flow citations are sporadic, which would result in noisier estimates. Moreover, cumulative citations provide clearer interpretation of overall knowledge spillovers over time after FDI occurs.

Due to computational constraints, our analysis is limited to the top 60 destinations by the number of granted patents from 2003-2022, representing over 99.9% of all patents. This selection is important for managing the dataset size, especially for analysis at the firm-country-sector level, which otherwise would exceed 180 million observations.[14] We therefore focus on this subset throughout the analysis. Overall, our worldwide sample include citations and FDI information for 12,656 firms globally, covering 12,696 brownfield FDI projects and 4,632 acquiring firms, as well as 87,415 greenfield FDI projects by 10,096 investing firms. The merged U.S. sample covers 1,872

[13] Arora et al. (2021b) made two main improvements in their datasets compared to previous matches. First, they account for the true owner of the patents. In many cases, the assignee of a patent listed in PATSTAT does not align with the ultimate owner of that patent, as the assignee may be a subsidiary or a firm controlled by the owner. The crosswalk developed by Arora et al. (2021b) addresses these issues and provides us with a set of PATSTAT patents matched to their ultimate owner for the period 1980 to 2015. Second, their dataset reassigns patents over time when firms undergo changes in their ownership structure.

[14] This number of observations comes from multiplying 1,889 assignees by 60 destinations, 80 ISIC 2-digit sectors, and 20 years.

firms, among which 1,143 firms engaged in a total of 2,850 M&A deals overseas, and 1,438 firms invested in a total of 16,433 greenfield investment projects.

2.4 Stylized Facts

Time-series evolution of FDI and patent citations To illustrate overall patterns of FDI and patent citations over time, we group the counties into advanced and developing countries, and trace investment and citation flows between and across groups, leading to four separate series in each panel in Figure I: from advanced to advanced countries (ADV⇒ADV); from advanced to developing countries (ADV⇒DEV); from developing to advanced countries (DEV⇒ADV); from developing to developing countries (DEV⇒DEV). They are all normalized to ADV⇒ADV flow value in 2003.

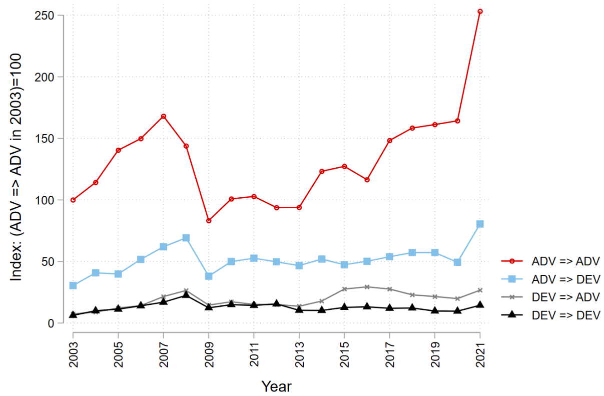
The top panel (Ia) shows that brownfield FDI between advanced countries dominates, followed by brownfield FDI from advanced to developing countries. Worldwide brownfield FDI had been increasing rapidly up until the global financial crisis when it plummeted. Since then, it has been slowly recovering to reach the pre-crisis peak only recently. On the other hand, the middle panel (Ib) reveals that greenfield FDI from advanced countries to developing countries outpaced greenfield between advanced countries until around the global financial crisis after which the former stagnated while the latter continued growing until the pandemic hits. For both brownfield and greenfield FDI, investment flows originating from developing countries tend to account for only a minor share of the total FDI. Likewise, the bottom panel (Ic) describes the evolution pattern of cross-country patent citations over the past two decades, which is predominantly driven by citations between advanced countries. It was not until a decade ago that developing countries began to cite patents belonging to advanced countries increasingly at a notable level.

Relationship between FDI and patent citations We exploit the firm-destination country level matched data to check if there is any relationship between FDI and patent citations after purging firm-year, destination country-year, and bilateral country-pair specific factors that could influence FDI and citation decisions independently and simultaneously. Specifically, we residualize FDI and patent citation measures using the Poisson pseudo-maximum likelihood (PPML) estimator *à la* Silva and Tenreyro (2006):

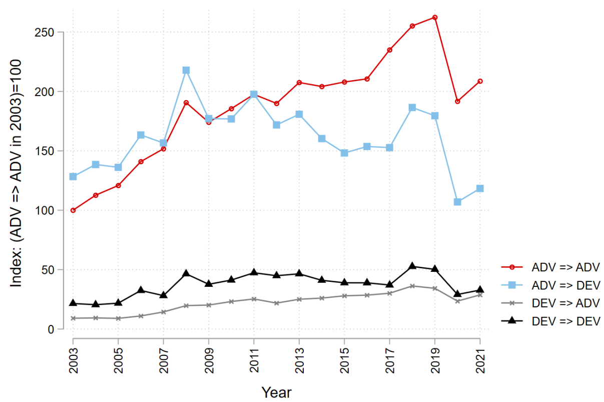
$$Y_{ioct} = \exp\left[FE_{it} + FE_{ct} + FE_{oc}\right] \times \varepsilon_{ioct}, \tag{1}$$

where Y_{ioct} is either a total number of FDI transactions by investor firm i from a source country o to a destination country c in year t for brownfield or greenfield FDI measures, or a total number of patent citation counts made from country c to firm i in country o in year t (i.e., knowledge flow

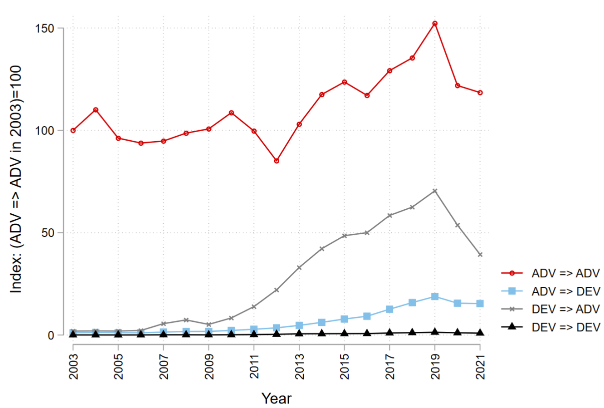
Figure I: FDI and Patent Citations over Time by Country-pair Income Levels



(a) Brownfield FDI



(b) Greenfield FDI



(c) Patent citations

Note: This figure plots the time-series evolution pattern of brownfield FDI (Ia), greenfield FDI (Ib), and patent citations (Ic), respectively, all of which are grouped by host country and source country income levels: from advanced to advanced countries (ADV⇒ADV); from advanced to developing countries (ADV⇒DEV); from developing to advanced countries (DEV⇒ADV); from developing to developing countries (DEV⇒DEV). Each of the measures is expressed in index value where ADV⇒ADV flow value in 2003 corresponds to 100.

from firm i to c) for citation measures. Explanatory variables are fixed effects terms: FE_{it} denotes firm-year fixed effects that can control for the fact that more innovative firms tend to invest abroad more; FE_{ct} is destination country-year fixed effects that can capture the fact that bigger countries tend to innovate more and receive FDI more; FE_{oc} is country-pair fixed effects that should absorb bilateral country-level variables such as geographical or cultural distance to take into account that a country pair between which more FDI activity takes place also tend to make and receive patent citations more intensively from each other.

Figure II plots binned scatters of residualized FDI and patent citations to illustrate their relationship at the firm-destination country level. The top row measures patent citations as a total number of citations made by new patents from a destination country to patents belonging to an investor firm. The bottom row measures patent citation as a total number of citations made by new patents by an investor firm to patents belonging to a destination country. The left and right column considers brownfield FDI and greenfield FDI, respectively.

For both brownfield and greenfield FDI, we find that an investor firm's patent tends to be cited more frequently from a country where the firm invested more, and, at the same time, the firm is more likely to cite patents belonging to a country with greater investment. This strongly suggests that such a positive correlation manifests the presence of underlying forces between FDI and patent citations above and beyond firm-, country-, or country pair-specific confounding effects, highly motivating us to identify the direction of causality between FDI and citations.

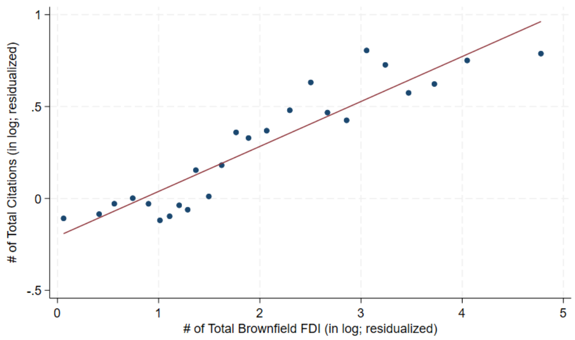
3 Empirical Strategy

Our empirical approach applies the local projection difference-in-differences (LP-DiD) framework for event studies recently proposed by Dube et al. (2023). This method combines local projections event studies with a careful selection of the sample to ensure that both treatment and control groups are unaffected by delayed treatment effects from previous events.

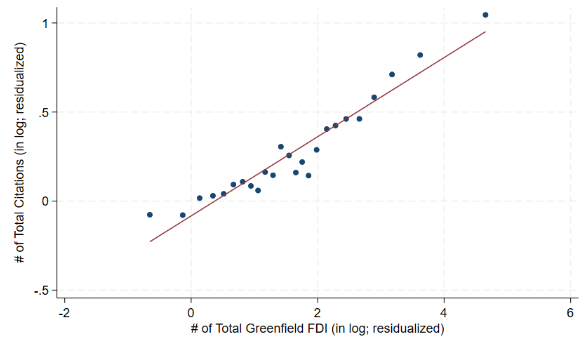
The LP-DiD methodology offers several advantages over traditional DiD and LP analyses. First, unlike DiD, it accommodates control for pre-treatment variables. A particularly significant case is the possibility to control for differences in observed trends in the dependent variable before the event, which is not possible in a standard DiD setting.[15] Compared to both LP and DiD specifications, the LP-DiD methodology avoids the bias arising from heterogeneous treatment effects across different treatment groups through appropriate selection of treatment and control groups. This bias is particularly relevant for our study, since it manifests most starkly in settings where

[15] A standard DiD cannot handle pre-treatment trends in dependent variables when units are treated at different times. Indeed, since a single regression is run, the only option is to control for trends in specific time periods.

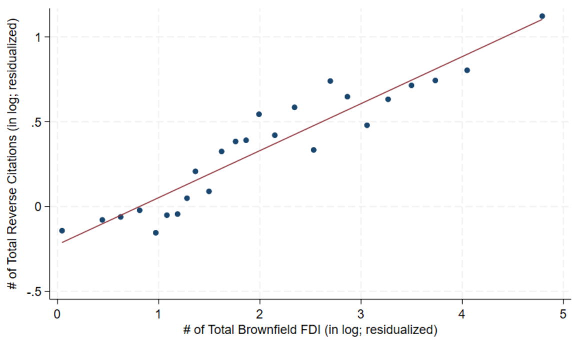
Figure II: Binned Scatters: Firm and Destination Country-level FDI and Citation



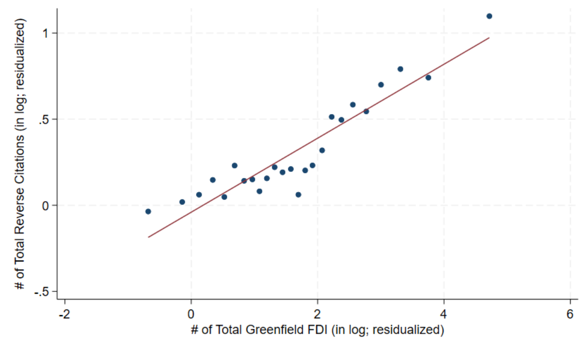
(a) Brownfield (Citations received)



(b) Greenfield (Citations received)



(c) Brownfield (Citations made)



(d) Greenfield (Citations made)

Note: This figure plots binned scatter plots for the relationship between FDI and citation at the firm and destination country level. The top row measures citation as a total number of citations made by a destination country to patents belonging to a given firm. The bottom row measures citation as a total number of citations made by a given firm to patents belonging to a destination country. The left and right column considers brownfield FDI (cross-country M&A) and greenfield FDI, respectively. Both the horizontal and vertical axes are residualized against source-destination country pair, destination country-year, and firm-year fixed effects by PPML estimator.

treatments are staggered and repeated (de Chaisemartin and D'Haultfœuille, 2020; Goodman-Bacon, 2021).

We adopt the LP-DiD approach precisely because multiple acquisitions or greenfield investments can target the same countries and sectors at various times. Our baseline specification examines how initial FDI by firm i in country c affects the citations flowing between the patents of country c and the patents of firm i, denoted as y_{ic}, h period after the event:

$$y_{ic,t+h} - y_{ic,t-1} = \beta_h D_{ic,t} + \sum_{k=1}^{p} \gamma_k^h y_{ic,t-k} + \eta' \mathbf{x}_{ic,t-1} + \delta_{it}^h + \delta_{ct}^h + \varepsilon_{ic,t}^h, \tag{2}$$

restricting the sample to observations that are either:

$$\begin{cases} \text{investment episodes} & D_{ic,t} = 1,\ D_{ic,t-j} = 0,\ 1 \leq j \leq L, \\ \text{or clean controls} & D_{ic,t-j} = 0,\ -h \leq j \leq L. \end{cases} \tag{3}$$

Here, $D_{ic,t}$ is a dummy indicating that firm i invested (or acquired another firm) in country c in year t for the first time over the period 2003-2022; δ_{it}^h and δ_{ct}^h denote firm-year and destination-year fixed effects to control for firm- and country-specific trends or shocks that could affect citation counts independent of the FDI event;and $\mathbf{x}_{ic,t-1}$ represents other controls which will be explaind below. In this regression, β_h identifies the cumulative impact of FDI on citations y_{ic} measured h years after the event.

Unlike ordinary LP analyses, the LP-DiD approach restricts the sample according to Equation (3) to ensure that neither the treatment group (firm-country pairs with FDI relationship) nor the control group have been "treated"—by firm i's first investment in country c—during the L periods preceding the FDI event under examination. This distinction is crucial as it avoids the inclusion of any firm-country pairs in the control group that experienced first-time FDI in the period between $t-1$ and $t-L$, thereby eliminating the risk of confounding the results with delayed effects from previous FDI activities. Similarly, we exclude any observations that were treated in the period from $t-L$ to $t+h$ to prevent contamination of the control group with delayed treatment effects. For example, suppose that FDI have a positive effect on citations that increases over time for L periods and the control group includes countries that have received FDI, e.g., at time $t-(L-1)$. These countries in the control would still be experiencing a positive treatment effect, causing an upward bias in the control group average citations, and a downward bias in β_h, as delayed treatment effects will effectively be subtracted from estimated coefficients (de Chaisemartin and D'Haultfœuille, 2020; Goodman-Bacon, 2021).[16] The parameter L represents the number of periods required for the effect of FDI to stabilize, after controlling for control variables. For this reason, we refer to L as the number of "stabilization lags" to signal that treatment effects stabilize at a constant level after L periods.

Like in any DiD analysis, the causal interpretation of our results hinges on the parallel trends assumption. This assumption posits that, in the absence of the treatment, the average outcome for the treated and control groups would have followed a similar trajectory over time, conditional on the control variables. Therefore, any divergence in outcomes after the treatment would be attributed

[16] Imposing the condition that controls have not been treated between $t-1$ and $t+h$ is not a key requisite of the specification. In facts, it potentially overly restricts the control group. In a previous version of the paper, we did not apply this restriction and obtained quantitatively similar results.

to the effect of the treatment itself, rather than pre-existing trends. In our context, this assumption would be compromised if trends in citations across different firm-country pairs diverge for reasons not accounted for by our empirical model. Our choice of controls and unit of analysis specifically aims at mitigating this concern about the parallel trends hypothesis. Indeed, firm-time fixed effects are included to control for the fact that firms engaging in FDI are inherently more productive and innovative, and hence, receiving more citations. Similarly, host country-time fixed effects account for the possibility that these firms may target more innovative countries which tend to generate and receive more citations. In addition, using relatively high levels of aggregation for the FDI destination mitigates endogeneity concerns that might arise if investors selectively conduct FDI in countries (or acquire firms) that can easily innovate on their own technology. Although the risk of endogeneity related to firm selection for FDI might still exist even with high-level aggregation, we believe that its impact would be attenuated when looking at aggregate citations coming from entire host (destination) countries. Moreover, in the following section, we take further steps to lessen this challenge, such as excluding citations from subsidiaries (or acquired firms), or focusing on industries that are not directly affected by the FDI activity.

