Knowledge Diffusion Through FDI: Worldwide Firm-level Evidence*

JaeBin Ahn Chan Kim Nan Li Andrea Manera

September 20, 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 global utility patent data with project-level information on greenfield FDI and cross-border mergers and acquisitions (M&A), covering 60 countries over two decades. Applying a new local projection difference-in-differences methodology, we find that FDI significantly enhances knowledge flows both from and to investing firms, with stronger effects observed for greenfield FDI compared to M&A. Our results also reveal substantial heterogeneity in FDI spillovers: host countries with higher innovation capacity or greater technological similarity to the investing firm benefit disproportionately, while countries with lower absorptive capacity capture fewer gains. In addition, we also uncover knowledge spillovers beyond the targeted firms and industries, particularly in sectors closely connected in the technology space.

Keywords: Greenfield FDI, Brownfield FDI, cross-border M&A, Inward FDI, Outward FDI, Knowledge spillover, Patent citation, LP-DiD

JEL Codes: F21, F23, F63, O31, O33

*JaeBin Ahn: International Monetary Fund, email: jahn@imf.org; Chan Kim: University of Maryland, email: chankim@umd.edu; Nan Li: International Monetary Fund, email: nli@imf.org; Andrea Manera: International Monetary Fund, email: 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.

1 Introduction

The global expansion of Foreign Direct Investment (FDI) has coincided with a surge in cross-border patent citations, suggesting an increase in international knowledge flows that are critical for economic growth and innovation. FDI has long been regarded as a potential key channel for facilitating these knowledge spillovers, particularly between investment firms and host countries (Keller, 2022). However, significant gaps remain in understanding the extent to which FDI fosters knowledge spillovers across diverse countries and industries. Robust causal evidence on the spillover effect of FDI is largely confined to specific contexts or country cases (e.g. Branstetter, 2006; Akcigit et al., 2024; Jiang et al., 2024), limiting the scope for broader generalization. Moreover, the role of absorptive capacity—defined as a host country's ability to absorb and utilize foreign knowledge—has not been thoroughly explored in a global context. There is also scant evidence regarding the comparative effectiveness of different modes of FDI—greenfield investments versus mergers and acquisitions (M&A)—in affecting knowledge flows. As rising geopolitical tensions threaten to disrupt cross-border investment, addressing these questions has become increasingly important (IMF, 2023; Gopinath et al., 2024).

This paper aims to fill these gaps by examining global firm-level data to uncover the dynamic effects of FDI on bi-directional knowledge flows. We first construct a novel dataset that combines the universe of project-level data on greenfield investments (hereafter, greenfield FDI) from fDi Markets and cross-border M&A (hereafter, brownfield FDI) from Refinitiv Eikon, with worldwide patent citation data from PATSTAT. Our dataset includes FDI transactions and bi-directional patent citations involving 60 countries over two decades, providing a high level of detail and extensive global coverage. To ensure robust identification, we employ the newly developed method of local projection difference-in-differences (LP-DiD) by Dube et al. (2023), which allows us to estimate the dynamic effects of FDI over time while addressing the biases inherent in staggered DiD models. Using this data and method, we analyze how citations to and from the investing firm evolve following its initial FDI in a country, whether these citation flows depend on the stock and content of existing knowledge in the host country, and whether the knowledge spillovers are confined to targeted industries or extend beyond them.

Our analysis contributes to the existing literature in three main ways. First, we strengthen the causal identification of FDI's impact on knowledge spillovers by applying 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 traditional staggered DiD analyses. In staggered DiD settings, treatment events—such as FDI—occurs at different times across projects, potentially yielding differential effects over time. Traditional DiD methods fall short in capturing these variations, potentially introducing bias into the results (Goodman-Bacon, 2021). The LP-DiD approach overcomes this by using local projections to estimate the dynamic effects of FDI, while employing a "clean control" condition to define appropriate treated and control groups.

Our global firm-level dataset enhances identification by allowing us to account for unobserved time-varying factors at both the host-country(-industry) and firm levels, which country-specific studies often struggle to control for. By incorporating an extensive set of fixed effects, we effectively account for variations in innovation rates, citation propensities, patent activities, and other unobserved shocks across countries, industries, and firms. This approach provides more accurate estimates of the causal effects of FDI on knowledge spillovers.

Second, 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 with different degrees of innovation capacity. This allows us to challenge the common perception that FDI promotes convergence between rich and poor countries by facilitating the transfer of advanced technologies. Moreover, the data enables examination of not only the spillovers from inward FDI (spillovers that host countries benefit from foreign investments) but also outward FDI (the reciprocal gains experienced by the investing firms). The latter remains relatively underexplored in the literature, with a few exceptions providing country-specific evidence (Branstetter, 2006; Griffith et al., 2006; Goldbach et al., 2019).

Third, the inclusion of both greenfield and brownfield FDI allows for a direct comparison of these distinct channels through which each type of investment influences knowledge flows. Both types of investments have the potential to facilitate knowledge flows within and beyond the boundaries of the affiliates and their industries (e.g., Javorcik, 2004; Keller and Yeaple, 2013; Phillips and Zhdanov, 2013; Bena and Li, 2014). However, greenfield investments often involve new operations, which are typically associated with higher levels of technology transfer, while brownfield FDI through mergers and acquisitions may yield different spillover dynamics due to the absorption of existing firms and assets (Neary, 2007; Antràs and Yeaple, 2014).

Our main findings are summarized as follows. First, we find that FDI significantly increases both outward and inward knowledge flows between investing firms and hosting countries. When comparing different modes of FDI, patent citations tend to increase more following greenfield investments than brownfield investments. This is likely because greenfield FDI typically involves building new capabilities from scratch, which requires greater technology transfers from the investing firm, while brownfield deals often focus more on market or asset acquisition. Nevertheless, in both cases, local firms may benefit by learning from foreign entrants by observing, imitating, and reverse-engineering their products, or even through employee transitions, generating positive spillovers at the country level. Specifically, Compared to the control group, patent citations from the host country to the investing firm rise by an average of 7.8% following brownfield FDI and 10.6% following greenfield FDI over the five years after the initial investment. Similarly, reverse citations—those made by investors to the host country—rise by an averages of 10.8% and 13.4% for brownfield and greenfield FDI, respectively. These baseline results are robust across a variety of alternative specifications and measures.

These findings complement earlier works that associate FDI with cross-border citations in specific country contexts. For example, Branstetter (2006) finds that Japanese affiliates in the U.S. increased patent citations both to and from their parent firms, with a particularly pronounced effect observed among greenfield affiliates compared to acquired firms. In contrast, Chen et al. (2022) did not find evidence of intra-industry knowledge spillovers in China. Most recently, Akcigit et al. (2024) find that the foreign investors of U.S. startups increased their citations to these startups post-investment, suggesting strong spillovers to foreign investors. While country-specific studies offer valuable insights, generalizing their findings can be challenges. By analyzing FDI globally using a comprehensive firm-level database, we address this limitation.

Second, while FDI is often regarded as a potential mechanism for fostering convergence between advanced and developing economies by transferring knowledge, our findings suggest that the reality is more nuanced. We find substantial heterogeneity in the degree of knowledge spillovers across countries. Knowledge spillovers are notably stronger in countries with higher level of pre-existing innovation capacity. Specifically, host countries in the top 10% of patent stocks receives increased spillovers that are two to ten times greater than those in the bottom 90%, providing robust evidence on the role of "absorptive capacities" (Keller, 1996). This implies that FDI may not necessarily promote convergence as often hypothesized, but may reinforce existing inequalities in innovation capacity between advanced and developing economies.

Adding to that literature, we further show that technological proximity (measured using the framework of Jaffe, 1986; Bloom et al., 2013) between investing firms and host countries significantly increases knowledge spillovers, with notable increases in citations when their technology compositions are similar. In fact, when technological proximity is low, knowledge spillovers appear to be insignificant. This finding adds to the literature documenting the uneven effect of inward FDI across countries, which often points to the importance of human capital and financial development (Borensztein et al., 1998; Alfaro et al., 2004). This suggests that policies aimed at improving absorptive capacity by investing in education, technological infrastructure and strengthening institutions are critical for benefiting more fully from FDI. Targeted FDI strategies, where policies focus on sectors where the country has some existing capabilities, would also allow domestic firms to better learn from foreign investors.

Finally, knowledge spillovers extend beyond directly targeted firms and industries of the foreign affiliates but extend more broadly across related sectors, particularly those that are technologically more closely linked to the investing firm. The industry-level estimation results suggest that the spillover effect outside the targeted sector can be, on average, as large as that found within the targeted sector.