In general, the presence or absence of parallel trends prior to treatment is an empirical question. Therefore, we report β_h coefficients for a number of periods before treatment to check for clear violations of the parallel trend hypothesis. In the following paragraphs, we discuss the choices and assumptions we maintain throughout the paper.

Treatment: First Entry Observed During 2003-2022. We define the initial entry of a firm into a destination as our treatment event, denoted by $D_{ic,t}$. This choice aligns with our LP-DiD restrictive sample criteria in Equation (3) and follows the methodological suggestion by Dube et al. (2023) for cases where a unit (a specific firm-country) can experience multiple events of treatments (i.e. treatment is "non-absorbing").[17] This definition contrasts with the alternative definition of $D_{ic,t}$ as *any*—rather than just the first—FDI, which we consider as part of a robustness in Appendix A. Our baseline does not consider investments following the first as separate events, under the assumption that subsequent FDIs are potentially connected to the initial one. Under this assumption, the effect on citations should still be attributed to the initial entry. In practice, we should note that 95 percent of subsequent investments take place within 5 years for any given firm-country pair with FDI entry. As we detail below, we set the number of stabilization lags, L to 5, which leads to the exclusion of these 95 percent subsequent investment from both control and treatment groups, leaving little scope for further investments to affect our results even in the case where we consider any event as our episode. Indeed, the robustness where $D_{ic,t}$ is an indicator for

[17]See Equations (24) and (25) in Dube et al. (2023) and the associated discussion.

any FDI reports statistically indistinguishable results.

To investigate the heterogeneous effects of different investment types, we estimate separate regressions for brownfield and greenfield FDI. Appendix A reports and discusses the results with two alternative definitions of the treatment variable. First, we consider each investment as an independent event, diverging from our initial approach where $D_{ic,t}$ represents the first investment by i in country c. Here, $D_{ic,t}$ includes *any* investment projects (the first and the subsequent ones). In this first robustness, we still estimate separate results for greenfield and brownfield FDI. Second, we estimate a single specification where $D_{ic,t}$ is any type of investment, regardless of whether it is greenfield or brownfield FDI. In both cases, we find results that are not significantly different from our baseline.

Dependent Variable: The Inverse Hyperbolic Sine of Cumulative Citations. Throughout the paper, we use as outcome the change in the inverse hyperbolic sine of a relevant measure of cumulative citations from 1995 onwards. In this study, we analyze the variation in cumulative citations since 1995, employing the inverse hyperbolic sine (asinh) transformation as our main outcome measure. We chose 1995 as the starting year to accumulate citations because it is the earliest year for which we can align M&A and patent data. The asinh transformation is commonly used as a substitute for the logarithmic transformation, accommodating data with zero and negative values.[18] While both transformations mitigate the impact of skewness in the distribution of observations, the interpretation of the asinh transformation's coefficients as percentage changes depends on the specific value around which effects are computed.[19] In our analysis, given that pre-entry citation counts average around 1, the estimated coefficients derived from this transformation would be close to, but slightly below, the actual percentage change in cumulative citations. Throughout the text, we will convert these coefficients back to percentage changes for clearer interpretation.

As reported in Section 4.4 and Appendix A, our results are robust to applying the $\log(1 + x)$ transformation, another common method used to handle zeros in the data. Furthermore, we explore the robustness of our results to focusing on citations between *triadic* patents—those registered simultaneously with the USPTO, JPO and EPO. The effects are qualitatively unchanged, and significantly higher when using the log transformation. Finally, to address potential truncation bias in citation counts towards the end of our sample, we run an alternative specification that only includes citations within five years after the initial patent application, excluding the final five years

[18]For instance, recent work such as Arora et al. (2021a), Azoulay et al. (2019), and Moretti (2021) used the asinh transformation for patent applications and scientific publications, which often have zero values.

[19]The inverse hyperbolic sine transformation approximates closely $\log(2x), -\log(-2x)$ for large positive and negative values of x, respectively, while it is close to linear around the origin. Chen and Roth (2024) examine how various log-like transformations, including the asinh transformation, affect estimates by assigning different weights to the intensive and extensive margins.

of data. These robustness checks yield broadly consistent results.

Choice of Stabilization Lags, L. We assume that treatment effects stabilize after five years, $L = 5$. We experimented with different assumptions with qualitatively unchanged results, and settled on this value to adopt the same specification to obtain all our results as well as to preserve a sufficiently high number of observations. To expand on this point, it is worth noting that the choice of L presents a variance-bias trade-off. The longer h is, the lower is the number of events included–which increases the variance of estimated coefficients–and the longer is the period of time for treatment effects to stabilize–which reduces bias in estimated coefficients. In the setting of this paper, we wish to show the path in the ten years around firm entries that take place over 2003 to 2022. When setting L to 5, by construction we can only include "clean" events from $2003 + L = 2008$ onwards. Furthermore, this only allows us to estimate the effects on citations for up to 14 periods after treatment, as our data is limited in terms of time span. In order to estimate effects for 5 years after first entry, we are therefore limited to a maximum value of $L = 14$. Further, PATSTAT citation data are reported with some lag, especially for what concerns USPTO data. Therefore, we should consider this data fully reliable only up to 2020, reducing the range of reasonable choices to a maximum of $L = 12$. Our choice of $L = 5$ therefore sits about the middle of the feasible ranges. To adopt a common specification, we are also limited by the US sample, which allows an accurate computation of citations made by investing firms up to 2015 only, as this is the year that the Arora et al. (2021b) database ends. As a result, the maximum number of lags that we can consistently set for all specification is $L = 6$. A previous version of this paper set $L = 6$ with largely unchanged results, but with substantially fewer events available for longer horizons in some settings. As reported in Section 4.4 and Appendix A, we verify the robustness of our baseline results to setting $L = 12$.

Choice of Controls. For all the results presented below, we follow the same specification, featuring as controls three lags of the asinh cumulative citations, setting $p = 3$ in the equation above, and the trend in this variable between 4 and 6 periods before the period considered, $y_{t-4} - y_{t-6}$. We make this choice based on observable differences in levels or trends that we detect in some of the settings that we consider. As discuss in the result section below, we find that adopting this specification broadly removes significant differences in pre-trends. In other words, in most cases, we do not detect violations of the parallel trend hypothesis conditional on these observables. In a previous version of the paper, we experimented with $p = 2$, dropping the term $y_{t-4} - y_{t-6}$ and replacing it with $y_{t-3} - y_{t-5}$, with qualitatively similar results.

4 Results

4.1 Main Result: Citation Flows Between Host Countries and Investing Firms

Our baseline specification estimates Equation (2) using the change in the inverse hyperbolic sine (asinh) of cumulative citations received by (or made by) firm i from (to) country c since 1995 as the dependent variable, and controlling for firm-year and country-year fixed effects, three lags of the dependent variable in levels and its trend based on changes 4 to 6 years prior to the event. We set the stabilization period, L, to 5 years.

Figures IIIa - IIId display the estimated coefficients β_h for the year of first entry of firm i in country c during our sample period (2003-2022). These figures demonstrate the positive effects of both greenfield and brownfield FDI on knowledge spillover from and toward the investing firms. For example, Figure IIIa shows that, five years after a brownfield investment, destination countries increase asinh cumulative citations to investing firms by 0.063. The pre-treatment $(t-1)$ average of asinh cumulative citations is 1.158, which corresponds to 1.43 untransformed cumulative citations. To interpret this coefficient, we translate these effects to percentage changes compared to pre-period untransformed cumulative citations applying the hyperbolic sine function as $\sinh(1.158+0.063)/\sinh(1.158)-1$. Following this procedure, we estimate that, five years after an investment, an investing firm receives, on average, around 7.8% more citations from the FDI destination country compared to other non-destination countries. Figure IIIb reports the coefficients for Greenfield FDI, which are statistically indistinguishable from the coefficients for Brownfield FDI. However, the pre-treatment average in this case is lower, totalling 0.66 cumulative citations before firm entry. As a result, the implied percentage change in citations from greenfield FDI is higher, at 10.6%. Considering the 95% confidence intervals for estimated coefficients returns an increase of $5.7-10\%$ and $9-12.2\%$ for brownfield and greenfield FDI respectively.

Turning to investors, Figures IIIc and IIId show increases in citations made to FDI host countries by investors of 10.8% and 13.4%, respectively. It is important to note that figure IIId displays a small pre-trend indicating that greenfield investors appear to select into destinations which they are increasingly citing over the years leading up to entry. However, extrapolating the small pre-trend still implies a positive treatment effect.[20] However, the presence of this trend raises doubts on the validity of the parallel trend hypothesis for this specification, cautioning against a causal interpretation of the results. In Section 4.2, we consider a specification at the firm-country-industry level where we are able to fully control for pre-trends. There, we find an increase of 10.8% in green-

[20]The trend implied by coefficients 4 and 5 years before entry corresponds to 0.0067 asinh cumulative citations per year. Subtracting this trend extrapolated to 5 periods after treatment implies an 11.8% increase in cumulative citations to host countries made by greenfield investors.

field investors' citations to the patents in the targeted country-industry. This smaller coefficient is line with our baseline finding, considering that, as we show also in Section 4.2, knowledge spillovers extend beyond directly targeted industries.

Figures IVa - IVd repeat the same specification for the sample of US firms identified through matches with the sample of Arora et al. (2021b). In addition to restricting to US firms and providing verified disambiguation of firm names, this matching also allows us to identify a patent's ultimate owner as reflected by corporate ownership structures. As a result, patents owned by the investors' branches, subsidiaries or owned companies will correctly be considered as belonging to the investor. Compared to the sample from direct name matches between PATSTAT and FDI data used for Figures IIIa - IIId, investing firms will generally have more patents attributed to them.[21] Carrying out the same computations as above, Figures IVa and IVb display an 8.4% and a 9.9% increase in citations, respectively. These numbers show little change compared to the ones mentioned earlier and are within the ranges implied by the results obtained using the global sample. When it comes to investors' citations, Figures IVc and IVd imply increases of 12.7% and 11.1%, respectively. Considering the uncertainty around these results, these estimates are not statistically different from those of Figure III.[22]

Once again, it should be noted that there is a slight pre-trend when it comes to greenfield FDI (Figure IVd). In contrast, we do not detect any violation of parallel trends in brownfield FDI cases. A possible explanation for this difference may lie in investment motives and the inherent characteristics of brownfield and greenfield FDI. The greenfield FDI data provided by fDi Markets provides motives of investment cited from announcements for a subset of greenfield FDI. These motives range from proximity to markets, to regulatory environment and skilled workforce availability, and cover 17,713 investments, accounting for 5.8% of the total. Among these investments, 43% of investors cite at least one of "skilled workforce availability," "technology and innovation," or "[proximity to] universities and research hubs." Additionally, 16.67% mention either of the latter two motives. These statistics reveal that many greenfield investors seek to acquire knowledge or skills. Furthermore, by nature, greenfield FDI involves opening a subsidiary of the investing firm, which usually operates in the same primary sector as the headquarters. These observations suggest

[21] Note that Arora et al. (2021b) check whether patents are reassigned as a result of ownership structures. As a result, some of the US firms in this sample may have less patents than in the world sample presented above. We say that firms will generally have more patents in the world sample because presumably acquiring firms are also firms that already own other companies and–indirectly–their patents.

[22] Note that there is a difference between the world sample and the US sample in terms of the countries of FDI investors. The former includes firms from around the world, while the latter only includes US firms. However, the estimated coefficients are similar, which is consistent with our country-level heterogeneity results, indicating that the characteristics of FDI investor countries do not lead to statistically distinguishable differences, while the host countries characteristics do.

that greenfield investors may target countries that have accumulated knowledge in the field of activity of the firm. This would manifest in the data as increasing citations leading up to the firm's entry. In contrast, brownfield FDI mostly targets firms in other primary SIC sectors. If correct, this reasoning would explain the difference in pre-trends between the two types of FDI.[23]

The reassignment of patent to ultimate owners in the Arora et al. (2021b) dataset may be problematic for our interpretation of results as knowledge spillovers. Indeed, if patents of the target firm are reassigned to investors, all citations made to those patents will also transfer to investors, mechanically increasing citations from country c to firm i. In Appendix A, we address this concern by only counting citations of patents that were *originally* assigned to US firms. As discussed there, and perhaps surprisingly, this restriction does not affect significantly the estimates of citations flowing from country c to firm i. Instead, we find significantly lower knowledge spillovers in the following direction. In other words, we find smaller spillovers when we consider only citations made patents originally assigned to US entities. This indicates that part of the effect found in Figures IVc and IVd is mechanical, and driven by reassignment of the destination country's patents to investors. It is important to note that this issue is not present in the world sample, where patent owners do not change over time and correspond to the original applicants.

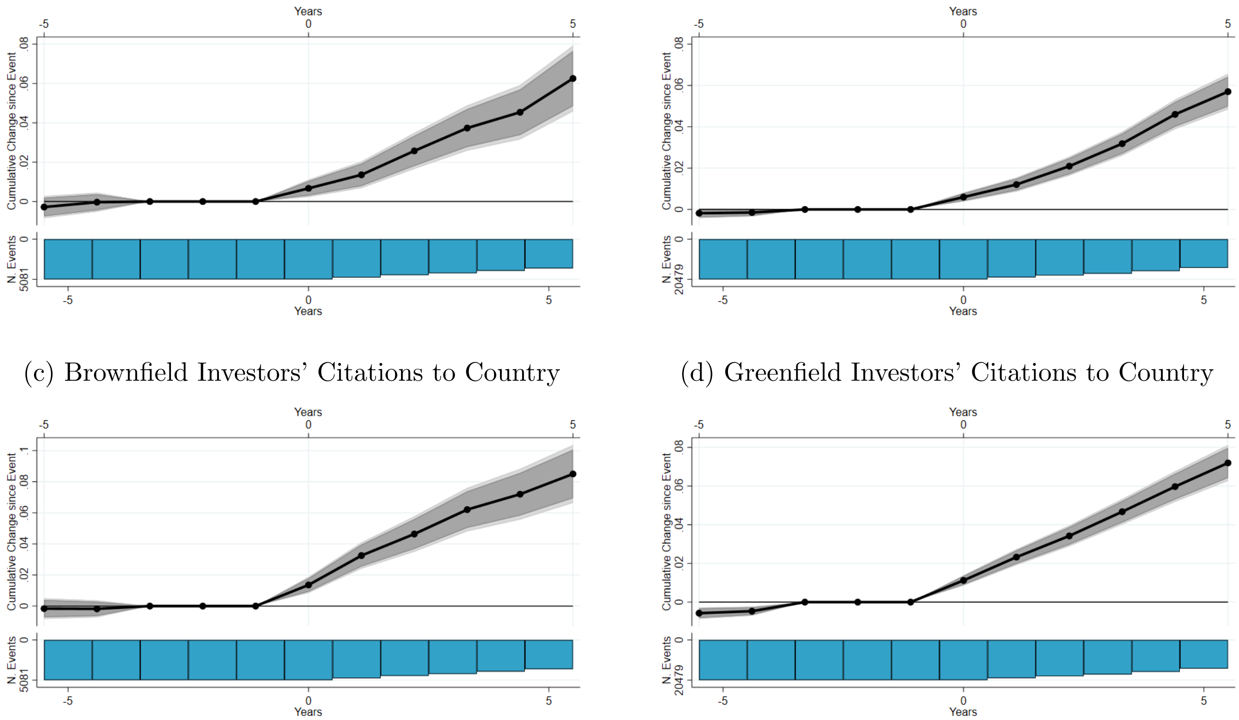
In summary, our baseline results suggest that FDI boosts knowledge diffusion both to and from FDI host countries with midpoints ranging from a 7.8% to a 12.7% increase in citations flowing between investors and host countries. However, when it comes to greenfield investment, we detect some pre-trends that threaten the causal interpretation of our results. In the next section, we pursue additional specifications and sample definitions to address this and other threats. In particular, we will present and discuss industry-level regressions that present no pre-trends for greenfield FDI.

4.2 The Scope of Knowledge Diffusion: Untargeted Firms and Country-Industry Results

In this subsection, we present two additional specifications aimed at refining the identification of knowledge spillovers and expanding the scope of our analysis. First, we assess the robustness of our results by excluding citations originating from firms that have been acquired in the brownfield FDI, for which target firm information is available. This exercise addresses the concern that increased citations from host countries to investors might result mechanically from citations from the newly acquired subsidiaries, representing self-citations rather than genuine knowledge spillovers. Second, we analyze citations at the country-industry level. This allows us to improve identification by controlling for additional country-industry-time fixed effects. Furthermore, this approach enables

[23]In our sample period, only 27.16% of M&A activities involve industries in the same primary SIC code.

Figure III: Impacts of FDI on Citations, World Sample



(a) Country's Citations to Brownfield Investors

(b) Country's Citations to Greenfield Investors

(c) Brownfield Investors' Citations to Country

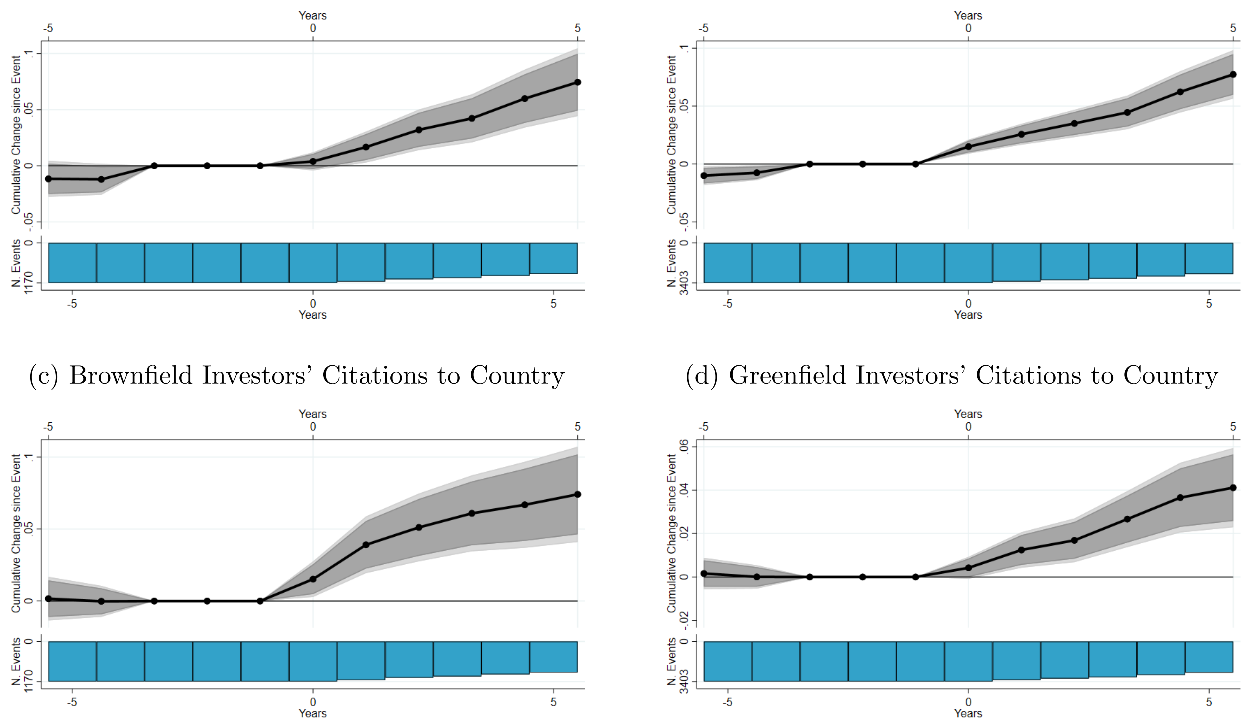
(d) Greenfield Investors' Citations to Country

Note: This figure presents the coefficients β_h estimated by our specification in Eq. (2) for the presented citation measures. "World sample" refers to the sample of firm-country pairs obtained matching PATSTAT company names with FDI data firm names following the procedure in Subsection 2.3. We include firm-time and country-time fixed effects as controls, as well as three lags of the dependent variable in levels and its trend estimated using the change in 6 to 4 years before the event. We set the stabilization period to five periods. The treatment variable is the first FDI investment carried out by each firm in each destination country. The bottom bars represent the number of treated samples.

us to explore whether these spillover effects extend beyond directly targeted industries and how they are influenced by technological linkages between industries. Specifically, we use an industry-level measure of technological relatedness to examine whether industries closely related to those directly impacted by FDI receive greater knowledge spillovers from these investments. This includes sectors both directly targeted by FDI and those indirectly linked through technology citation networks.

In the next subsection, we present additional results that demonstrate the impact of similarity between the patent stocks of host countries and investors on knowledge spillovers. Together with the results from this subsection, these findings suggest that knowledge spillovers are not limited to directly related firms and industries, further supporting our conclusion that FDI leads to cross-border knowledge spillovers.

Figure IV: Impacts of FDI on Citations, US Sample



(a) Country's Citations to Brownfield Investors

(b) Country's Citations to Greenfield Investors

(c) Brownfield Investors' Citations to Country

(d) Greenfield Investors' Citations to Country

Note: This figure presents the coefficients β_h estimated by our specification in Eq. (2) for the presented citation measures. "US sample" refers to the sample of firm-country pairs obtained from the Arora et al. (2021b) dataset, which are then matched to firm names in our FDI datasets following the procedure in Subsection 2.3. We include firm-time and country-time fixed effects as controls, as well as three lags of the dependent variable in levels and its trend estimated using the change in 6 to 4 years before the event. We set the stabilization period to five periods. The treatment variable is the first FDI investment carried out by each firm in each destination country. The bottom bars represent the number of treated samples.

Dropping Citations From Target Firms. In order to address the mechanical threat arising from self-citations through acquired subsidiaries, we assess the robustness of our findings to excluding citations from acquired firms. We do so applying our name matching strategies to exclude citations coming from any company that has a name similar to the acquired target or to the acquiring firm. For greenfield investment, the data does not include the names of the newly established entities. Therefore, this robustness exercise is limited to brownfield investment episodes.

We carry out this exercise for both our world sample and the US sample discussed above. In this case, the greater reliability of ownership structures provided by the US sample is even more valuable, since we are concerned about potential self-citations by acquiring firms. We report the resulting coefficients in Figure V. We estimate an increase in host country's citations to brownfield investors' patents of 8.9% and 7.3% in the world and US samples, respectively. In both cases, the bounds implied by coefficient confidence intervals cover the estimates of 7.8% and 8.4% for the same samples presented in the previous section. When it comes to investors' citations to host countries, the corresponding numbers are 10.8% (the same as baseline) and 4.5%, respectively. The latter estimate is the only significantly different from the baseline. Its smaller size suggests that about two thirds of the effect in Figure IVc is due to citations made by investors to acquired firms in destination countries.

Firm-Industry Evidence: Setup and Specifications. We now move to investigate whether knowledge spillovers transmit more intensely to sectors that are directly targeted by FDI activities, and whether knowledge spillovers extend beyond targeted industries. To do so, we enrich our specification by adding the ISIC 2-digit industry dimension as follows:

$$y_{icl,t+h} - y_{icl,t-1} = \beta_h T_{icl,t} + \sum_{k=1}^{3} \gamma_k^h y_{icl,t-1-k} + \delta_{ilt}^h + \delta_{clt}^h + \varepsilon_{icl,t}^h, \tag{4}$$

where now $y_{ilc,t}$ is the asinh cumulative citations made to patents owned by firm i by patents assigned to country c's patents for two-digit ISIC code l within a 5-year period including t, or vice versa. Notably, we can now include firm-industry-time and country-industry-time fixed effects. These terms allow us to address some of the selection concerns that we mentioned in the discussion of our baseline results. In particular, we can now correct for selection concerns stemming from target industries growing in country c at a specific time, as well as for firm i becoming more technologically relevant in sector l worldwide. We define the treatment as the first year when a US firm enters a destination country c's industry l. The resulting coefficients can be interpreted as differences in citations in the treated destination-industry pairs compared to other non-treated destination-industry pairs before and after FDI, while controlling for overall firms' patenting activities across specific industries and destination firms' overall patenting capacities across specific industries. We obtain citations at the industry level using sector classifications reported by PATSTAT. The definition of industry used to attribute treatment, $T_{icl,t}$ varies according to the type of FDI we consider. The data provided by fDi Markets includes a detailed description of the sectors targeted by FDI, which can vary for the same firm i across different projects and destinations, which we map to two-digit ISIC codes. When it comes to M&A data, the information is less precise, since our data source contains standardized information only on the *primary* SIC sector of the investing and target firm. As a result, this sector is not guaranteed to be the effective sector that receives FDI, since the target firm may be active in several sectors and the brownfield FDI may target industries that do not necessarily coincide with the primary SIC code of the target entity. For this reason, in what follows we focus our presentation on greenfield FDI, and discuss the less reliable results on brownfield FDI briefly. When it comes to brownfield FDI, we estimate two specifications where l is set to the industry of the investor or of the target firm.

Knowledge Input Coefficients. We also check whether industries that rely more on knowledge from treated industries experience larger knowledge flows. To do so, we calculate "knowledge input coefficients" of industry l from industry k at the ISIC 2-digit level within destination country c as:

$$a^c_{l \leftarrow k} \equiv \frac{\text{Total citations made by industry } l \text{ towards } k \text{ in country } c \text{ between 1990-2000}}{\text{Total citations made by industry } l \text{ in country } c \text{ between 1990-2000}}, \quad (5)$$

that is, the share of overall citations by sector l that are made to sector k. Therefore, $a^c_{l \leftarrow k}$ ranges between 0 and 1 and the sum of coefficients across k is 1. In other words, $a^c_{l \leftarrow k}$ measures the strength of the knowledge connection between the FDI target industry k and sector l for which we wish to measure knowledge spillovers. To avoid potential endogeneity issues, we calculate the knowledge input coefficients using citations from the period 1990-2000, which predates our empirical analysis. This measure introduces an additional dimension of treatment intensity for sector l, under the assumption that the effect of investment on citations should be mediated by how strongly sector l's technology relies on the knowledge produced by industries k that receive FDI. By considering treatment effects beyond directly targeted industries k, this specification accounts for knowledge spillovers that may extend beyond the directly targeted industry.

In this case, the treatment is defined as $T_{icl,t} \equiv \sum_k D_{ick,t} \times a^c_{l \leftarrow k}$, where $D_{ick,t}$ is the dummy for the first year of FDI for each firm-destination-industry triplet used in the previous analysis. As a result, the intensity of treatment, $T_{icl,t}$ will be stronger for industries l that more intensely cite sector k. These knowledge coefficients including the same-industry knowledge coefficient, $a^c_{l \leftarrow l}$, are also smaller than 1. Therefore, industries directly targeted by investment will also have their treatment $T_{icl,t}$ scaled by the intensity at which their patents cite same-industry patents.

Sample Selection. In order to estimate the specification above, we must fill in all possible firm-country-industry-time combinations, as these represent observations with 0 citations. The resulting dataset for all possible combinations would be prohibitively large and extend beyond the computational capabilities that we have access to. We therefore choose to restrict attention to the US sample and consider the top 30 innovative destinations, as identified by the share of worldwide patents of assignees' patent offices. This leaves us with around 276 million observations, comprising investments from US firms to the top 30 destinations across 98 ISIC 2-digit industries. It is also important to note that, in the knowledge input coefficient specification, we choose to restrict to firms that make only an investment to one target industry per country, which encompasses 90% of total observations. This appears to be the least arbitrary choice to compute a meaningful measure of treatment intensity using knowledge coefficients. If we were to keep multiple-industry investors, we would need to make some (arbitrary) decisions on how to aggregate input coefficients referring to industry l's connection to multiple treated industries k, complicating the interpretation of the treatment variable.

Industry-level Greenfield FDI Results. Figure VI reports the industry-level results for greenfield FDI obtained estimating Equation (4). The results for directly targeted industries are qualitatively unchanged relative to the baseline. In addition, we do not detect any clear violation of the parallel trend hypothesis. We believe that this is due to the tighter identification afforded to us by the inclusion of country-industry-time and firm-industry-time fixed-effects. In particular, this inclusion would address the fact that country-industries targeted by greenfield FDI may be acquiring knowledge related to investors' patent in the years leading up to the investment, the mechanism which we believe led to detectable pre-trends in our baseline results. Quantitatively, we obtain an average 9% increase in citations made by host industries to greenfield investors. In the opposite direction, investors increase their citations to target industries by 10.8%. These results are at the lower end of the effects that we estimated for our baseline considering destination countries as a whole, as should be expected if knowledge spillovers extend beyond the confines of targeted industries.

The results for the treatment interacted with the input coefficient show clear positive effects without a pre-trend. These positive effects provide evidence for knowledge spillovers extending beyond directly targeted sectors. Indeed, as visible from the blue bars reported under each graph, the number of country-industries that are treated in this specification is much larger than the number of directly targeted industries, implying a strong contribution of technologically related–but not directly targeted–industries to estimated treatment effects.[24] Quantitatively, these estimates imply that investments targeting an industry with an input coefficient of 1 cause a 32.7% increase in citations made by related industries to greenfield investors. As discussed, $a_{l \leftarrow k} = 1$ if and only if treated sector l directs all its citations to sector k, which is never the case empirically. Instead, we propose two interpretations of these estimates using average input coefficients. First, we compute the effect of FDI on directly targeted sectors implied by the results of this specification. On

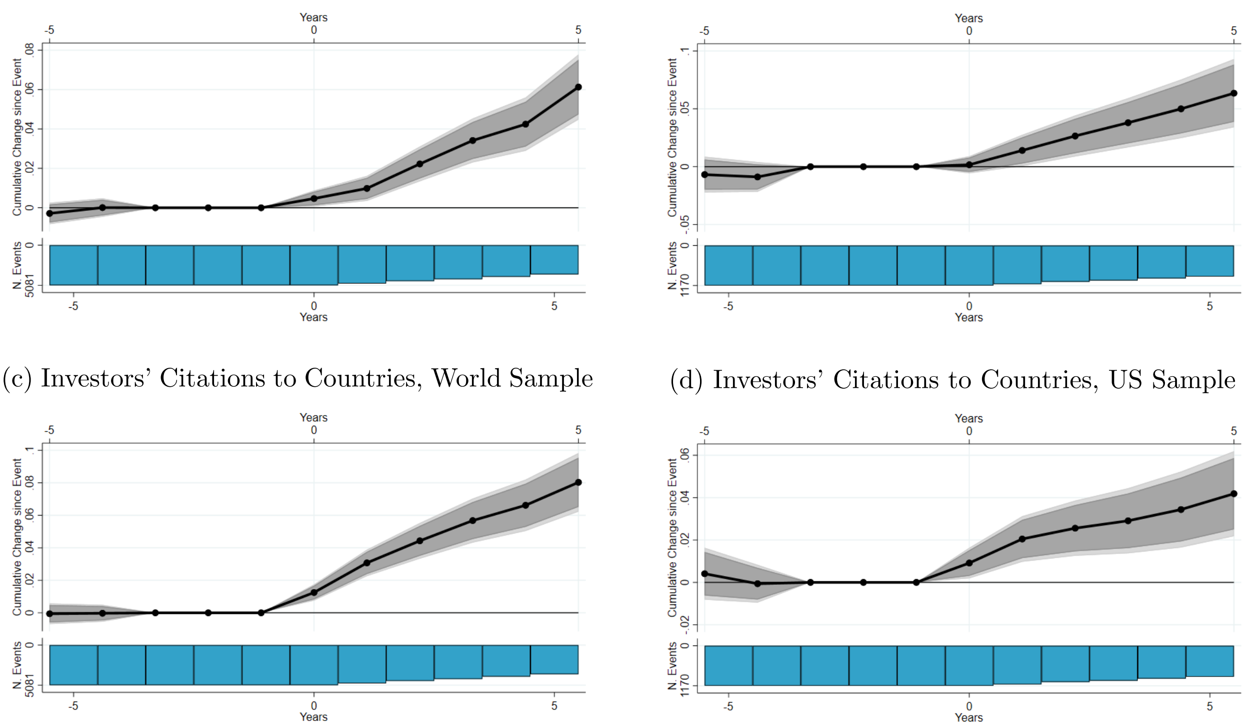
[24]For example, comparing Figures VIa and VIb, we see that $3,487$ industry-country pairs are directly targeted by greenfield FDI, a mere 2.75% of the total number of industry-country pairs that are treated according to the knowledge input specification $(126,598)$.

average, the share of citations that an industry directs towards itself, $a_{l \leftarrow l}$, is 22.4%, which implies a 7.2% increase in citations for sectors directly targeted by FDI. This value is smaller than the 9% obtained above for the industry-level analysis, but it falls within the 6.6% lower bound implied by confidence interval in Figure VI. Second, we can compute the effect for the average treated sector, that is industries that have $T_{icl,t} > 0$. In this case, the average of the treatment variable, $T_{icl,t}$ is 5.7%. This number means that, on average, industries connected to FDI target sectors direct 5.7% of their citations to those sectors. This number implies a 1.8% increase in citations made to investors by sectors that are affected directly or indirectly by FDI activity. Coincidentally, the percent treatment effects for investors citations to industries result in identical effects at the first decimal digits, arising from slightly higher coefficient and citations in asinh terms.