This paper adds to the extensive literature evaluating the externalities of FDI, including the spillovers through backward and forward linkages and the pro-competitive effect (Rodriguez-Clare, 1996; Glass and Saggi, 1998; Navaretti and Venables, 2004). Most studies of knowledge spillover effects of FDI use changes in firm or industry-level productivity to indirectly infer knowledge spillovers (Aitken and Harrison, 1999; Javorcik, 2004; Haskel et al., 2007; Keller and Yeaple, 2009; Harrison and Rodríguez-Clare, 2010). Our paper directly traces knowledge flows using citations between patents (*à la* Aghion et al., 2023; Cai et al., 2022; 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. More importantly, much of the exiting literature relies on traditional DiD methodologies. Our application of the LP-DID represents a significant methodological advancement. This allows us to address common issues such as bias from staggered treatment and heterogeneous dynamic effects, which often complicate conventional DiD estimates. The LP-DiD methodology not only enhances the causal identification of FDI's impact on knowledge flows but also provides a more accurate dynamic description of these spillovers over time.

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, discusses its suitability and advantages relative to other methods and our choices of empirical specifications. Section 4 reports the results. After describing the baseline effects of brownfield and greenfield FDI on citations, we move to analyze the heterogeneity of our findings depending on countries' absorptive capacity and also present industry-level results. 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.[1] 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 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.

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).[2] 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).

2.2 Patent 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 granted and applied patents in 90 patent-issuing authorities worldwide, including those in non-European countries. These patent-issuing authorities submit data on patent applications to the EPO, which

[1]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).

[2]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.

extracts information on the applicants based on the self-reported data provided to the patent authorities. The EPO then uses a "name disambiguation" procedure to consolidate various versions of an applicant's name that refer to the same entity and assign a unique ID. For the purposes of our analysis, we focus on applicants classified as companies, excluding those identified as 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.

We then count citations generated or received by a patenting firm to or from a given country. As common in the literature, we measure citations at the level of a "patent family" in order to avoid issues of double counting.[3] In addition, to measure knowledge spillover as accurately as possible, we only include citations made by the patent applicants, ensuring our analysis reflects the knowledge base of inventors at the time of invention. This excludes any citations added later by patent examiners, as those may reflect technological relevance but not direct knowledge spillover.

When counting patents and citations, we employ fractional counting, dividing each patent family proportionally across applicants and across different sectors, originally classified based on International Patent Classification (IPC). Specifically, for a patent applicable to N IPC codes and belonging to M companies, $1/(NM)$ patents are assigned to each firm-IPC observation. We obtain counts at the firm-ISIC code level applying the probabilistic crosswalk of Lybbert and Zolas (2014) to firm-IPC counts. For higher levels of aggregations, we sum the relevant fractional counts. In addition, in some analysis we use triadic patents—patents registered in the US Patent Office (USPTO), Japanese Patent Office, and EPO—as indication of high-quality patents (de Rassenfosse et al., 2014).

2.3 Linking Citations and FDI Datasets

To analyze citations related to specific firms, we first need to align firm names from FDI datasets with those in PATSTAT. To manage this, we standardize the names following established procedures described by Arora et al. (2021b). In addition, we restrict the PATSTAT sample to include only firms with five or more patents, which accounts for majority of total patent counts in the data. We then apply two different string similarity metrics and retain firms whose matches exceed a certain similarity score.[4] In addition, we conduct a manual check to filter out mismatches. We refer to

[3]A patent family groups patent applications that refer to the same invention, which can appear at different times in different patent offices. Counting patent families is the preferred approach to measure inventions in the patent literature (see de Rassenfosse et al., 2014).

[4]The two similarity metrics are (1) cosine similarity with term-frequency-inverse document frequency (TF-IDF) weighting at the tri-gram character level and (2) a matching algorithm developed by the Dutch central bank to identify potential matches between datasets (see https://github.com/DeNederlandscheBank /name_matching for more details.) We kept the matches with a similarity score over the threshold of 0.7 (out of 1) from either metric.

the resulting matches as the "world sample", indicating their comprehensive coverage of corporate entities globally across the datasets.

For robustness and for some of the analyses where we need to reduce dimensionality and focus on a smaller sample, we zoom in on a separate sample specifically for U.S. investing firms by utilizing the well-established matches from previous studies. This sample builds on the firm name matches between PATSTAT and U.S. Compustat firms developed by Arora et al. (2021b). We then match this set of U.S. firms with FDI events from fDi Markets and Refinitiv Eikon. In addition to applying the two string similarity metrics mentioned earlier, we refine the matches using online search similarity. Specifically, we confirm that each pair of firm names shares at least one common webpage among the first ten results from Bing's search API. This procedure enhances the accuracy for both patent and FDI information for U.S. firms, which we refer to as the "U.S. 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 destination countries based on the number of granted patents from 2003 to 2022, representing over 99.9% of all patents. Overall, our world sample includes 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 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 abroad.

2.4 A First Look at the Relationship between FDI and Patent Citations

We begin by examining firm-destination country level matched data to identify stylized facts that motivate our empirical investigation into the knowledge diffusion channel of FDI spillovers.[5] Specifically, we investigate whether there exists any relationship between FDI and patent citations after controlling for firm-year, destination country-year, and bilateral country-pair specific factors that

[5] A time-series description of the patterns of FDI and patent citations is provided in the Appendix A.

could independently and simultaneously influence FDI and citation decisions. This involves residualizing FDI and patent citation measures using the Poisson Pseudo-Maximum Likelihood (PPML) estimator *à la* Silva and Tenreyro (2006):

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

where Y_{ioct} is the total number of FDI transactions by investor firm i from a source country o to a destination country c in year t for either brownfield or greenfield FDI measure. For citation measures, we consider either the total number of patent citation counts made from country c to firm i in country o in year t (i.e., knowledge flow from firm i to c) or the total number of patent citation counts made by firm i in country o in year t to patents in country c, which we refer to as reverse citations. Explanatory variables are fixed effects terms: δ_{it} denotes firm-year fixed effects that can control for the tendency of more innovative firms to invest abroad more and have their patents cited more frequently; δ_{ct} is destination country-year fixed effects that can account for factors such as the tendency of larger countries to innovate more and attract more FDI.; δ_{oc} is country-pair fixed effects that should absorb bilateral country-level variables, such as geographical or cultural distance, to take into account the fact that country pairs engaging in more FDI activity tend to exchange patent citations more intensively as well.

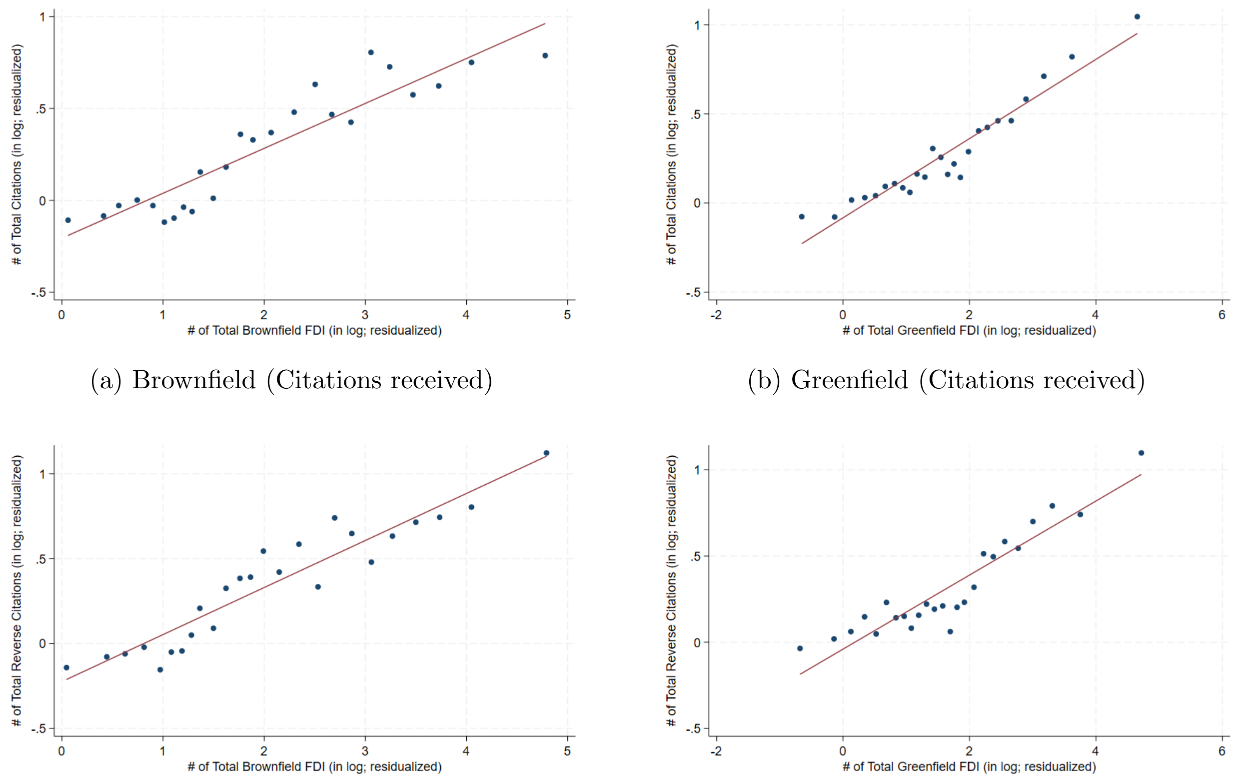
Figure 1 plots binned scatters of residualized FDI and patent citations to illustrate their relationship at the firm-destination country level. The top row measures patent citations as a total number of citations made by new patents from a destination country to patents belonging to an investor firm. 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 receive citations 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 positive correlation strongly suggests the presence of underlying factors linking FDI and patent citations, which persist beyond firm-, country-, or country pair-specific confounding effects. This observation highly motivates us to investigate 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, as recently proposed by Dube et al. (2023). This method combines local projec-