We also estimate a specification which includes an interaction between $T_{icl,t}$ and a dummy denoting whether industry l is directly targeted by greenfield FDI to ensure that our coefficients are not driven by directly targeted industries only. In facts, we find that the percent increase in citations for non-targeted industries with $a_{l \leftarrow k} = 1$ is 28.5% while the corresponding number is a lower 22.7% for targeted sectors. If anything, this points to the fact that spillovers identified by the input coefficients specifications are larger (in percent terms) in related sectors compared to directly targeted sectors. The figures corresponding to this latter specification are available upon request.

Industry-level Brownfield FDI Results. As mentioned above, industry information for brownfield FDI is limited to the *primary* industry of investors and target firms, which are therefore our only available choices for l in this setting. Unlinke in the FDI case, the actually treated industry may differ widely from the industries we assign treatment to, leading to potential attenuation bias. First, we shall discuss results where we using acquiring firms' industries. In this case, we find positive but not significant effects for citations from industries that are directly invested in to investor firms, and imprecisely-estimated zero effects for the reverse citation flow from investors to destination country-industries. In both cases, we believe that this is due to acquiring firms' industries not providing an accurate indication of investment industries. When it comes to the knowledge input coefficient specification, we recover results similar to the greenfield case. This is likely because the actual target industries rely on knowledge related from the acquiring firms' industry, so assigning treatment $T_{icl,t} > 0$ attenuates the measurement error arising from using only the industry of the acquiring firm. When we assign direct treatment to industry l using the primary industry of the target firm, we again find no effect on citations from investing firms to the target country-industry. However, we do find positive effects on citations from the target country-industry to investing firms. This discrepancy may be due to the fact that the industry actually subject to FDI is technologically related to the primary industry of the target firm, but does not coincide with it. In this case, the industry of the target firm would still receive indirect knowledge spillovers, but would not generate spillovers back to the investors' industry.[25] This conjecture is supported by the fact that we recover positive knowledge spillovers in both directions when employing the knowledge input coefficient specification. In summary, subject to the caveats of a less precise industry identification, the industry-level results described for greenfield FDI appear to extend to the brownfield context. To streamline the exposition, we do not report the graphs related to these specification, but the related results are available upon request.

Figure V: Impacts of Brownfield FDI on Citations, Excluding Targeted Firms



(a) Countries' Citations to Investors, World Sample

(b) Countries' Citations to Investors, US Sample

(c) Investors' Citations to Countries, World Sample

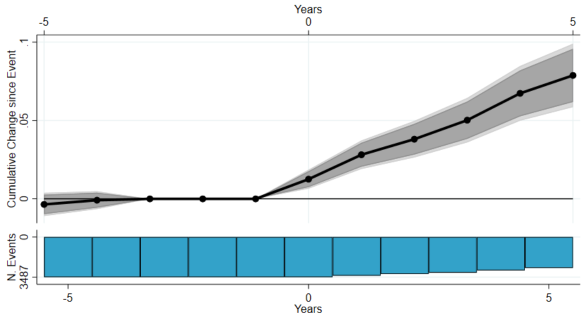
(d) Investors' Citations to Countries, US Sample

Note: This figure presents the coefficients β_h estimated by our specification in Eq. (2) for the citation measures excluding citations from potential acquired firms. "World sample" and "US sample" refer to the sample of firm-country pairs obtained matching PATSTAT company names and names in the Arora et al. (2021b)'s data set with FDI data firm names, respectively. The matching process follows the procedure in Subsection 2.3. We include firm-time and country-time fixed effects as controls, as well as three lags of the dependent variable in levels and its trend estimated using the change in 6 to 4 years before the event. We set the stabilization period to five periods. The treatment variable is the first FDI investment carried out by each firm in each destination country. The bottom bars represent the number of treated samples.

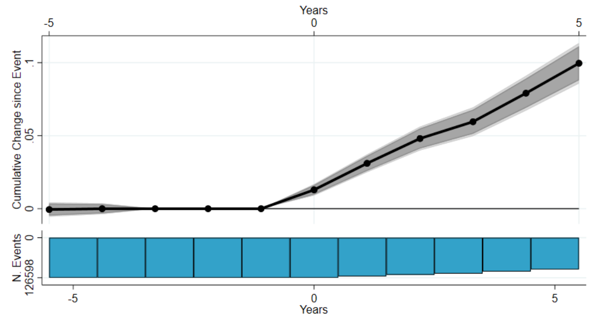
[25] It is important to note that the target and investors' primary industry do not coincide in 75% of cases. Thus, the fact that the target firm industry is related to the treated sector does not imply that the investors' industry is related to the sector actually treated by FDI.

Figure VI: Citation Flows Between Greenfield Investors and Countries' Industries

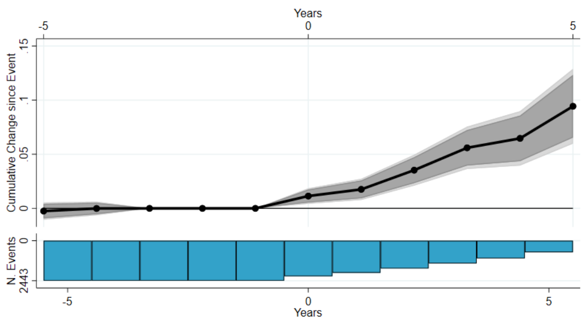
(a) Host Industries' Citations to Investors



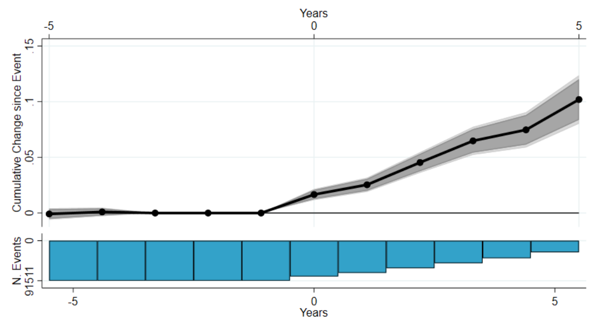
(b) Related Industries' Citations to Investors



(c) Investors' Citations to Host Industries



(d) Investors' Citations to Related Industries



Note: This figure presents the coefficients β_h estimated by our specification in Eq. (4) for the citation measures presented. The sample used in this figure is the "US sample" described in Subsections 2.3 and 4.2. We include firm-industry-time and country-industry-time fixed effects as controls, as well as three lags of the dependent variable in levels and its trend estimated using the change in 6 to 4 years before the event. We set the stabilization period to five periods. The treatment variables are the first FDI investment carried out by each firm in each country-industry pair and its interaction with the knowledge input coefficient described in Subsection 4.2 for the left and right columns, respectively. The bottom bars represent the number of treated samples.

4.3 Heterogeneity by Host Countries' Absorptive Capacity

In this subsection, we explore the heterogeneity of our results depending on host countries' characteristics that may help them reap the benefits of FDI. These characteristics are commonly referred to as "absorptive capacities." For instance, Alfaro (2017) discusses the role of absorptive capacity in the context of the effect of FDI on economic growth. Among the factors commonly discussed in the literature as necessary to realize the benefits of FDI, two are particularly relevant to our focus on knowledge spillovers: a business environment receptive to the adoption of new technologies; and adequate levels of human capital. In the following, we use the existing size of a country's patent stocks and the similarity of those stocks to those of investors as proxies for these absorptive capacities, respectively.

We use the following specification to allow for heterogeneity in the absorptive capacities of host countries:

$$y_{ic,t+h} - y_{ic,t-1} = \beta_h D_{ic,t} + \beta_h^I D_{ic,t} \times I_{ic,t} + \sum_{k=1}^{p} \gamma_k^h y_{ic,t-k} + \eta' \mathbf{x}_{ic,t-1} + \delta_{it}^h + \delta_{ct}^h + \varepsilon_{ic,t}^h, \qquad (6)$$

where I_{ic} is an interaction term capturing absorptive capacity, which is also included in the vector of controls $\mathbf{x}_{ic,t-1}$ where relevant. In a first set of specifications,the term $I_{ic,t}$ is a dummy equal to 1 when destination c belongs to the top 10% of destinations by the quantity of patents worldwide listed in the PATSTAT data.[26] This dummy variable is used to distinguish host countries that have a high absorptive capacity in terms of innovation capability. Since the interaction variable is constant in this case, we do not include it in the sets of controls as it would be collinear with a set of country fixed effects. We employ this first measure to capture the capacity of destination countries to absorb new technologies in general.

The second set of results utilizes the similarity between countries' patent stocks and investors' patent stocks as the interaction variable $I_{ic,t}$. This similarity is calculated as the cosine similarity between the vectors of the number of investors' patents and countries' patents across different ISIC 2 digit industries.[27] For each time period t, this similarity is computed based on the patent stock accumulated in the previous ten years. We use this measure to capture the technological suitability of destinations to investors' technologies and as an indicator of the presence of relevant human capital. In this case, we include $I_{ic,t}$ in the control vector, $\mathbf{x}_{ic,t-1}$.

Heterogeneity by Destinations' Patent Stocks. Figures VII and VIII display our results where $I_{ic,t}$ is a dummy equal to 1 when destination c belongs to the top 10% of destinations. In this and the following figures, the upper panels report the coefficients for the baseline where $I_{ic,t} = 0$, β_h, while the lower panels report the coefficient on the interaction term, β_h^I. The interaction term is positive and significant, pointing to the fact that more citation flows arise in both directions when the FDI destination is among the top 10% patent producers. Starting from Figure VIIa, we see that being among top patent producing destinations boosts knowledge spillovers in brownfield target countries, more than doubling the estimated baseline effect in asinh terms. Since the pre-treatment number of citations differs across destinations, we compute the effects in corresponding percentage terms, inverting our transformation as discussed in Section 4.1.[28] Compared to the pre-

[26] The top five destinations in our sample are the U.S., China, Japan, Germany and Korea

[27] For example, assume there are two ISIC industries, I1 and I2. Investor A has 3 patents in I1 and 4 patents in I2, while host country B has 4 patents in I1 and 3 patents in I2. To calculate the cosine similarity of their patent vectors, we use the formula $\frac{\mathbf{A}\cdot\mathbf{B}}{\|\mathbf{A}\|\|\mathbf{B}\|} = (12+12)/(5*5) = 0.96$. Here, $\mathbf{A} = (3,4)$ and $\mathbf{B} = (4,3)$ are the patent vectors of investor A and country B, where each element represents the number of patents in industries I1 and I2, respectively.

[28] Specifically, we compute percentage changes corresponding to baseline results (upper panels) as

investment period, brownfield FDI host countries in the bottom 90% of patenting increase their citations to investors by 5.1% after five years. The corresponding increase for top 10% destinations by patenting is more than double, at 12.3%. This percentage increase accounts for the fact that pre-investment citations are higher (almost double) in destination countries. When it comes to the greenfield FDI effects displayed in Figure VIIb, the difference is even starker. Most innovative destinations increase their citations by 21.3%, while the bottom 90% of patent producers only increase citations by 2.73%. Figure VIII shows that this heterogeneity applies also to investors' citations to FDI host destinations. In the case of brownfield investments, citation increase by 6.8% in less innovative destinations and by 17.6% among the top 10%. When it comes to greenfield FDI, we obtain estimates of 4.9% and 24.5%, respectively.

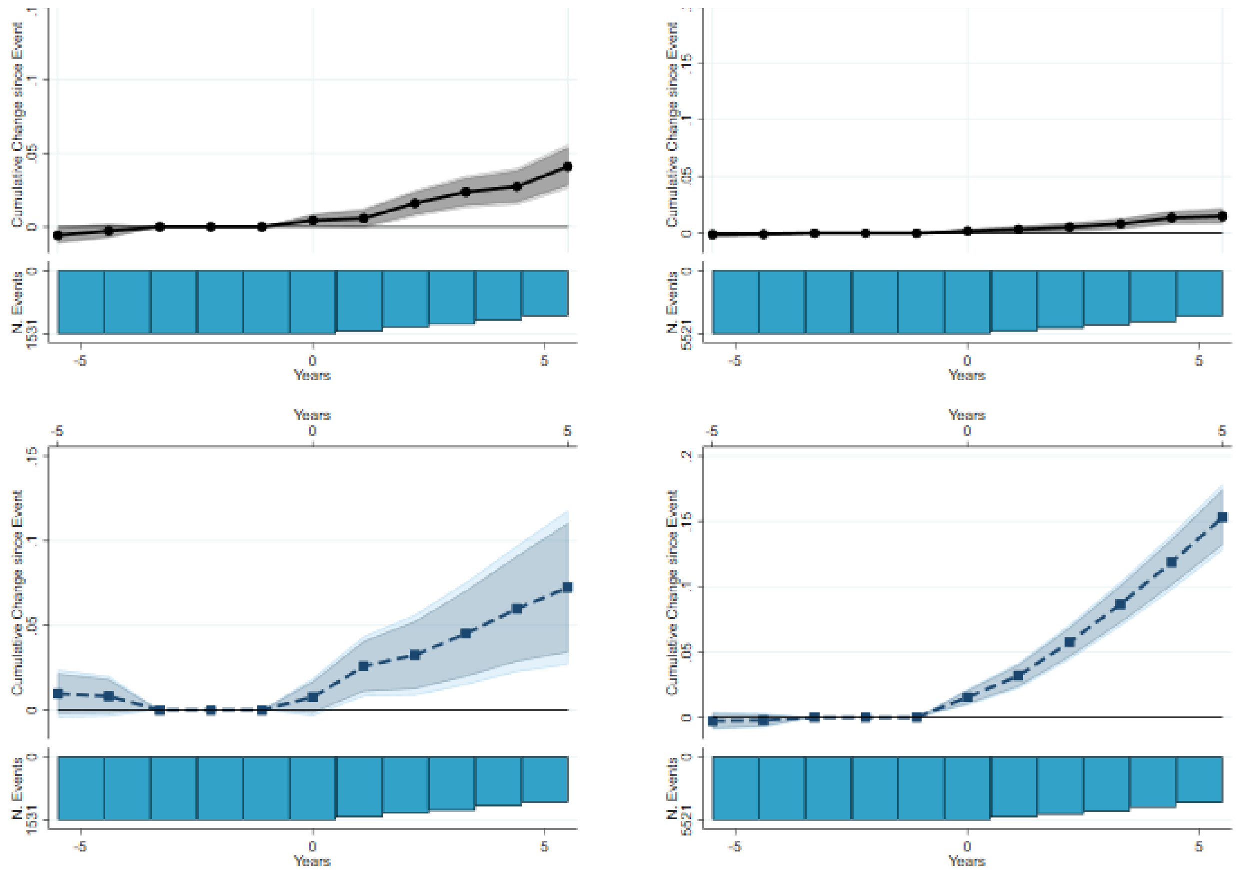
It is worth noting that the interaction used for this exercise compares both the same firm across different destinations, as well as the overall effect of firms investing in top 10% destination versus other firms that do not invest in these destinations. To partially alleviate this concern, we repeat the same specification restricting attention to firms investing in multiple destinations, which allows us to focus more closely on cross-destination heterogeneity. The results are qualitatively unchanged–albeit more noisy due to the smaller sample–and available upon request.

Heterogeneity by Destinations' Technological Similarity. Figures IX and X report the coefficients from the specification where $I_{ic,t}$ is a continuous variable representing the average cosine similarity between investing firms' and host countries' patent stocks, based on the ISIC 2-digit code of the patents' industries in the ten years preceding time t. We multiply the cosine similarity by a factor of 10 for readability, so $I_{ic,t}$ ranges between 0 and 10. For reference, $I_{ic,t}$ averages 1.26 in our sample. In the figures below, the upper plots report the coefficients corresponding to 0 patent stock similarity, while the lower panels report the coefficients β_h^I on the interaction term. Once again, we confirm that higher absorptive capacity generates larger citation flows. Interestingly, it appears that countries benefit from knowledge spillovers even in cases where technological similarity is very low, as visible from the upper panels of Figure IX. The same is not true for knowledge spillovers from countries to investors, in which case we see no investors' citations from greenfield FDI investors when the destination's patent stock is 0. Perhaps surprisingly, technological similarity seems to play a lesser role in the case of brownfield FDI. Moving from the 25^{th} to the 75^{th} percentile of similarity (from 0.33 to 1.66) only increases the citations generated by brownfield FDI by about 32%, while it more than doubles the effect of greenfield FDI. Considering the difference in pre-investment

$\sinh(\bar{y}^{\text{base}}_{ic,t-1} + \hat{\beta}_h)/\sinh(\bar{y}^{\text{base}}_{ic,t-1}) - 1$ and results for the lower panels as $\sinh(\bar{y}^{\text{high}}_{ic,t-1} + \hat{\beta}_h + \hat{\beta}^I_h)/\sinh(\bar{y}^{\text{high}}_{ic,t-1}) - 1$, where $\bar{y}^{\text{base}}_{ic,t-1}$ and $\bar{y}^{\text{high}}_{ic,t-1}$ represent the sample averages of the outcome variable in the period before first firm-entry for the sample of target countries.

Figure VII: Host Countries' Citations to Investors, by Host Countries' Patent Stock. Upper panels: Bottom 90%; Lower Panels: Top 10%

(a) Countries' Citations to Brownfield Investors (b) Countries' Citations to Greenfield Investors



Note: This figure presents the coefficients β_h estimated by our specification in Eq. (6) for the presented citation measures. The sample used in this figure is the "World sample", which refers to the sample of firm-country pairs obtained matching PATSTAT company names with FDI data firm names following the procedure in Subsection 2.3. We include firm-time and country-time fixed effects as controls, as well as three lags of the dependent variable in levels and its trend estimated using the change in 6 to 4 years before the event. We set the stabilization period to five periods. The treatments variable are the first FDI investment carried out by each firm in each destination country and its interaction with a dummy dividing host countries based on their ten-year patent stocks. The bottom bars represent the number of treated samples.

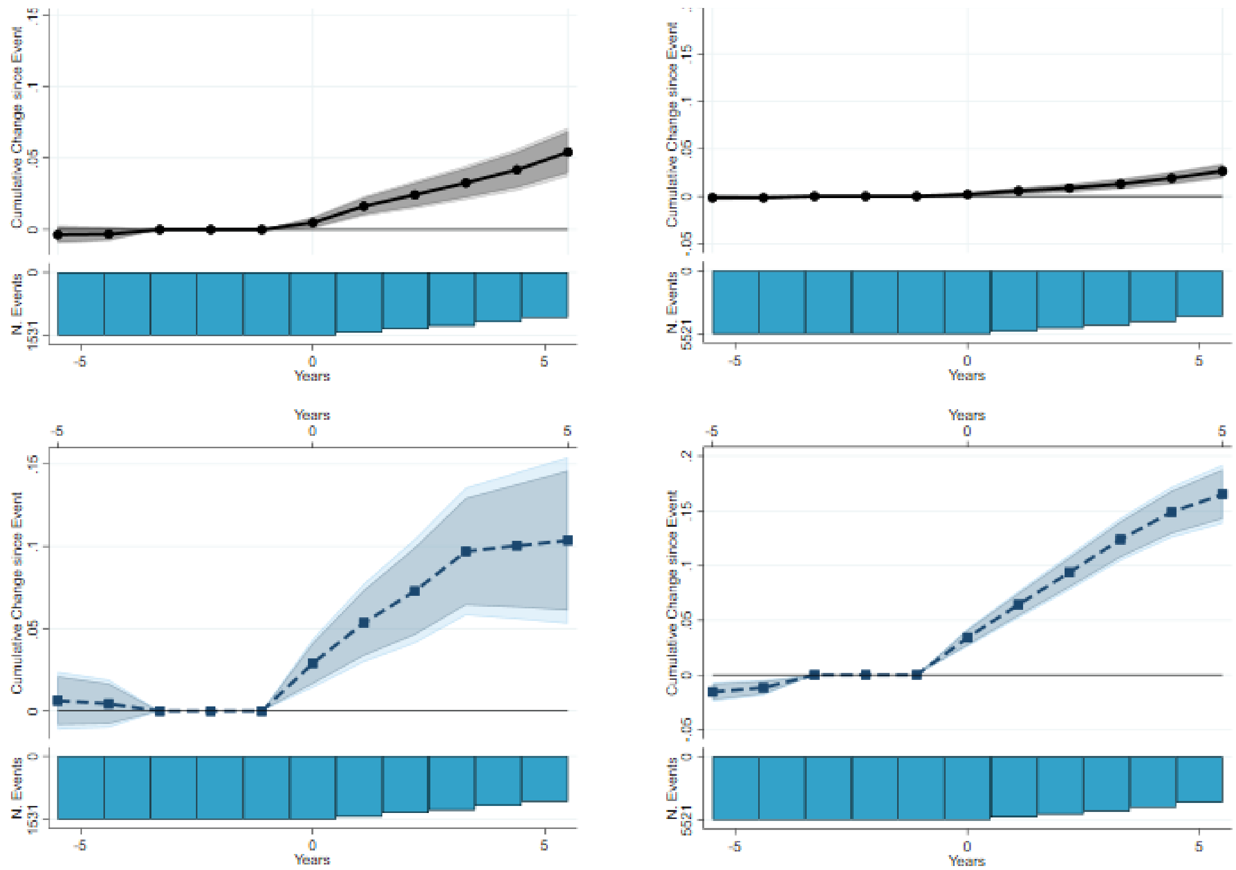
citation flows, this implies that percentage effects are actually slightly lower in countries with higher similarities. The percent effect of brownfield FDI on citations from countries at the 25^{th} percentile of similarity is 8%, while at the 75^{th} percentile citations increase by 7.7%. Considering that the pre-treatment average of citations is more than double in the latter case, this still represent a higher absolute value change than what we estimate for the 25^{th} percentile.[29] When it comes

[29]We compute these average effects starting from the average number of citations for countries with similarity within a 0.01 radius of the corresponding similarity percentile. For example, the first quartile of similarity is 0.33,

Figure VIII: Investors' Citations to Host Countries, by Host Countries' Patent Stock. Upper panels: Bottom 90%; Lower Panels: Top 10%

(a) Brownfield Investors' Citations to Countries

(b) Greenfield Investors' Citations to Countries



Note: This figure presents the coefficients β_h estimated by our specification in Eq. (2) for the presented citation measures. The sample used in this figure is the "World sample", which refers to the sample of firm-country pairs obtained matching PATSTAT company names with FDI data firm names following the procedure in Subsection 2.3. We include firm-time and country-time fixed effects as controls, as well as three lags of the dependent variable in levels and its trend estimated using the change in 6 to 4 years before the event. We set the stabilization period to five periods.The treatments variable are the first FDI investment carried out by each firm in each destination country and its interaction with a dummy dividing host countries based on their ten-year patent stocks.

to greenfield FDI, we find a more pronounced effect of technological similarity. The corresponding citation effects in this setting are 6.7% (at the 25^{th} percentile) and 9.2% (at the 75^{th}). The case of investors' citation to host countries confirms a more important role of technology similarity for knowledge diffusion in the context of greenfield FDI. Indeed, Figure Xb shows that citations do not increase if there is no similarity in patent stocks between greenfield investors and host countries.

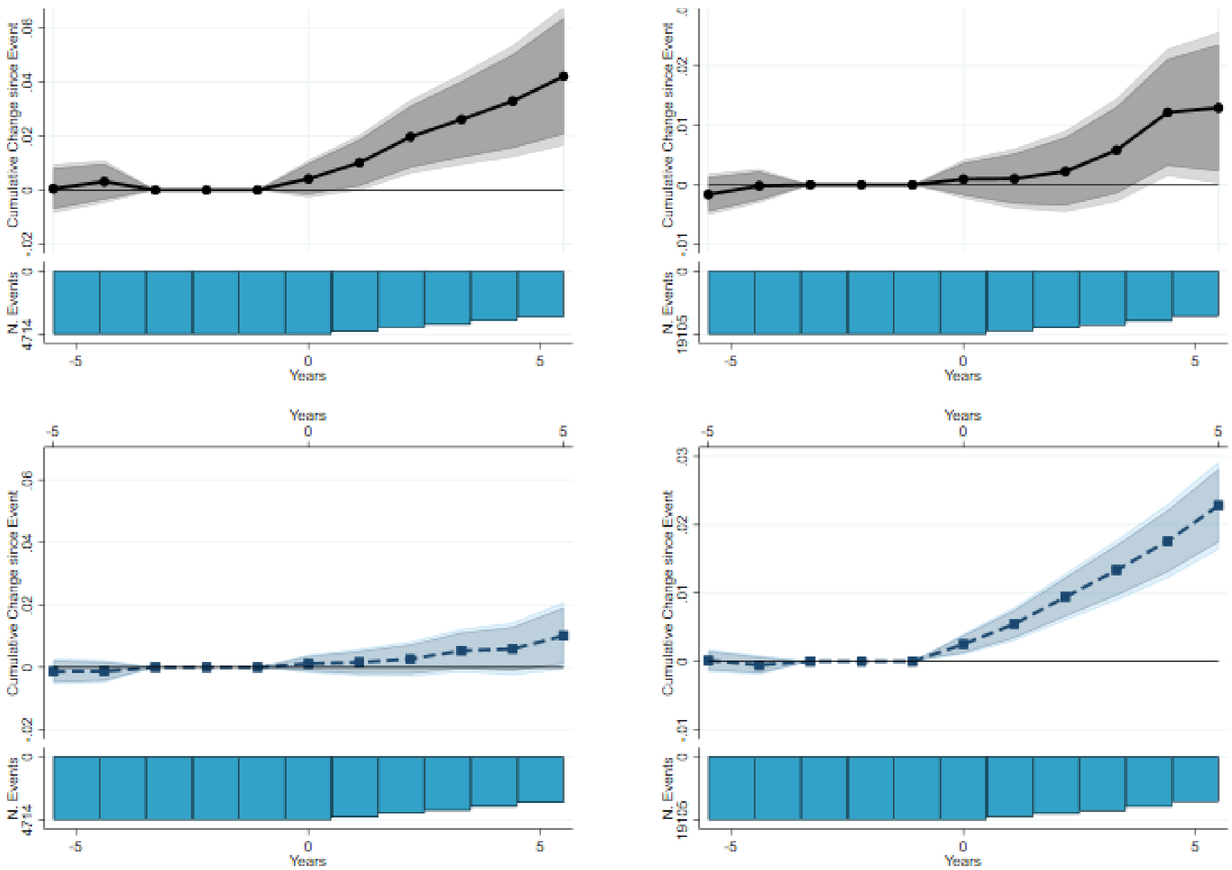
so we compute the average effect at the 25^{th} percentile around the average of asinh citations for countries that: (i) receive brownfield FDI, and (ii) have average similarity between 0.32 and 0.34. This number is computed for the year before the first investment.

In quantitative terms, the coefficients in this figure imply that citations increase by 4% at the 25^{th} percentile of similarity and 11.1% at the 75^{th}. In the case of brownfield investments reported in Figure Xa we obtain increase of 12% and 10%, respectively. As in the case of hosts' citations to investors, the effect on the absolute number of citations remains larger at higher patent similarity, even though the percentage effect declines slightly.

Figure IX: Host Countries' Citations to Investors, by Host Countries' Technology Similarity. Upper panels: Baseline; Lower Panels: Interaction Effect.

(a) Countries' Citations to Brownfield Investors

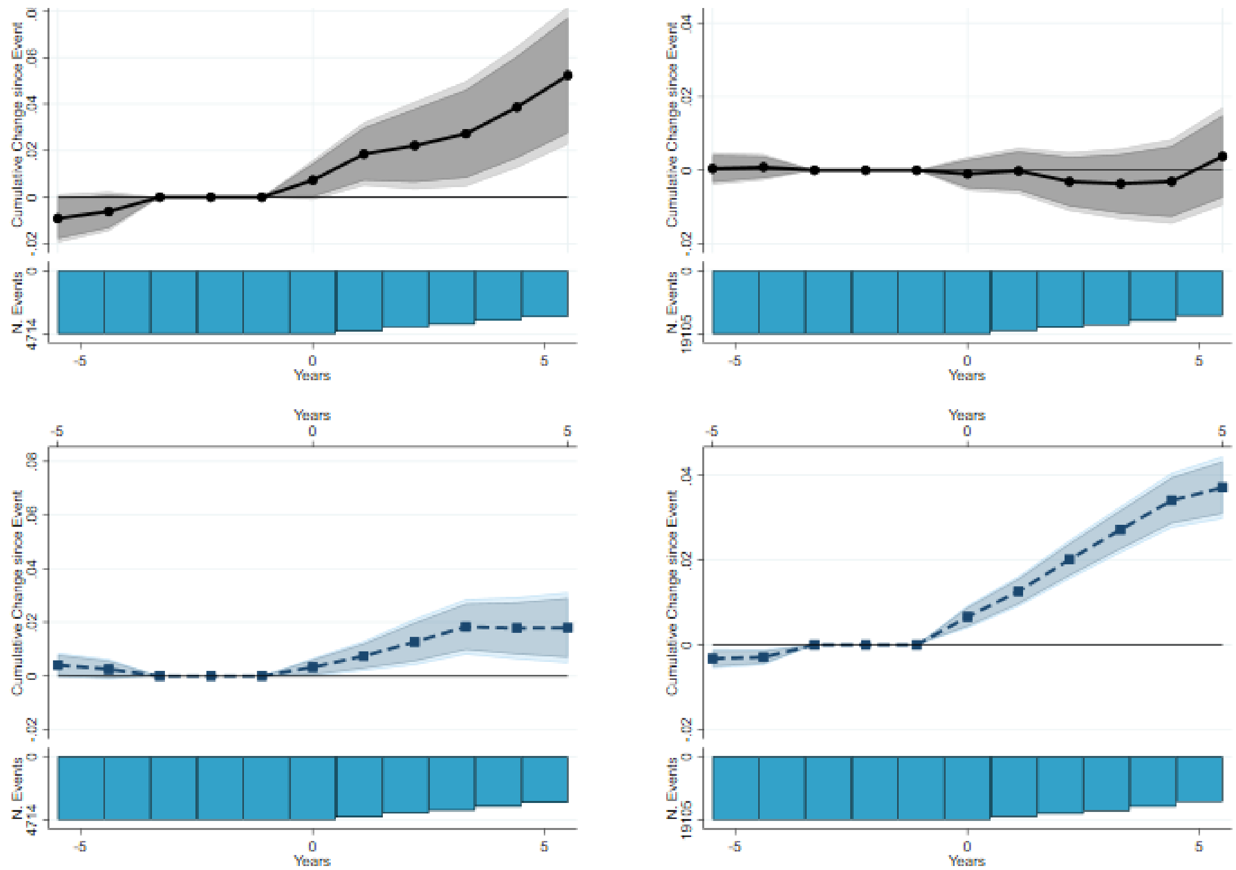
(b) Countries' Citations to Greenfield Investors



Note: This figure presents the coefficients β_h estimated by our specification in Eq. (2) for the presented citation measures. The sample used in this figure is the "World sample", which refers to the sample of firm-country pairs obtained matching PATSTAT company names with FDI data firm names following the procedure in Subsection 2.3. We include firm-time and country-time fixed effects as controls, as well as three lags of the dependent variable in levels and its trend estimated using the change in 6 to 4 years before the event. We set the stabilization period to five periods. The treatments variable are the first FDI investment carried out by each firm in each destination country, as well as its interaction with the similarity between technologies of the investors and host countries. The bottom bars represent the number of treated samples.