Figure 1: 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.

tions 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. The new method offers several advantages over traditional DiD and LP analyses. First, unlike DiD, it allows for the control for pre-treatment variables, particularly in controlling for differences in observed trends in the dependent variable before the event—something that is not feasible in a standard DiD setting.[6] Second, compared to both LP and DiD specifications, the LP-DiD methodology reduces bias from heterogeneous treatment effects across different treatment groups by appropriately selecting treatment and control groups.

[6] A standard DiD cannot account for pre-treatment trends in dependent variables when units are treated at different times. Since it involves running a single regression, the only option is to control for trends in specific time periods.

This advantage is particularly relevant for our study, since it addresses challenges in settings with staggered and repeated treatments (de Chaisemartin and D'Haultfœuille, 2020; Goodman-Bacon, 2021), where multiple acquisitions or greenfield investments take place in the same countries (and sectors) at various times.

To investigate the heterogeneous effects of different investment types, we estimate separate regressions for brownfield and greenfield FDI. Our baseline specification examines how initial FDI by firm i in country c affects the citations between patents from country c and those 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:

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

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; $\mathbf{x}_{ic,t-1}$ represents other controls which are explained below. In this regression, the coefficient of primary interest is β_h that 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 (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. 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).

To avoid this, we exclude any observations that were treated in the period from $t - L$ to $t + h$ to prevent contamination of the control group.[7] 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.

As with any DiD analysis, the validity of our causal interpretation depends on the parallel trends assumption. This assumption posits that, in the absence of the treatment, the average outcomes for the treated and control groups would have followed a similar trajectory over time, once other independent variables are controlled for. Therefore, any observed divergence in outcomes after the treatment can be attributed to the effect of the treatment itself, rather than pre-existing trends. In our context, this assumption would be compromised if citation trends across different firm-country pairs diverged for reasons not captured by our empirical model.

Our choice of controls and unit of analysis specifically aim 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 typically more productive and innovative, which could naturally lead to higher citation counts. Similarly, host country-time fixed effects are included to control for the possibility that these firms may target more innovative countries, which tend to both generate and receive more citations. By analyzing the impact at the host-country or host-industry-country level, rather than at the individual subsidiary level, we reduce the risk of bias from investors selectively acquiring or investing in particularly innovative affiliates. In addition, in the following section, we further mitigate this challenge by excluding citations from subsidiaries or acquired affiliates and by focusing on industries that are not directly affected by 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 a firm's initial entry 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 specific firm-country can experience multiple events of treatments (i.e. when treatment is "non-absorbing").[8] Alternatively, one can define the event as *any* FDI, not

[7] Imposing the condition that controls have not been treated between $t - 1$ and $t + h$ is not a key requisite of the specification. In fact, we obtained quantitatively similar results without this restriction.

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

just the initial one, which we explore as a robustness check in Appendix B. In our baseline, we do not consider subsequent FDIs as separate events, operating under the assumption that they are likely connected to the initial entry. Under this assumption, the effect on citations should still be attributed to the initial entry.[9] As shown in Appendix B, the results indeed do not significantly differ from those based on our baseline definition of the treatment.

Dependent Variable: The Inverse Hyperbolic Sine Transformation of Cumulative Citations. Throughout the paper, we use Asinh (inverse hyperbolic sine) transformation of cumulative citations (from 1995 onwards) as the outcome variable. The Asinh transformation is commonly used as an alternative to the logarithmic transformation to reduce the impact of skewness in the distribution of observations, while also accommodating data with zero values.[10] However, unlike the log transformation, the interpretation of coefficients derived from the Asinh transformation as percentage changes depends on the specific value around which effects are computed.[11] For clarity, we will convert these coefficients back to percentage changes throughout the text.

In addition, we examine the robustness of our results to focusing on citations between *triadic* patents—those registered simultaneously with the USPTO, JPO and EPO. To address potential truncation bias in citation counts towards the end of our sample, we also run an alternative specification that includes only citations within five years after the initial patent application, excluding the final five years of data. These robustness checks yield broadly consistent results.

Choice of Stabilization Lags. Our baseline estimation assumes that treatment effects stabilize after five years, with $L = 5$. Varying this lag around 5 years yields qualitatively similar results (see Section 4.4 and Appendix B). The choice of L generally involves balancing a variance-bias trade-off. As the horizon h increases, the number of included events decreases, raising the variance of the estimated coefficients. On the other hand, a longer h allows more time for treatment effects to stabilize, reducing potential bias in estimated coefficients due to heterogeneous treatment effects. Given that our data includes firm entries between 2003 and 2022, setting L to 5 allows us to estimate the effects of FDI on citations for up to 14 periods after the initial entry. Extending the stabilization lags beyond this would significantly reduce the time horizon over which we can study

[9] In practice, 95 percent of subsequent investments occur within five years of the initial entry. As discussed below, the number of stabilization lags, L, is set to 5, which results in the exclusion of most subsequent FDI from both the control and treatment groups. This exclusion leaves little room for further investments to affect the results even if we consider any FDI event as the treatment.

[10] For example, recent works like Azoulay et al. (2019); Arora et al. (2021a); Moretti (2021); Jiang et al. (2024) have used the Asinh transformation for patent applications and scientific publications, which often contain zero values. The Asinh transformation of a variable x is defined as $\text{asinh}(x) = \log(x + \sqrt{x^2 + 1})$.

[11] As reported in Section 4.4 and Appendix B, our results remain robust when applying the $\log(1 + \textit{citations})$ transformation to handle zeros in the data. Although the effects are qualitatively similar, they are significantly higher with the log transformation.

such effects.

Choice of Controls. For all the results presented below, we use the same specification, which includes three lags of the asinh cumulative citations (with $p = 3$) 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 discussed 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.[12]

4 Results

4.1 Increased Citation Flows Between Host Countries and Investing Firms

Our baseline specification estimates Equation (2) using the asinh transformation of cumulative citations received by (or made by) firm i from (to) country c since 1995 as the outcome variable. We control for firm-year and country-year fixed effects, three lags of the outcome variable, and its trend 4 to 6 years prior to the event. The stabilization period, L, is set to 5 years.

Figure 2 presents the estimated coefficients, β_h, demonstrating the positive effects of both greenfield and brownfield FDI on knowledge spillovers from and toward the investing firms. Specifically, Figure 2a shows that, five years after a brownfield investment, destination countries increase their asinh cumulative citations to investing firms by 0.063. This translates into an average increase of 7.8% in citations from the FDI destination country compared to non-destination countries.[13] Similarly, Figure 2b shows the estimated coefficients for greenfield FDI, which are statistically similar to those for brownfield FDI. However, the lower pre-treatment average for greenfield investments implies a higher percentage change in citations, at 10.6%. Based on the 95% confidence intervals for the estimated coefficients, the results suggest an increase in citations of 5.7% to 10% following brownfield FDI and 9% to 12.2% following greenfield FDI. Turning to citations made by investors to host countries, Figures 2c and 2d show that citations made to host countries by investors increase by around 10.8% and 13.4% for brownfield and greenfield FDI, respectively.[14]

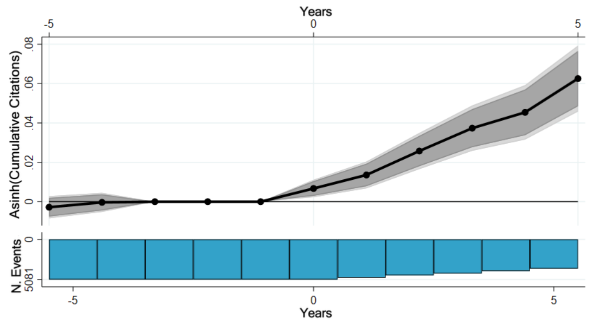
[12] When setting $p = 2$ and replacing the term $y_{t-4} - y_{t-6}$ with $y_{t-3} - y_{t-5}$, we obtain qualitatively similar results.