Figure X: Investors' Citations to Host Countries, by Host Countries' Technology Similarity. Upper panels: Baseline; Lower Panels: Interaction Effect.

(a) Brownfield Investors' Citations to Countries (b) Greenfield Investors' Citations to Countries



Note: This figure presents the coefficients β_h estimated by our specification in Eq. 2 for the presented citation measures. The sample used in this figure is the "World sample", which refers to the sample of firm-country pairs obtained matching PATSTAT company names with FDI data firm names following the procedure in in Subsection 2.3. We include firm-time and country-time fixed effects as controls, as well as three lags of the dependent variable in levels and its trend estimated using the change in 6 to 4 years before the event. We set the stabilization period to five periods. The treatments variable are the first FDI investment carried out by each firm in each destination country, as well as its interaction with our measure of similarity between technologies of the investors and host countries. The bottom bars represent the number of treated samples.

4.4 Other Results and Robustness

Patent Creation. We do not find conclusive evidence on the effects of FDI on patent creation, largely due to the difficulty in defining a tight identification strategy to investigate this question. Indeed, the identification strategy pursued in this paper relies on heterogeneity at the firm-destination level, which we cannot leverage for patent creation. The stock of patents for destination countries does not vary at the investing firm level, and would be absorbed by the destination-time fixed effects included in our specification. We therefore tried using our brownfield FDI data to evaluate the effect of acquisition on target firms. In this case, we found unreliable results, pointing to positive or negative effects on patent creation depending on the measure used for the patent stock. In particular, the results appeared negative for granted patents and positive for triadic ones, potentially indicating that FDI increases valuable inventions, but might overall lower patenting. However, the latter result may mechanically arise if target firms cease operations as a separate entity, which might often be the case. These results are available upon requests.

Robustness. The results of the baseline analysis remain robust under various conditions: (1) expanding the set of destination countries to all those available in PATSTAT; (2) limiting forward citations to five years after the application of a patent to address truncation issues for citations; (3) increasing the number of lags for the effects to stabilize, L, to 12; (4) applying the transformation $\log(1 + x)$ to cumulative citations instead of $\operatorname{asinh}(x)$; (5) using only citations between triadic patents; (6) using each investment–instead of just the first FDI–as a separate event; (7) grouping greenfield and brownfield FDI in a single "FDI investment" variable that does not distinguish different investment types. In all these cases, the results are qualitatively unchanged. With few exceptions, we find coefficients and implied percentage changes that are not statistically different from our baseline results. Finally, we check the robustness of our results for the US sample of firms to considering only patents originally assigned to US firms. In this case, the results for citations made by US firms to destination countries decrease in magnitude, but remain positive and statistically significant. The details and figures corresponding to each specification, the rationale behind these exercises and the implied percentage changes are reported in Appendix A.

5 Conclusion

In this paper, we explore the effects of brownfield and greenfield FDI on knowledge diffusion, measured by citation flows between investing firms and FDI destinations. Our results show that, following the first entry of an investor into an FDI destination, citations made by the destination to investors increase by 7.8%-10.6% on average, depending on the type of investment. In addition, investors increase their citations to FDI hosts by 4.5%-13%. These results are robust across different specifications. We also demonstrate that knowledge spillovers extend beyond targeted firms and industries, benefiting other sectors in destination countries. Finally, we highlight the role for the destinations' absorptive capacity in mediating technology diffusion. FDI hosts with larger pre-existing patent stocks and technologies more similar to those of investors benefit more from both brownfield and greenfield investments, and vice versa.

We view our study as a first exploration of the granular effects of capital flows on knowledge diffusion across countries. Our findings of the sizable knowledge flows induced by FDI suggest an important additional channel through which potential fragmentation of investment flows, amidst rising geopolitical tensions and attempts to strengthen national and supply chain security, could harm productivity and growth (IMF, 2023; Gopinath et al., 2024). Our work also underscores the critical role of destinations' capacity to reap the benefits from FDI activities and opens the door for further investigation into the channels through which knowledge transmits from investors to target entities. Specifically, we find that these heterogeneous effects of investments appear mediated by the *type* of investment itself (brownfield or greenfield). However, the lack of widely available data on FDI motives and characteristics limited our analysis of the reasons behind these differences. Accurate text analysis and classification of original FDI announcements present as a promising avenue to address this gap in knowledge.

In addition, our paper focuses on patent activity, which is available for the subset of countries that produce and report sufficient numbers of patents. We still know little about the diffusion of other forms of organizational knowledge, such as know-how. One question that warrants further exploration is whether acquired firms and industries increase their overall patenting activities following investments. Another related and challenging question for future research is the role of FDI in catalyzing the development of new economic sectors.

References

Aghion, P., A. Bergeaud, T. Gigout, M. Lequien, and M. Melitz (2023): "Exporting ideas: Knowledge flows from expanding trade in goods," .

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

Aiyar, S., D. Malacrino, and A. Presbitero (2023): "Investing in friends: The role of geopolitical alignment in fdi flows," Tech. rep., CEPR Discussion Paper.

Akcigit, U., S. T. Ates, J. Lerner, R. R. Townsend, and Y. Zhestkova (2024): "Fencing off Silicon Valley: Cross-border venture capital and technology spillovers," *Journal of Monetary Economics*, 141, 14–39.

Akcigit, U. and M. Melitz (2022): "International trade and innovation," in *Handbook of International Economics*, Elsevier, vol. 5, 377–404.

Alfaro, L. (2017): "Multinational Activity in Emerging Markets: How and When Does Foreign Direct Investment Promote Growth?" in *Geography, Location, and Strategy*, Emerald Publishing Limited, vol. 36 of *Advances in Strategic Management*, 429–462.

Alfaro, L., A. Chanda, S. Kalemli-Ozcan, and S. Sayek (2004): "FDI and economic growth: the role of local financial markets," *Journal of international economics*, 64, 89–112.

Antràs, P. and S. R. Yeaple (2014): "Multinational firms and the structure of international trade," *Handbook of international economics*, 4, 55–130.

Arora, A., S. Belenzon, and L. Sheer (2021a): "Knowledge spillovers and corporate investment in scientific research," *American Economic Review*, 111, 871–898.

——— (2021b): "Matching patents to Compustat firms, 1980–2015: Dynamic reassignment, name changes, and ownership structures," *Research Policy*, 50, 104217.

Autor, D., D. Dorn, G. H. Hanson, G. Pisano, and P. Shu (2020): "Foreign competition and domestic innovation: Evidence from US patents," *American Economic Review: Insights*, 2, 357–374.

Azoulay, P., C. Fons-Rosen, and J. S. G. Zivin (2019): "Does science advance one funeral at a time?" *American Economic Review*, 109, 2889–2920.

Belz, T., L. A. Robinson, M. Ruf, and C. Steffens (2013): "Tax avoidance as a driver of mergers and acquisitions," *Available at SSRN 2371706*.

Bena, J. and K. Li (2014): "Corporate innovations and mergers and acquisitions," *The Journal of Finance*, 69, 1923–1960.

Bergant, K., P. Mishra, and R. Rajan (2023): "Cross-border Spillovers: How US Financial Conditions affect M&As Around the World," Tech. rep., National Bureau of Economic Research.

Bernard, A. B., J. B. Jensen, and P. K. Schott (2009): "Importers, exporters and multinationals: a portrait of firms in the US that trade goods," in *Producer dynamics: New evidence from micro data*, University of Chicago Press, 513–552.

Blalock, G. and P. J. Gertler (2008): "Welfare gains from foreign direct investment through technology transfer to local suppliers," *Journal of international Economics*, 74, 402–421.

Bollaert, H. and M. Delanghe (2015): "Securities Data Company and Zephyr, data sources for M&A research," *Journal of Corporate Finance*, 33, 85–100.

Borensztein, E., J. De Gregorio, and J.-W. Lee (1998): "How does foreign direct investment affect economic growth?" *Journal of international Economics*, 45, 115–135.

Branstetter, L. (2006): "Is foreign direct investment a channel of knowledge spillovers? Evidence from Japan's FDI in the United States," *Journal of International economics*, 68, 325–344.

Chen, J. and J. Roth (2024): "Logs with zeros? Some problems and solutions," *The Quarterly Journal of Economics*, 139, 891–936.

Chen, S.-F. S. and M. Zeng (2004): "Japanese investors' choice of acquisitions vs. startups in the US: The role of reputation barriers and advertising outlays," *International Journal of Research in Marketing*, 21, 123–136.

Chen, Y., H. Jiang, Y. Liang, and S. Pan (2022): "The impact of foreign direct investment on innovation: evidence from patent filings and citations in China," *Journal of Comparative Economics*, 50, 917–945.

Cunningham, C., F. Ederer, and S. Ma (2021): "Killer acquisitions," *Journal of Political Economy*, 129, 649–702.

de Chaisemartin, C. and X. D'Haultfœuille (2020): "Two-Way Fixed Effects Estimators with Heterogeneous Treatment Effects," *American Economic Review*, 110, 2964–2996.

de Rassenfosse, G., H. Dernis, and G. Boedt (2014): "An Introduction to the Patstat Database with Example Queries," *Australian Economic Review*, 47, 395–408.

Deng, L., Y. Lu, and Y. Tang (2024): "Does FDI increase product innovation of domestic firms? Evidence from China," *Journal of Economic Behavior & Organization*, 222, 1–24.

Dikova, D. and K. Brouthers (2016): "International establishment mode choice: Past, present and future," *Management International Review*, 56, 489–530.

Dube, A., D. Girardi, and A. M. Taylor (2023): "A Local Projections Approach to Difference-in-Differences Event Studies," NBER Working Paper 31184.

Erel, I., Y. Jang, and M. S. Weisbach (2022): "Cross-border mergers and acquisitions," Tech. rep., National Bureau of Economic Research.

Glass, A. J. and K. Saggi (1998): "International technology transfer and the technology gap," *Journal of development economics*, 55, 369–398.

Globerman, S., A. Kokko, and F. Sjöholm (2000): "International technology diffusion: evidence from Swedish patent data," *Kyklos*, 53, 17–38.

Goldbach, S., A. J. Nagengast, E. Steinmüller, and G. Wamser (2019): "The effect of investing abroad on investment at home: On the role of technology, tax savings, and internal capital markets," *Journal of International Economics*, 116, 58–73.

Goodman-Bacon, A. (2021): "Difference-in-Differences with Variation in Treatment Timing," *Journal of Econometrics*, 225, 254–277.

Gopinath, G., P.-O. Gourinchas, A. Pescatori, and P. Topalova (2024): "Changing Global Linkages: A New Cold War?" *IMF Working Paper*, WP/24/76.

Griffith, R., R. Harrison, and J. Van Reenen (2006): "How special is the special relationship? Using the impact of US R&D spillovers on UK firms as a test of technology sourcing," *American Economic Review*, 96, 1859–1875.

Harrison, A. and A. Rodríguez-Clare (2010): "Trade, foreign investment, and industrial policy for developing countries," *Handbook of development economics*, 5, 4039–4214.

Haskel, J. E., S. C. Pereira, and M. J. Slaughter (2007): "Does inward foreign direct investment boost the productivity of domestic firms?" *The review of economics and statistics*, 89, 482–496.

IMF (2018): "Chapter 4: Is Productivity Growth Shared in a Globalized Economy?" in *World Economic Outlook, April 2018: Cyclical Upswing, Structural Change.*

——— (2023): "Chapter 4: Geoeconomic Fragmentation and Foreign Direct Investment," in *World Economic Outlook, April 2023: A Rocky Recovery.*

Jaffe, A. B., M. Trajtenberg, and R. Henderson (1993): "Geographic localization of knowledge spillovers as evidenced by patent citations," *the Quarterly journal of Economics*, 108, 577–598.

Javorcik, B. S. (2004): "Does foreign direct investment increase the productivity of domestic firms? In search of spillovers through backward linkages," *American economic review*, 94, 605–627.

Javorcik, B. S., A. Lo Turco, and D. Maggioni (2018): "New and improved: does FDI boost production complexity in host countries?" *The Economic Journal*, 128, 2507–2537.

Jiang, K., W. Keller, L. D. Qiu, and W. Ridley (2018): "International joint ventures and internal vs. external technology transfer: Evidence from China," Tech. rep., National Bureau of Economic Research.

Jinji, N., X. Zhang, and S. Haruna (2015): "Trade patterns and international technology spillovers: evidence from patent citations," *Review of World Economics*, 151, 635–658.

Keller, W. (1996): "Absorptive capacity: On the creation and acquisition of technology in development," *Journal of development economics*, 49, 199–227.

——— (2021): "Knowledge spillovers, trade, and FDI," Tech. rep., National Bureau of Economic Research.

Keller, W. and S. R. Yeaple (2009): "Multinational enterprises, international trade, and productivity growth: firm-level evidence from the United States," *The review of economics and statistics*, 91, 821–831.

LaBelle, J., I. Martinez-Zarzoso, A. M. Santacreu, and Y. V. Yotov (2023): "Cross-border Patenting, Globalization, and Development," .

Lybbert, T. J. and N. J. Zolas (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," *Research Policy*, 43, 530–542.

Maurseth, P. B. and B. Verspagen (2002): "Knowledge spillovers in Europe: a patent citations analysis," *Scandinavian Journal of Economics*, 104, 531–545.

Moretti, E. (2021): "The effect of high-tech clusters on the productivity of top inventors," *American Economic Review*, 111, 3328–3375.

Navaretti, G. B. and A. J. Venables (2004): *Multinational firms in the world economy*, Princeton University Press.

Neary, J. P. (2007): "Cross-border mergers as instruments of comparative advantage," *The Review of Economic Studies*, 74, 1229–1257.

Ni, B., H. Kato, and Y. Liu (2021): "Does It Matter Where You Invest? The Impact of FDI on Domestic Job Creation and Destruction," *The Impact of FDI on Domestic Job Creation and Destruction (January 24, 2021)*.

Phillips, G. M. and A. Zhdanov (2013): "R&D and the incentives from merger and acquisition activity," *The Review of Financial Studies*, 26, 34–78.

Rodriguez-Clare, A. (1996): "Multinationals, linkages, and economic development," *The American economic review*, 852–873.

Silva, J. S. and S. Tenreyro (2006): "The log of gravity," *The Review of Economics and statistics*, 88, 641–658.

Toews, G. and P.-L. Vézina (2022): "Resource discoveries, FDI bonanzas, and local multipliers: Evidence from Mozambique," *Review of Economics and Statistics*, 104, 1046–1058.

A Appendix: Robustness Results

Before describing each of our robustness exercises, it is worth recalling here our baseline findings in terms of percentage changes, and the panels that they correspond to in Figure III, whose structure is preserved in all the graphs reported below. We refer to the percent changes implied by 95% confidence intervals as "95% C.I. bounds" and report them in parentheses after midpoint estimates. Following the panel lettering, in Figure III we found that:

(a) Citations to brownfield investors increase on average by 7.8% (5.7 – 10%);

(b) Citations to greenfield investors increase on average by 10.6% (9 – 12.2%);

(c) Citations to countries made by brownfield investors increase by 10.8% (8.4 – 13.3%);

(d) Citations to countries made by greenfield investors increase by 13.4% (11.6 – 15.2%).

We will refer to these letters in the descriptions below for brevity. All the robustness exercises below adopt the same specification, sample definition, treatment and control variables, and number of stabilization lags, L, as the baseline, with the exception of the deviations that are expressly noted in each paragraph.

Robustness Including All Available Countries. In our baseline, we restrict attention to 60 top destinations and origin countries, representing 99.7% of global patenting activity. Figure XI includes all the destinations available in PATSTAT. As expected, the results are indistinguishable from the baseline, with midpoints corresponding to an increase in citations of: (a) 7.8%; (b) 9.9%; (c) 11.3%; and (d) 13.1%.

Robustness Fixing the Citation Window to 5 Years. The results presented in the main text may suffer from so-called "truncation bias." This bias arises from the fact that patent citations build up and increase as time passes. Further, the pattern of this increase may not be linear. As a result, using cumulative citations without restricting to a specific horizon after a patent is registered may provide inaccurate estimates, since for example older patents will have more time to accumulate citations. Another more relevant example of bias may arise if citations increase more markedly at specific horizons after their first application. To tackle this issue, Figure XII reports estimates that accumulate citations up to five years following the initial application. In addition, we also restrict the sample to 2015, to avoid truncation bias for patents that may be registered starting in 2015, and for which we would not have enough reliable citation data for at least five years. Indeed, we find patent data to be most reliable up to 2020, even though we employ the

Spring 2022 edition of PATSTAT. We find the following percentage increases: (a) 11.3% (b) 12% (c) 15.1% (d) 7.7%. All the first three cases report larger estimates than the baseline. In particular, estimates referring to citations to (a) and from (c) brownfield investors are even above the upper end of the 95% C.I. bounds from the baseline. By contrast, citations from greenfield investors (d) are significantly below the baseline results. This might point to the fact that in greenfield ventures, investors take more time to learn about newer local technologies than in the case of brownfield investments, where they acquire established firms and can benefit from the target firms' knowledge about latest developments. This would result in a lower reaction of truncated citation measures, which by construction capture citations made to more recent patents.

Robustness to Increasing L to 12 years. Figure XIII reports the results of our main specification when we extend the number of years required for treatment effects to stabilize, L, to 12. As discussed in the main text, this is the highest L that we believe reasonable to estimate effects up to 5 years after firms' investments. The estimates correspond to the following increases in citations: (a) 5.6%; (b) 12.5%; (c) 10.8% (d) 12.5%. These results are broadly consistent with our baseline, with results for citations made by investors, (c) and (d) within the baseline's 95% C.I. bounds. When it comes to citations received by investors, we find that result (a) sits just below the lower bound of 5.7% implied by our baseline, while (b) is slightly above the 12.2% we originally found.

Robustness to Using $\log(1+x)$ Transformation. In Figure XIV, we explore the robustness of our results to using the $\log(1+x)$ transformation to deal with zeros in our outcome variable, instead of the asinh transformation. We find midpoint estimates that are always larger than the baseline, and significantly so for panels (b) and (d), the cases related to greenfield investments. In particular, we find increases in citations of: (a) 9.4% ; (b) 12.7%; (c) 13.1%; (d) 16.2%. If anything, it appears that our transformation may understate the effects of FDI on patent citations.

Robustness to Using Triadic Patents. In the main text, we considered citations made by all granted patent families in destination countries to all granted patent families by origin firms. This already represents a correction for quality relative to using raw patent applications in destination countries. In this robustness, we further restrict both citing and cited patents to be triadic, that is registered at the USPTO, the JPO and the EPO. Triadic patent counts are generally considered to be a better gauge of innovation across countries (de Rassenfosse et al., 2014). Indeed, this registration would require a firm to incur the costs related to application, grant and patenting fees in multiple offices, signalling that the firm may attach a higher value to the invention in question. Therefore, triadic patents would be more valuable compared to non-triadic patents based on the applicants' *ex-ante* evaluation. Figure XV reports results corresponding to the following citation increases: (a) 6.6%; (b) 9.5%; (c) 10%; (d) 9.5%. In all cases, the midpoint estimates are lower than the baseline, and significantly so only in case (d), representing the increase in citations made by greenfield investors to destination countries. This finding is likely the result of a relative scarcity of triadic patents in greenfield destination countries.

Robustness to Using Each Investment as a Separate Event. In the main text, we presented the response to the the *first* FDI made by each firm. We did so under the presumption that following investment may be partly caused by the firm's first entry. In this case, considering each investment as a separate result would understate the effect of FDI investment. However, we only have a limited time coverage for our events, so only some of FDI included in our dataset will genuinely be first entries, making the approach in our baseline potentially inconsistent. At the same time, it is worth noting that the majority of firms does not carry out more than one investment in each destination over the time frame considered, so we expect this robustness not to impact our findings *a priori*. Consistent with this observation, Figure XVI displays a larger number of events (about 12% more for brownfield investments and about 20% more for greenfield) and estimated percentage increases in citations that are not significantly different from baseline, and are quite close to the midpoints we found there. In particular, the numbers for each panel are: (a) 6.8%; (b) 9.9%; (c) 9.5%; (d) 13.2%.

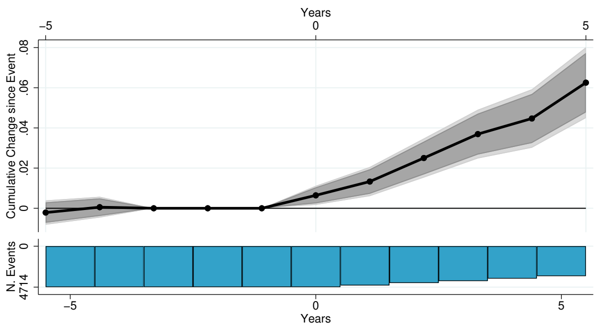
Robustness to Grouping Greenfield and Brownfield FDI. The LP-DiD methodology rests on the assumption that treatment effects stabilize after L periods. In our baseline, we considered greenfield and brownfield FDI as separate events. This may create a bias in our results, as we have shown that both brownfield and greenfield FDI cause increased citations to investing firms. As a result, firm-destination pairs that undergo both types of investments may be incorrectly assigned to clean treatment or control groups. For example, suppose that we are considering the effects of greenfield FDI. In this case, the clean control group includes firm-country pairs for which the treatment effects of previous greenfield FDI investments have stabilized. However, this group may include firm-country pairs with brownfield FDI events with effects that have not stabilized, potentially biasing the results if, e.g., brownfield FDI cause subsequent greenfield FDI. To address this concern, we now replace our treatment variable with a dummy that denotes *any* FDI occurring, and impose that no other FDI has occurred in the previous five years to allow units in our treatment or control group. Further, like in the previous robustness exercise, we consider each FDI as a separate event. Figure XVII shows that the resulting number of events $(29,573)$ is quite close to the sum of brownfield and greenfield FDI events taken separately and reported below the panels in Figure XVI $(30,845)$.

This suggests that in most cases firms engage in only one type of investment at a time. We find that citations from destinations to FDI investors increase by 9.4% following and investment, and investors' citation to target countries rise 12.9%. These numbers are quite close to the simple average of estimates (a) an (b)–9.1%–and (c) and (d)–12.1%, respectively. Both estimates are also comfortably withing the 95% C.I. bounds from the baseline.

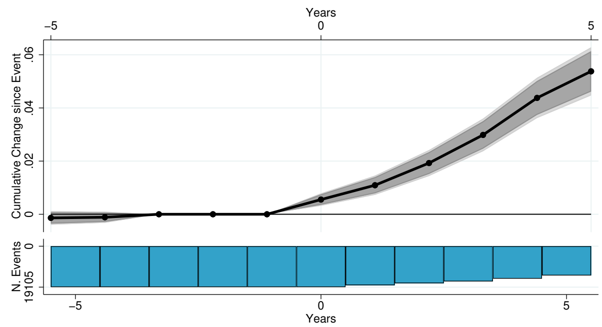
Robustness to Using Only Patents Originally Assigned to US Firms in the US Sample. We further consider the robustness of our results for the US sample to limiting our citation measure to only include patents that are originally assigned to US firms. Indeed, as noted in the main text, the Arora et al. (2021b) dataset tracks the ownership of patents over time, which means that the ultimate owner of each patent is not fixed, but may change over time if the patent is transferred. This may threaten our interpretation of increased citations as knowledge spillovers if such citations arise from the reassignment of patents in target countries to the investing firms. In this case, increased citations to the investing firm may result mechanically from patents of target firms being reassigned to their new owner. In the case of greenfield investment, a similar issue may arise if investors simultaneously carry out other brownfield FDI activity–or otherwise acquire patents–in the destination country. It is important to note that including only patents originally assigned to US firms should provide us with highly conservative estimates, because this exercise also excludes patent that the investing firms may have acquired from other destinations. Our estimates from Figure XVIII imply the following increases in lifetime citations: (a) 5.7%; (b) 9.1%; (c) 3.5%; (d) 5.4%. In this case, the numbers should be compared to those implied by Figure IV (with C.I. bounds in parentheses): (a) 8.4% ($4.9 - 11.9$%); (b) 10% ($7.2 - 12.7$%); (c) 12.7% ($7 - 18.6$%), (d) 11.1% ($6.2 - 16.1$%). Considering these estimates and the substantially higher uncertainty surrounding them, we see that only the results on citations flowing from FDI investors to destination countries are substantially affected by the restriction to original US owners. The lower coefficients in (c) and (d) compared to the baseline signal that increased citations by acquiring firms in the Arora et al. (2021b) dataset partly stem from the mechanical reassignment of patents from destination countries or other countries to investing U.S. firms.

Figure XI: Impacts of FDI on Citations, World Sample, Using All Available Destinations

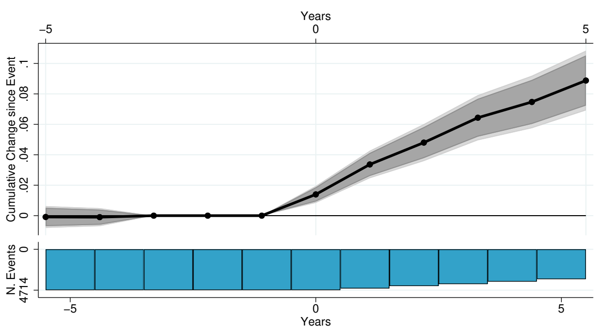
(a) Country's Citations to Brownfield Investors



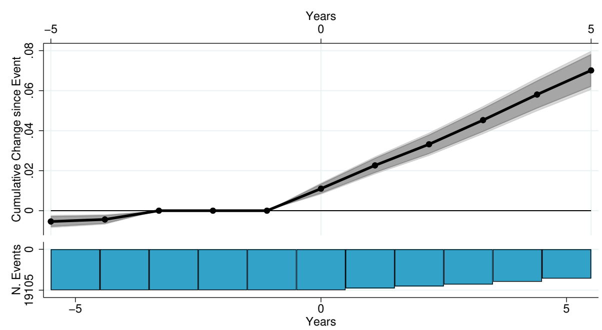
(b) Country's Citations to Greenfield Investors



(c) Brownfield Investors' Citations to Country



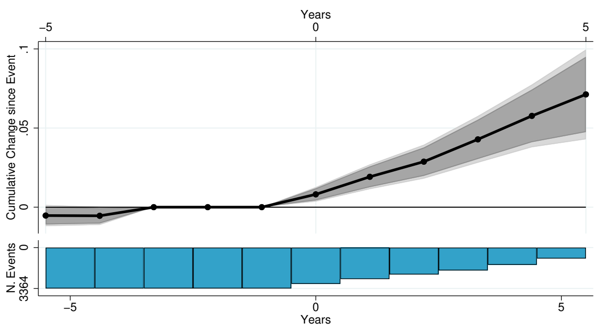
(d) Greenfield Investors' Citations to Country



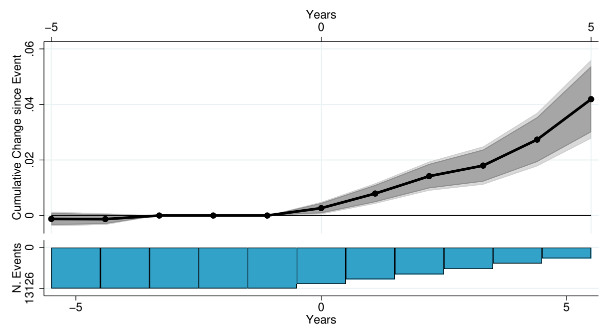
Note: This figure presents the coefficients β_h estimated by our specification in Eq. 2 for the presented citation measures. Controls and treatment are unchanged relative to Figure III. We extend the sample to cover all destinations available in PATSTAT.

Figure XII: Impacts of FDI on Citations, World Sample, Using Citations Within 5 Years of Patent Registration

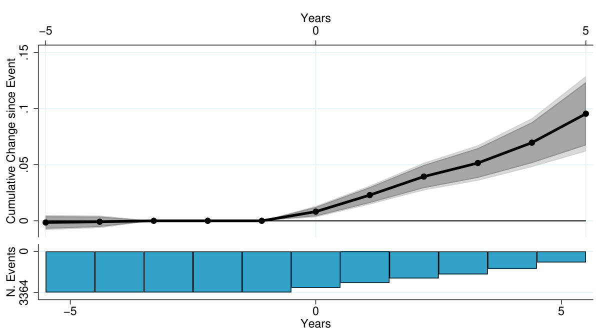
(a) Country's Citations to Brownfield Investors



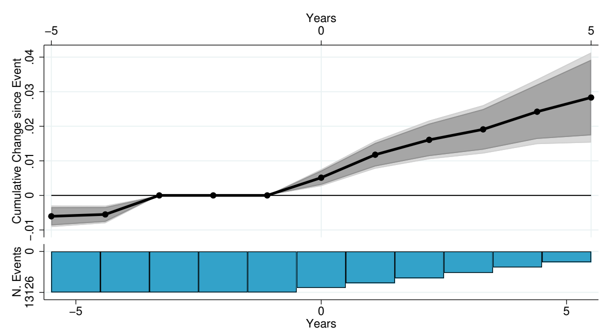
(b) Country's Citations to Greenfield Investors



(c) Brownfield Investors' Citations to Country



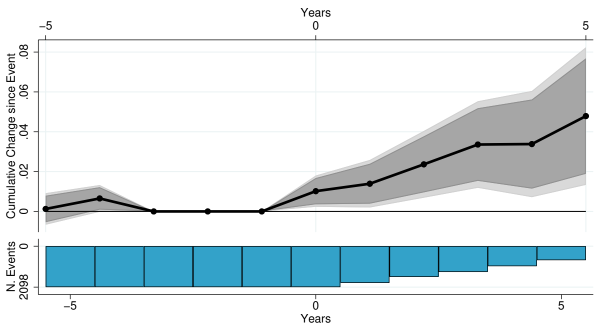
(d) Greenfield Investors' Citations to Country



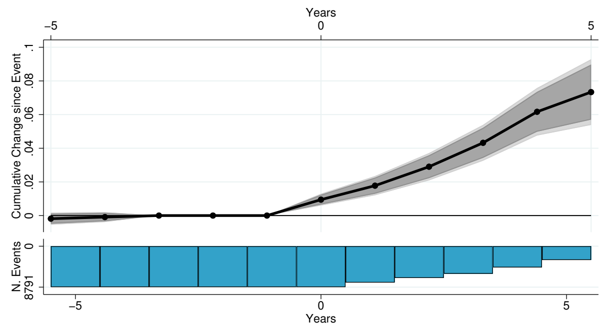
Note: This figure presents the coefficients β_h estimated by our specification in Eq. 2 for the presented citation measures. Controls and treatment are unchanged relative to Figure III. We now restrict only to citations received within 5 years of patent registration and shorten the sample to cover 2003-2015 to reduce truncation bias.