[13] Given the pre-treatment average asinh cumulative citations of 1.158, we translate these effects into percentage changes relative to pre-event untransformed cumulative citations using the formula $\sinh(1.158 + 0.063)/\sinh(1.158)$-1.

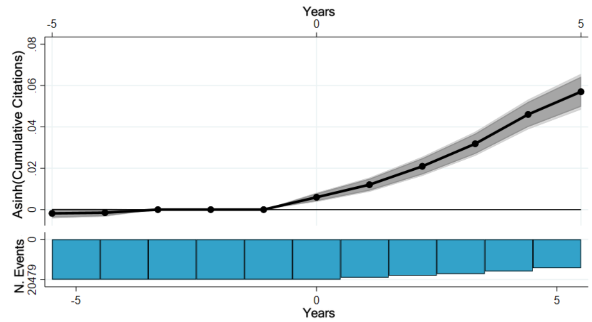
[14] Figure 2d indicates a slight pre-trend, suggesting that greenfield investors may select destinations where they have already benefited from knowledge inflows in the years prior to their entry. However, even when accounting for this pre-trend, the analysis still points to a sizable positive treatment effect. Moreover, Section 4.3 examines a specification at the firm-country-industry level, where we can fully control for pre-trends and the results remain consistent with the overall findings.

Figure 2: Impacts of FDI on Citations, World Sample

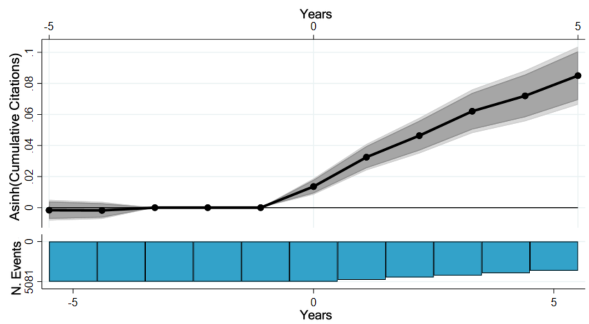
(a) Country's Citations to Brownfield Investors



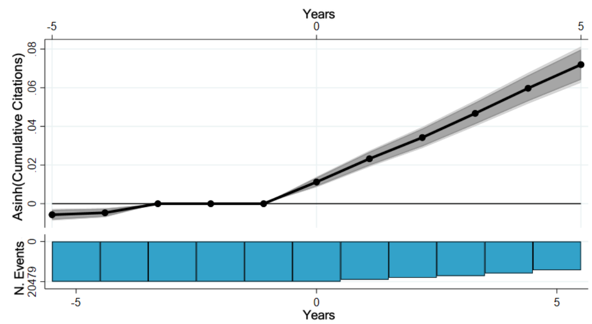
(b) Country's Citations to Greenfield Investors



(c) Brownfield Investors' Citations to Country



(d) Greenfield Investors' Citations to Country

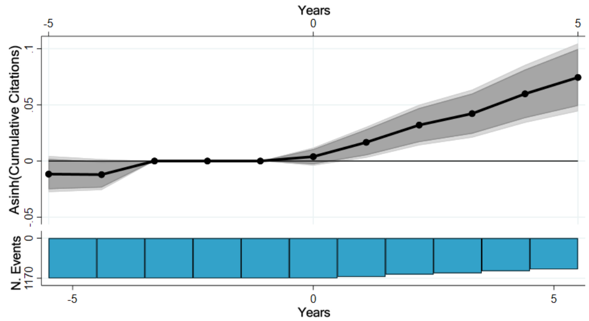


Note: This figure presents the estimated coefficients β_h in Eq. (2). The y-axis in each panel represents the change in the Asinh (inverse hyperbolic sine) transformation of cumulative citations from period $t-1$. "World sample" refers to the sample of firm-country pairs obtained matching PATSTAT company 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.

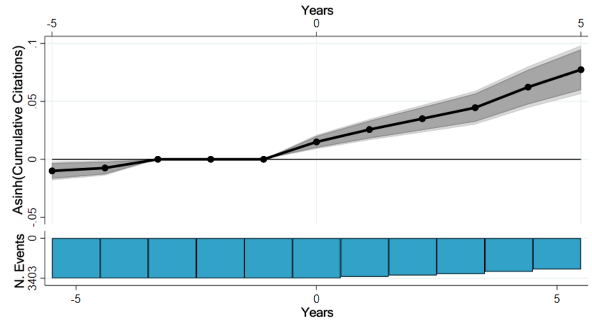
Figure 3 repeats the same analysis for the sample of U.S. firms identified through matches with the sample constructed by Arora et al. (2021b). Carrying out the same calculations as before, they show an 8.4% and a 9.9% increase in citations from host countries toward brownfield and greenfield investors, respectively. Regarding citations by investors, Figures 3c and 3d imply that citations by brownfield investors increase by 12.7% and by greenfield investors by 11.1%. Overall, these estimates are not statistically different from those obtained using the world sample summarized in Figure 2.

Figure 3: Impacts of FDI on Citations, US Sample

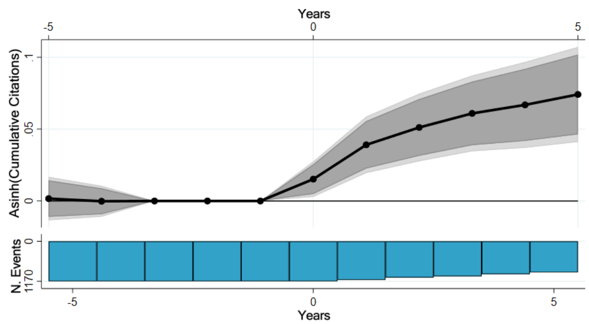
(a) Country's Citations to Brownfield Investors



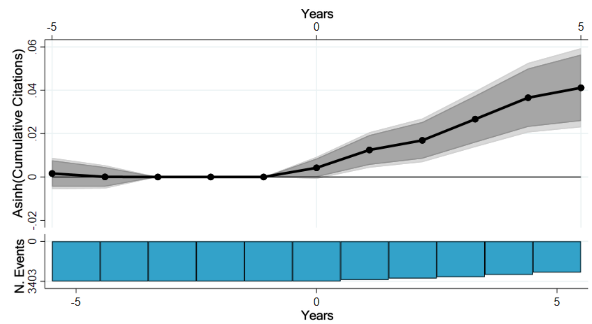
(b) Country's Citations to Greenfield Investors



(c) Brownfield Investors' Citations to Country



(d) Greenfield Investors' Citations to Country



Note: This figure presents the estimated coefficients β_h in Eq. (2). The y-axis in each panel represents the change in the Asinh (inverse hyperbolic sine) transformation of cumulative citations from period $t-1$. "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.

4.2 Heterogeneity by Host Countries' Absorptive Capacity

We now examine the heterogeneity in the baseline estimation results based on host countries' characteristics, often referred to as "absorptive capacities." For example, Alfaro (2017) highlights the role of absorptive capacity in determining FDI's impact on economic growth. Two key factors commonly cited in the literature as necessary for realizing the benefits of FDI are particularly relevant to our focus of knowledge spillovers: a business environment conducive to adopting new technologies and adequate levels of human capital. In the following, we measure a country's absorptive capacity using (1) its knowledge stock, represented by total patent stocks, and (2) the technological proximity between the host country and the foreign investor, measured by "technological closeness" as defined below.

We use the following specification to account for heterogeneity in host countries' absorptive capacities:

$$y_{ic,t+h} - y_{ic,t-1} = \beta_h D_{ic,t} + \beta_h^I D_{ic,t} \times I_{ic,t} + \sum_{k=1}^{p} \gamma_k^h y_{ic,t-k} + \eta' \mathbf{x}_{ic,t-1} + \delta_{it}^h + \delta_{ct}^h + \varepsilon_{ic,t}^h, \qquad (4)$$

where the coefficients of the interaction term, β_h^I, capture how the spillover effects vary across host countries with different absorptive capacity, I_{ic}. It is also included in the vector of controls $\mathbf{x}_{ic,t-1}$ where relevant.

Heterogeneity by Host Countries' Patent Stocks. In the first set of specifications, $I_{ic,t}$ is a time-invariant dummy variable equal to 1 when the host country c is among the top 10% of countries worldwide by the total number of patents listed in the PATSTAT data during the sample period. This dummy variable helps distinguish host countries with high absorptive capacity in terms of capability to innovate and absorb new technologies.

Figures 4 and 5 show our results. The upper panels of each figure report the coefficients β_h, corresponding to the bottom 90% of countries where $I_{ic,t} = 0$, while the lower panels display the coefficients for the interaction term, β_h^I. The positive and significant β_h^I indicates that more citation flows occur in both directions when the FDI host country is among the top 10% of patent producers.