Figure XIII: Impacts of FDI on Citations, World Sample, Using 12 Stabilization Periods

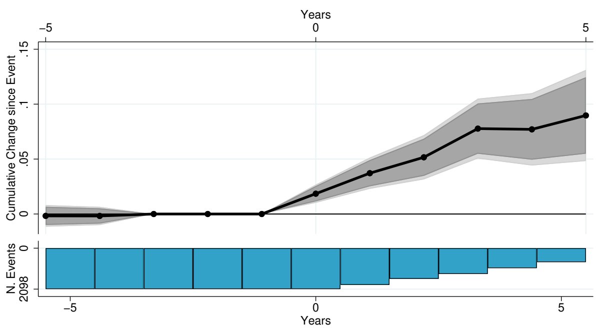
(a) Country's Citations to Brownfield Investors



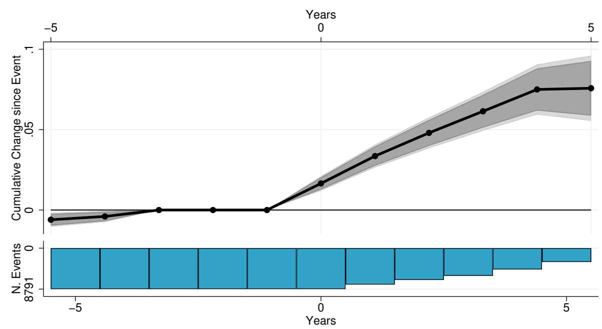
(b) Country's Citations to Greenfield Investors



(c) Brownfield Investors' Citations to Country



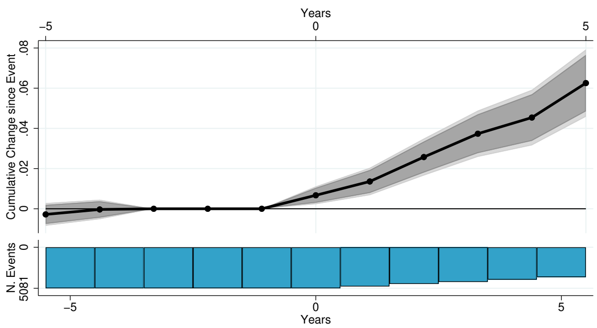
(d) Greenfield Investors' Citations to Country



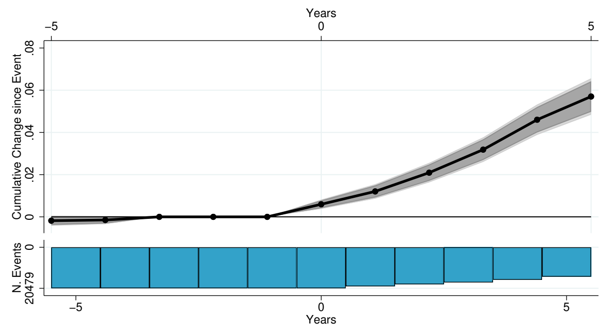
Note: This figure presents the coefficients β_h estimated by our specification in Eq. 2 for the presented citation measures. Controls, sample selection and treatment are unchanged relative to Figure III. We now set $L = 12$ instead of $L = 5$.

Figure XIV: Impacts of FDI on Citations, World Sample, Using the $\log(1+x)$ Transformation

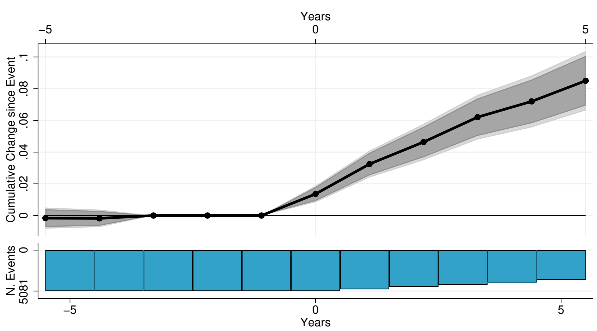
(a) Country's Citations to Brownfield Investors



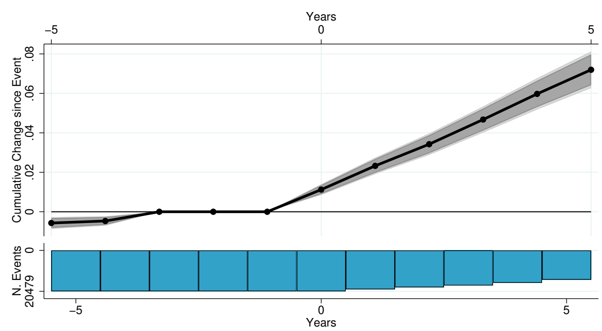
(b) Country's Citations to Greenfield Investors



(c) Brownfield Investors' Citations to Country



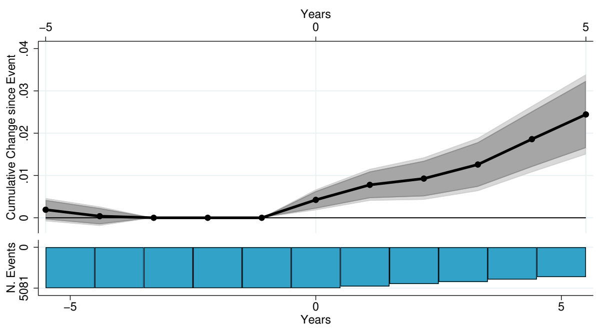
(d) Greenfield Investors' Citations to Country



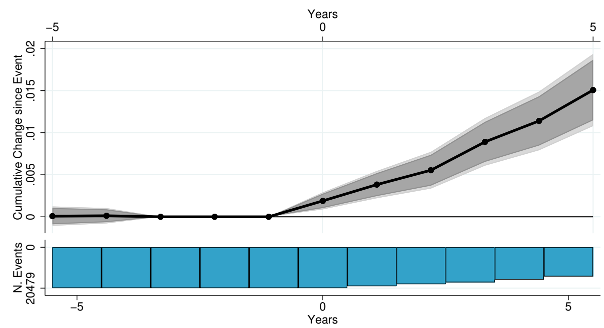
Note: This figure presents the coefficients β_h estimated by our specification in Eq. 2 for the presented citation measures. Controls and sample selection are unchanged relative to Figure III. We now use as outcome the the logarithm of one plus the number of cumulative citations since 1995 instead of the asinh transformation.

Figure XV: Impacts of FDI on Citations, World Sample, Only Citations Between Triadic Patents

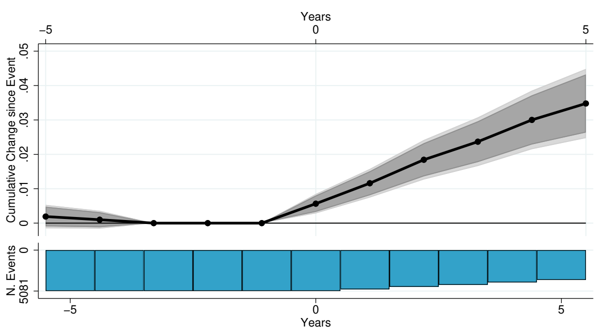
(a) Country's Citations to Brownfield Investors



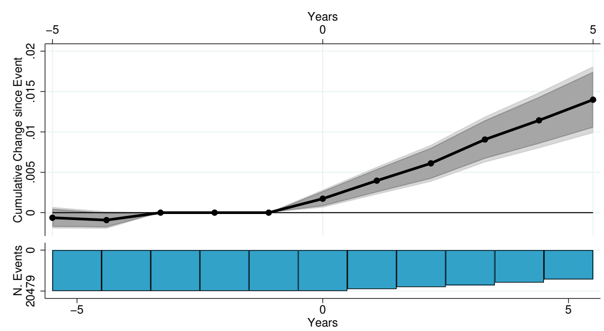
(b) Country's Citations to Greenfield Investors



(c) Brownfield Investors' Citations to Country



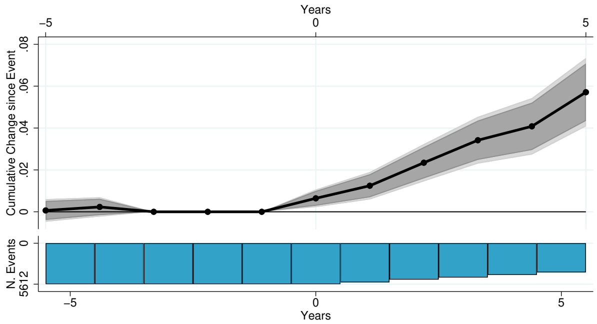
(d) Greenfield Investors' Citations to Country



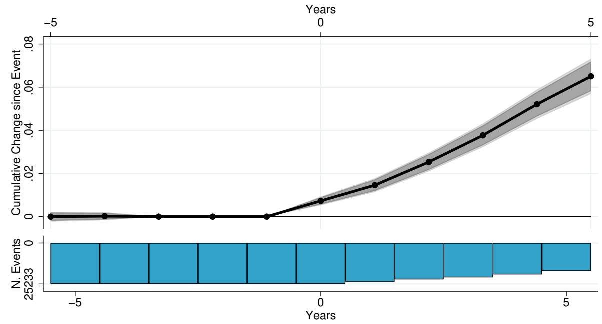
Note: This figure presents the coefficients β_h estimated by our specification in Eq. 2 for the presented citation measures. Controls, sample selection and treatment are unchanged relative to Figure III. We now restrict citations to those made and received by triadic patents, that is, patents simultaneously registered at the USPTO, EPO, and JPO.

Figure XVI: Impacts of FDI on Citations, World Sample, Each Investment is a Separate Event

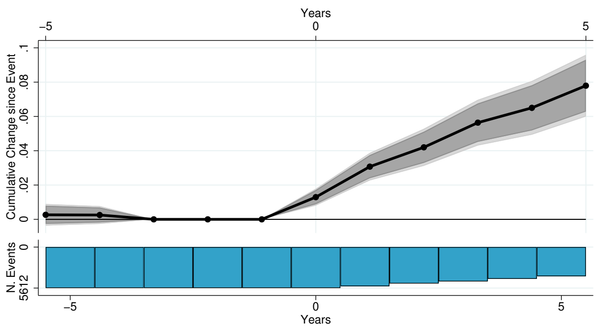
(a) Country's Citations to Brownfield Investors



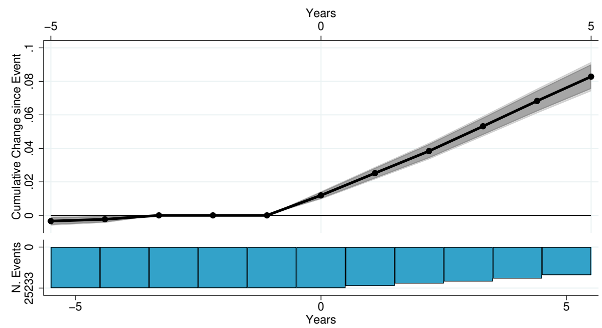
(b) Country's Citations to Greenfield Investors



(c) Brownfield Investors' Citations to Country



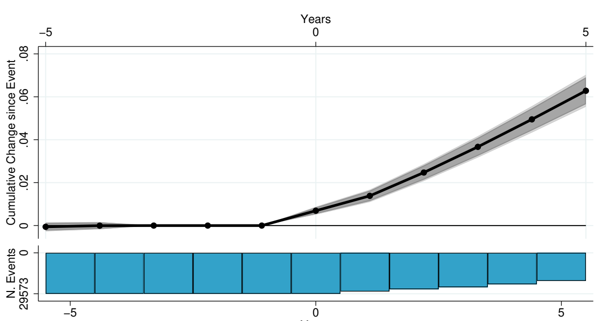
(d) Greenfield Investors' Citations to Country



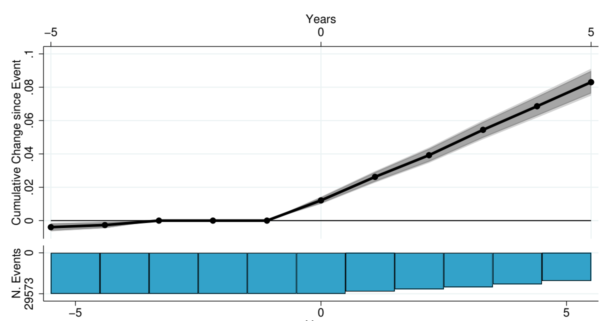
Note: This figure presents the coefficients β_h estimated by our specification in Eq. 2 for the presented citation measures. Controls and sample selection are unchanged relative to Figure III. We now use as treatment *any* FDI, not just the first investment observed in our sample period.

Figure XVII: Impacts of FDI on Citations, World Sample, Using a Single FDI Measure and Considering Each Investment as a Separate Event

(a) Country's Citations to FDI Investors



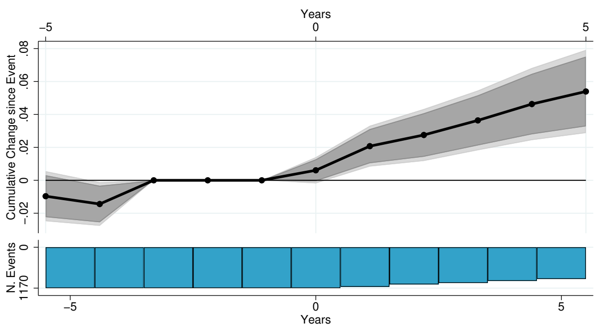
(b) FDI Investors' Citations to Country



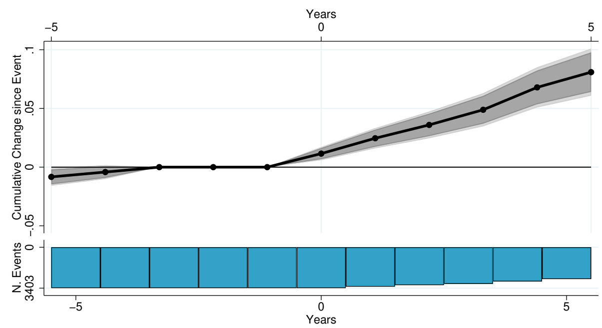
Note: This figure presents the coefficients β_h estimated by our specification in Eq. 2 for the presented citation measures. Controls and sample selection are unchanged relative to Figure III. We now use as treatment a dummy denoting an FDI by firm i, be it brownfield or greenfield, instead of defining separate treatments. We also use as treatment *any* FDI, not just the first investment observed in our sample period.

Figure XVIII: Impacts of FDI on Citations, US Sample, Patents Originally Assigned to US Entities.

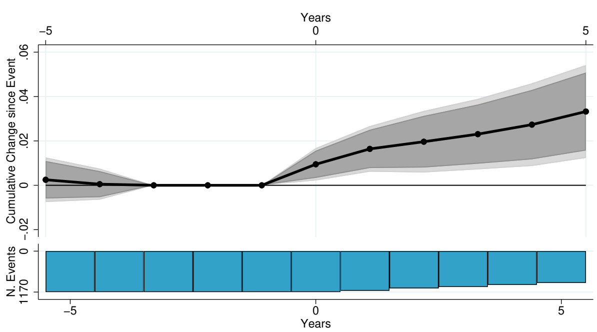
(a) Country's Citations to Brownfield Investors



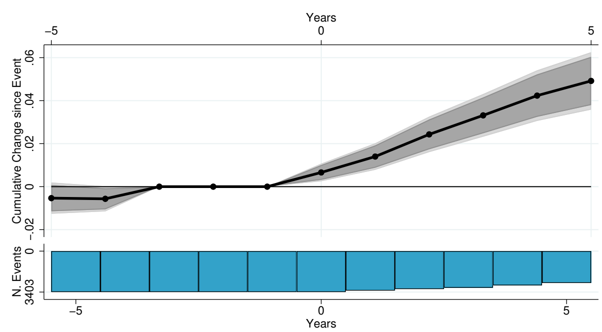
(b) Country's Citations to Greenfield Investors



(c) Brownfield Investors' Citations to Country



(d) Greenfield Investors' Citations to Country



Note: This Figure presents the coefficients β_h estimated by our specification in Eq. 2 for the presented citation measures. The specification and sample selection correspond to Figure IV. The only difference is that we replace the outcome variable with citations only directed or made by patents that were originally assigned to US ultimate owners.

PUBLICATIONS

Knowledge Diffusion Through FDI: Worldwide Firm-Level Evidence

Working Paper No. WP/2024/152