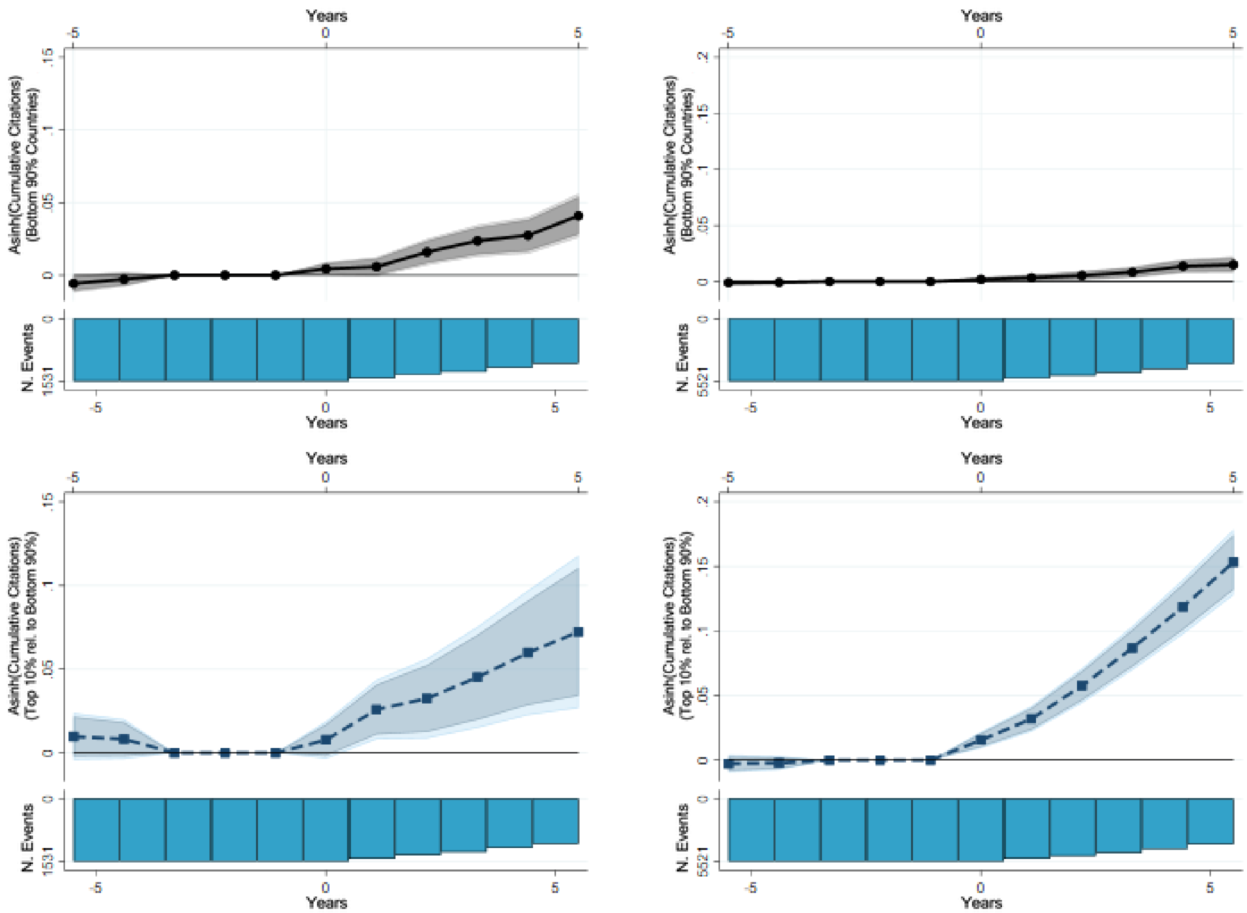
Starting from Figure 4a, we find that being a top patent-producing destination significantly boost knowledge spillovers from brownfield FDI, almost doubling the estimated baseline effect in asinh terms. Since the pre-treatment number of citations varies across these two country groups, we calculate the effects in percentage terms by inverting our transformation as discussed in Section 4.1.[15] Compared to the pre-investment period, brownfield FDI destination countries with low initial patent stocks increase their citations to investors by 5.1% after five years. The corresponding increase for the top patent-producing countries is more than double, at 12.3%. This percentage increase accounts for the fact that pre-investment citations are higher in these top countries. For greenfield FDI effects shown in Figure 4b, the difference is even more pronounced: the most innovative host countries increase their citations by 21.3%, while the less innovative countries only increase citations by 2.7%.

Figure 5 shows that the reverse knowledge flows (from host countries to foreign investors) are also subject to similar heterogeneity. In the case of brownfield investments, citations made by investors increase by 6.8% in less innovative destinations five-years post entry, while they rise by 17.6% among the top 10%. For greenfield FDI, the increases of 4.9% and 24.5%, respectively.

[15] Specifically, we compute percentage changes corresponding to upper panels as $\sinh(\bar{y}_{ic,t-1}^{\text{low}} + \hat{\beta}_h)/\sinh(\bar{y}_{ic,t-1}^{\text{low}}) - 1$ and results for the lower panels as $\sinh(\bar{y}_{ic,t-1}^{\text{high}} + \hat{\beta}_h + \hat{\beta}_h^I)/\sinh(\bar{y}_{ic,t-1}^{\text{high}}) - 1$, where $\bar{y}_{ic,t-1}^{\text{low}}$ and $\bar{y}_{ic,t-1}^{\text{high}}$ represent the sample averages of the outcome variable in the period before treatment for the respective sample of host countries.

Figure 4: Host Countries' Citations to Investors, by Host Countries' Patent Stock

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



Note: This figure presents the estimated coefficients β_h and β_h^I in Eq. (4). The y-axis in each panel represents the change in the Asinh (inverse hyperbolic sine) transformation of cumulative citations from period $t-1$. 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 4 to 6 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 patent stocks. The bottom bars represent the number of treated samples.

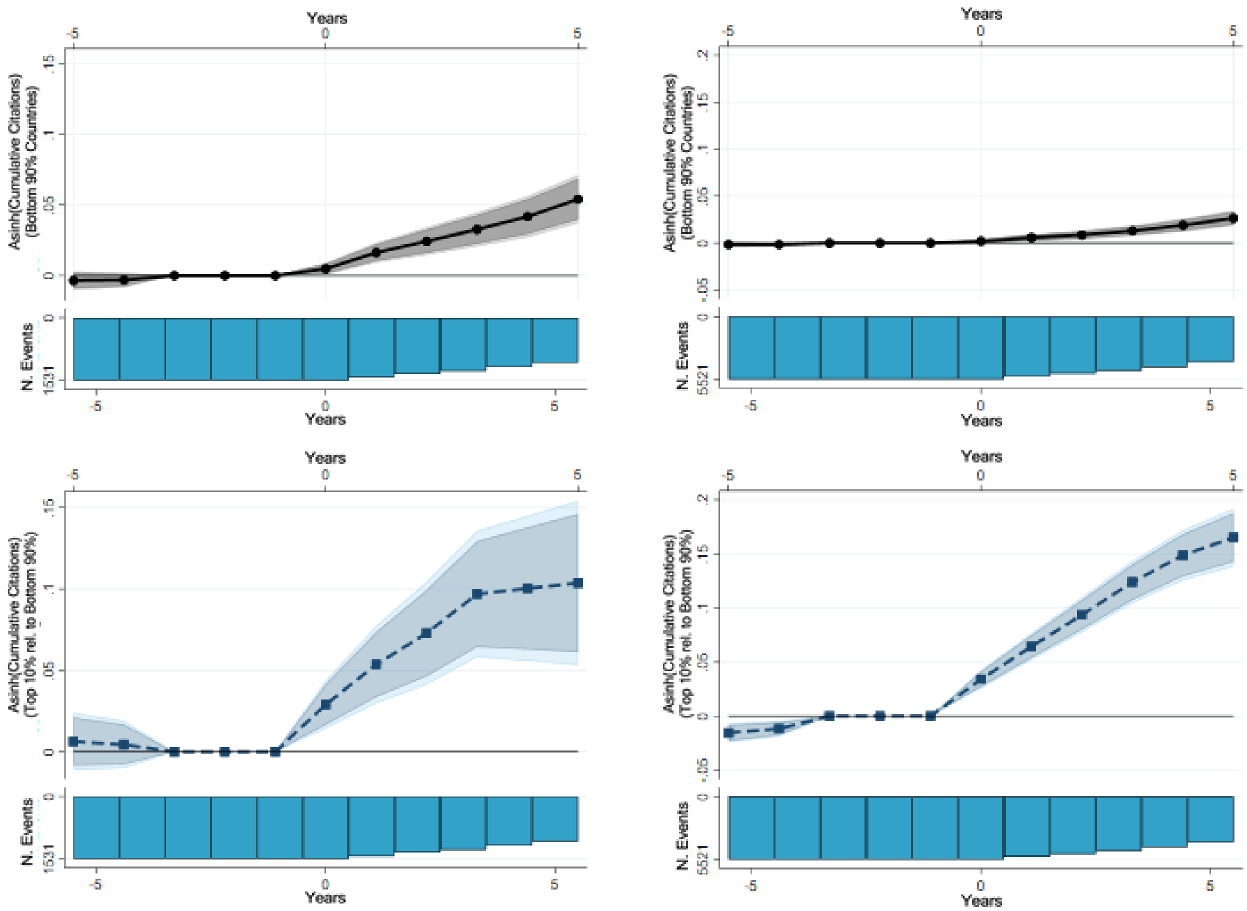
Heterogeneity by Host Countries' Technological Closeness. The second set of results uses the technological closeness between the patent stocks of host countries and those of investors as $I_{ic,t}$. This closeness measure is calculated using cosine similarity between the vectors representing patent portfolios of investors and host countries across different IPC 4-digit subclasses, following the approach of Jaffe (1986) and Bloom et al. (2013).[16] The technological closeness for each

[16]For example, consider two IPC 4-digit subclasses, I1 and I2. Investor A has 3 patents in I1 and 4 patents in I2,

Figure 5: Investors' Citations to Host Countries, by Host Countries' Patent Stock

(a) Brownfield Investors' Citations to Countries

(b) Greenfield Investors' Citations to Countries



Note: This figure presents the estimated coefficients β_h and β_h^I in Eq. (4). The y-axis in each panel represents the change in the Asinh (inverse hyperbolic sine) transformation of cumulative citations from period $t-1$. 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 4 to 6 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 patent stocks. The bottom bars represent the number of treated samples.

firm-country pair is computed for each period t based on the patent stock accumulated over the previous ten years. It captures the technological compatibility between host countries and investors and serves as an indicator of the presence of relevant human capital. In this specification, $I_{ic,t}$ is

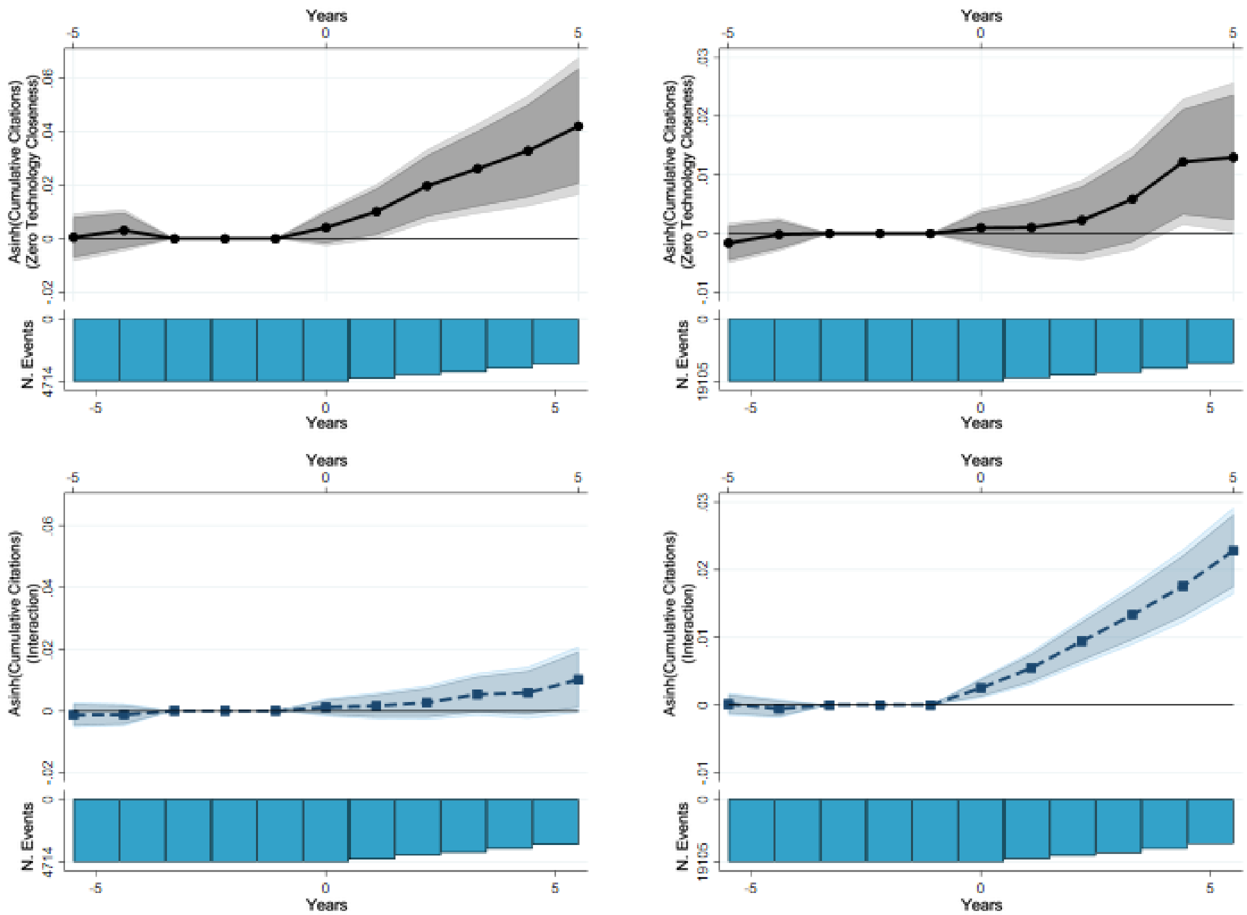
while host country B has 4 patents in I1 and 3 patents in I2. The cosine similarity of their patent vectors is calculated using the formula $\frac{\mathbf{A}\cdot\mathbf{B}}{\|\mathbf{A}\|\|\mathbf{B}\|} = (12+12)/(5*5) = 0.96$. Here, $\mathbf{A} = (3,4)$ and $\mathbf{B} = (4,3)$ are the patent vectors of investor A and country B, with each element represents the number of patents in industries I1 and I2, respectively. In our analysis, these vectors span 628 dimensions, corresponding to the total number of unique IPC 4-digit classes.

also included in the control vector, $\mathbf{x}_{ic,t-1}$.

Figure 6: Host Countries' Citations to Investors, by Technology Closeness. Upper panels: Zero Closeness; Lower Panels: Interaction Effect.

(a) Countries' Citations to Brownfield Investors

(b) Countries' Citations to Greenfield Investors



Note: This figure presents the estimated coefficients β_h and β_h^I in Eq. (4). The y-axis in each panel represents the change in the Asinh (inverse hyperbolic sine) transformation of cumulative citations from period $t-1$. 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 closeness between technologies of the investors and host countries. The bottom bars represent the number of treated samples.

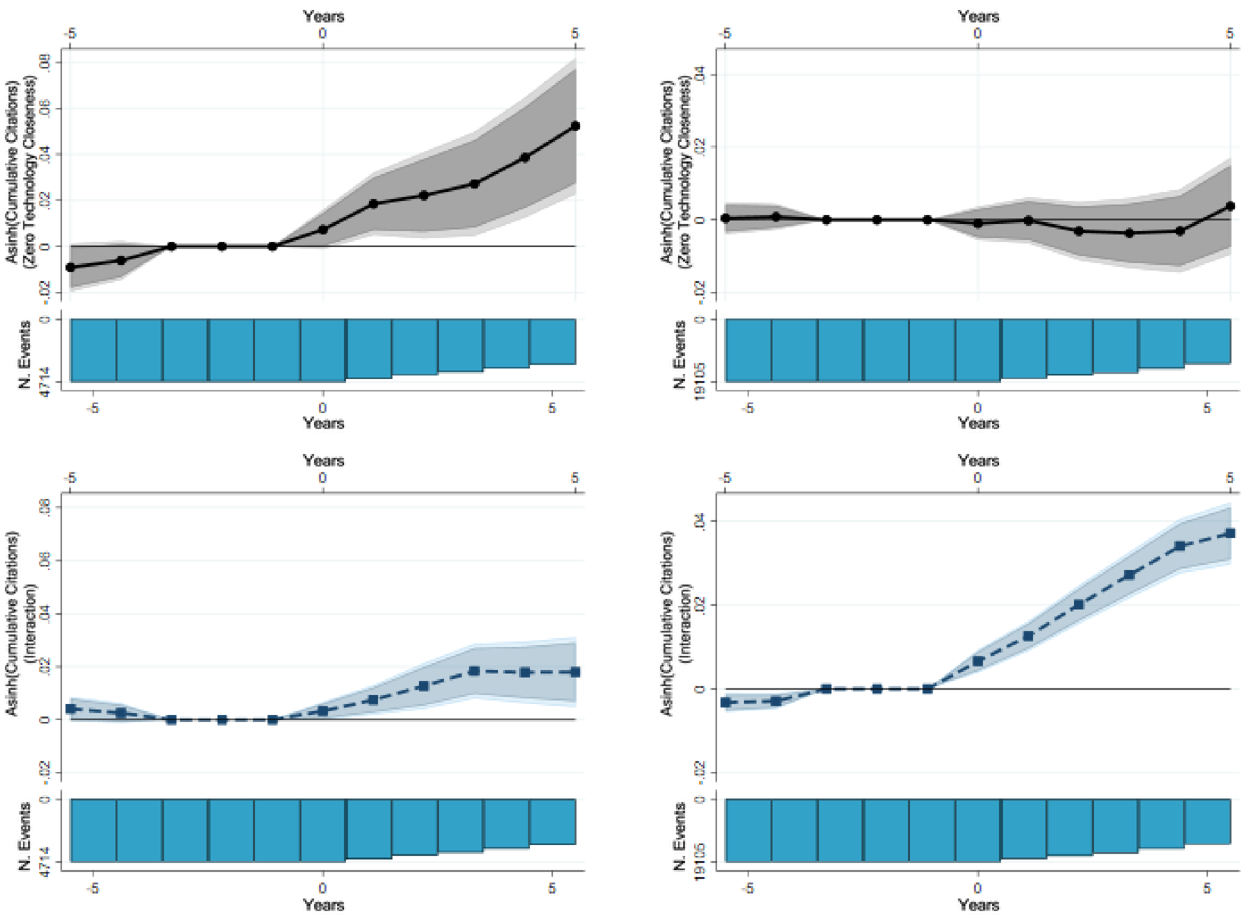
Figures 6 and 7 report the estimated coefficients.[17] We find that higher absorptive capacity, particularly in the case of greenfield FDI, leads to larger citation flows (Figure 6b and 7b). When technological proximity is low, there is no significant increase in either inward or outward knowledge

[17]We multiply this proximity measure by a factor of 10 for readability.

Figure 7: Investors' Citations to Host Countries, by Technology Closeness. Upper panels: Zero Closeness; Lower Panels: Interaction Effect.

(a) Brownfield Investors' Citations to Countries

(b) Greenfield Investors' Citations to Countries



Note: This figure presents the estimated coefficients β_h and β_h^I in Eq. (4). The y-axis in each panel represents the change in the Asinh (inverse hyperbolic sine) transformation of cumulative citations from period $t-1$. 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 closeness between technologies of the investors and host countries. The bottom bars represent the number of treated samples.

spillovers from greenfield FDI (see the upper-right panels of Figure 6 and 7). In contrast, citations increase by 6.7% at the 25^{th} percentile of similarity and 9.2% at the 75^{th}. In terms of absolute citation counts, this means that the FDI effect on outward citations at the 75^{th} percentile is more than double that at the 25^{th}, as pre-treatment citations to investing firms are already higher in destinations with more similar patent stocks. Reverse citation (from investor to destination) increase by 4% at the 25^{th} percentile of similarity and 11.1% at the 75^{th} percentile.

In contrast, technological similarity seems to play a smaller role in the case of brownfield FDI, where destination countries benefit from knowledge spillovers of even when technological closeness is low, as shown in the upper-left panels of Figure 6 and 7. The percentage effect on citations from countries at the 25^{th} percentile of similarity is 8%, while at the 75^{th} percentile, the increase is 7.7%. Since pre-treatment citation flows are larger with greater similarity, moving from the 25^{th} to the 75^{th} percentile increases citations by about 32% in absolute value. Reverse citations exhibit a similar pattern, without notable increase as technological similarity increases.

Table 1 provides a summary of the estimated effects of FDI on knowledge spillovers as measured by the subsequent increase in patent citations after five years.

Table 1: Knowledge Spillover Effects of FDI: Percentage Increase in Patent Citations

Knowledge Direction		**Brownfield**	**Greenfield**
A. Baseline Effects			
Investors → Destination		7.8%	10.6%
Destination→ Investors		10.8%	13.4%
B. Heterogeneity Effects by Destination Innovation Capacity			
Investors → Destination	Bottom 90%	5.1%	2.7%
	Top 10%	12.3%	21.3%
Destination→ Investors	Bottom 90%	6.8%	4.9%
	Top 10%	17.6%	24.5%
C. Heterogeneity Effects by Firm-Destination Technology Similarity			
Investors → Destination	25th	8%	6.7%
	75th	7.7%	9.2%
Destination→ Investors	25th	12%	4%
	75th	10%	11.1%

Notes: This table presents the average percentage increase in citation flows five years after the initial FDI investment. These results use the "World sample," comprising firm-country pairs matched by PATSTAT company names with FDI data firm names. Part A1 summarizes baseline results at the firm-destination level, as shown in Figure 2. Parts B and C provide a summary of heterogeneity results from Figures 4 through 7.

4.3 The Scope of Knowledge Spillovers: Country-Industry-level Results

This section presents additional results aimed at refining the identification of knowledge spillovers and broadening the scope of our analysis by examining citations at the country-industry level. This approach enhances identification by controlling for additional country-industry-time fixed effects.

Furthermore, it enables us to investigate whether these spillover effects extend beyond the directly targeted industries and how they are influenced by technological linkages between industries. Specifically, we use an industry-level measure of technological relatedness to determine whether industries closely related to those directly impacted by FDI receive greater knowledge spillovers from these investments. This analysis includes both sectors directly targeted by FDI and those indirectly connected through technology citation networks. Due to the significant computational challenges posted by the curse of dimensionality, we limit this analysis to the U.S. sample.

In the preceding section, we presented results demonstrating the impact of technology closenss between 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 strengthening our conclusion that FDI leads to cross-border knowledge spillovers.

Firm-Industry Evidence: Setup and Specifications. We investigate whether knowledge spillovers transmit more intensely to industries that are directly targeted by FDI, 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:

$$y_{icl,t+h} - y_{icl,t-1} = \beta_h T_{icl,t} + \sum_{k=1}^{3} \gamma_k^h y_{icl,t-1-k} + \delta_{ilt}^h + \delta_{clt}^h + \varepsilon_{icl,t}^h, \tag{5}$$

where now $y_{ilc,t}$ is the asinh cumulative citations made to firm i's patents by country c's patents in a two-digit ISIC industry 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 control for industry-specific time-varying factors that potentially affect citation trends. For example, 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 a firm enters a specific industry l in country c. The resulting coefficients can be interpreted as the differences in citations between treated destination-industry pairs and untreated pairs before and after FDI, while controlling for time-varying factors specific to the firm-industry and the destination-industry time-varying factors (such as changes in patenting activities). To obtain industry-level citations, we convert IPC codes to ISIC codes using the crosswalk provided by Lybbert and Zolas (2014). The definition of an industry varies depending on the type of FDI under consideration. For greenfield FDI, the data provided by fDi Markets includes detailed descriptions of targeted sectors, which may vary for the same firm across different projects and destinations. We map these sectors to two-digit ISIC codes. For M&A data, only the primary sector of the investing and target firms is available, which may not always correspond to the actual sector where the FDI occurs—a potential limitation of the data.

Spillovers from FDI in Other Industries and Knowledge Linkages. We also examine spillovers from FDI in industries other than the treated industry and whether the effects depend on knowledge linkages between the FDI industry and other industries. Specifically, we investigate whether industries that rely more on knowledge from the treated industries experience greater knowledge flows. To do so, we first calculate the "knowledge input coefficients" of industry l from industry k at the ISIC 2-digit level within destination country c as:

$$a^c_{l \leftarrow k} \equiv \frac{\text{Total citations made by industry } l \text{ towards } k \text{ in country } c \text{ between 1990-2000}}{\text{Total citations made by industry } l \text{ in country } c \text{ between 1990-2000}}, \tag{6}$$

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

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 used in the previous analysis. As a result, the treatment intensity $T_{icl,t}$ will be stronger for industries l that more intensely cite sector k.

Sample Choices. Adding the additional dimension of industry significantly increases the number of firm-industry-country-time combinations, extending beyond the computational capabilities available to us. Therefore, we have chosen to restrict our analysis to the U.S. sample and focus on the top 30 innovative destinations, as identified by their patent stocks. This approach leaves us with around 276 million observations.

Industry-level Greenfield FDI Results. Figure 8a and 8c report the knowledge spillover effects of greenfield FDI, estimated using Equation (5) where $T_{icl,t} = 1$ when the foreign firm i started FDI in industry l of country c. These results are qualitatively consistent with our baseline findings. In addition, we do not detect any violation of the parallel trend hypothesis, in contrast to the slight pre-trend noted in the baseline case of reverse citations. We attribute this improvement to the tighter identification achieved by including host country-industry-time and firm-industry-time fixed-effects, as opposed to just country-time and firm-time fixed effects. Quantitatively, we find an average increase of 9% in citations made by host industries to greenfield investors. In the opposite direction, foreign investors increase their citations to target industries by 10.8%. These results are at the lower end of the effects that we estimated in our baseline, where the comparison is made at the firm-country level rather than firm-industry-country level.

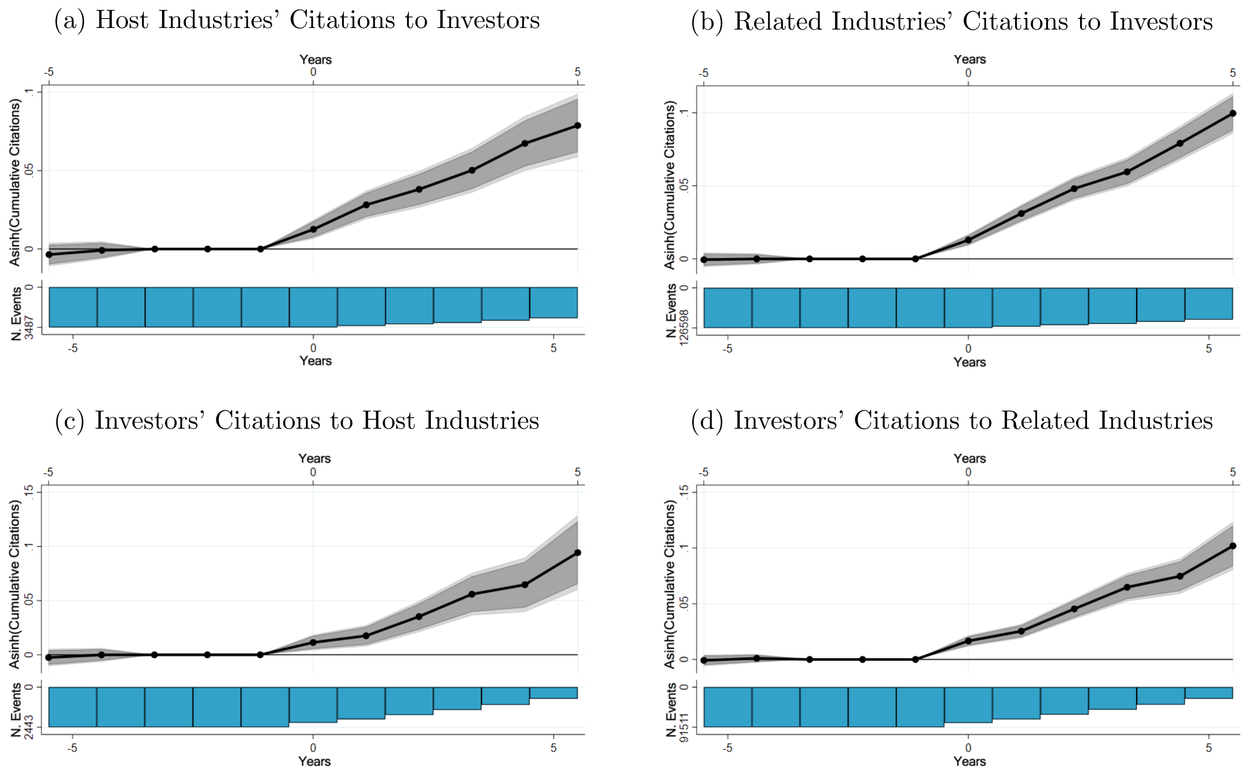
The results for the treatment interacted with the knowledge input coefficients, shown in Figure 8b and 8d, demonstrate clear positive effects, providing evidence that knowledge spillovers extend beyond directly targeted industries. As indicated by the blue bars in each graph, the number of treated country-industries in this specification is significantly larger than the number of directly targeted industries, suggesting a strong contribution of technologically related—but not necessarily directly targeted—industries to estimated effects at the country level.

We also estimate a specification that includes an interaction between $T_{icl,t}$ and a dummy variable indicating whether industry l is directly targeted, to ensure our results are not solely driven by these industries. Our findings show that the percentage increase in citations for non-targeted industries remains substantial and statistically significant, with an increase of 28.5% when $a_{l \leftarrow k} = 1$. This further supports the idea that knowledge spillovers extend beyond directly involved industries, benefiting other sectors through the knowledge input network.

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. Unlike in the greenfield FDI case, the actually treated industry may differ widely from the industries we assign treatment to, leading to potential attenuation bias. With these caveats in mind, Figure 9 presents the industry-level results for brownfield FDI derived from estimating Equation (5).

When we assign direct treatment to industry l using the primary industry of the target firm, the results for directly targeted industries' citations toward investors remain qualitatively consistent with the baseline estimation results reported earlier, although there is a slight indication of a potential violation of the parallel trends hypothesis. However, we find no effect on citations from investing firms to the target country-industry. This discrepancy may be due to the fact that the industry actually subject to brownfield FDI is technologically related to the primary industry of the target firm, but does not coincide with it. In this case, the primary industry of the target firm would still receive indirect knowledge spillovers, but would not generate spillovers back to the

Figure 8: 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 estimated coefficient β_h from the greenfield FDI version of the specification in Eq. (5). The y-axis in each panel represents the change in the Asinh (inverse hyperbolic sine) transformation of cumulative citations from period $t-1$. The sample used in this figure is the "US sample" described in Subsections 2.3 and 4.3. 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 4.3 for the left and right columns, respectively. The bottom bars represent the number of treated samples.

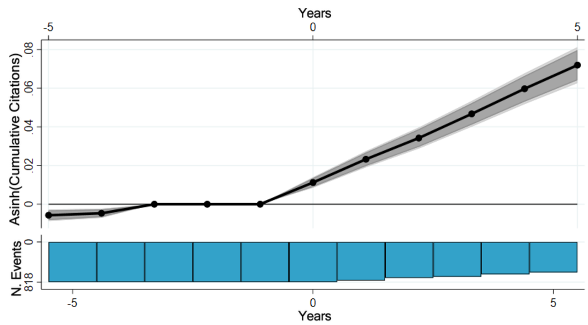
investors' industry.[18] This conjecture is supported by the fact that we recover positive knowledge spillovers in both directions when employing the knowledge input coefficient specification.

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, with the exception of reverse citations from the targeted industry to investors.

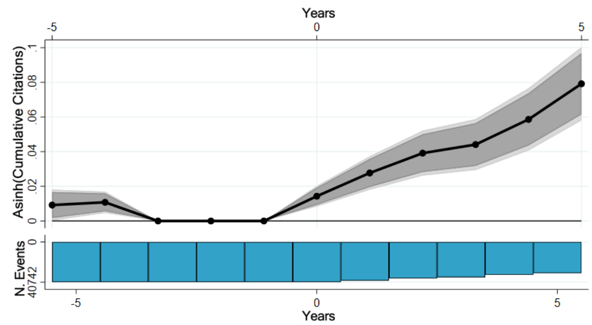
[18] It is important to note that the target and investors' primary industry do not coincide in 75% of cases. Thus, the fact that the target firm industry is related to the treated sector does not imply that the investors' industry is related to the sector actually treated by FDI.

Figure 9: Citation Flows Between Brownfield Investors and Countries' Industries

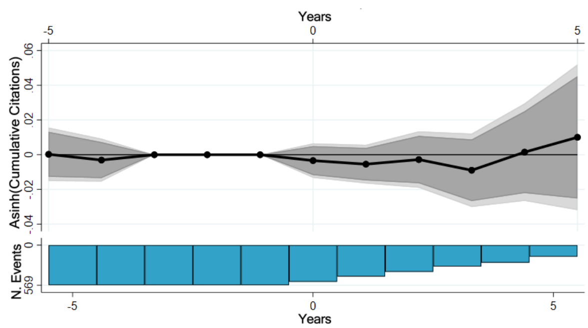
(a) Target Industries' Citations to Acquirers



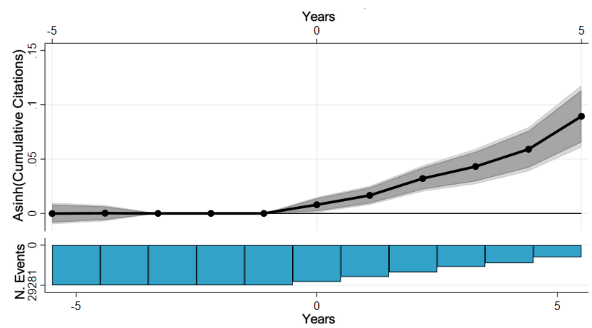
(b) Related Industries' Citations to Acquirers



(c) Acquirers' Citations to Target Industries



(d) Acquirers' Citations to Related Industries



Note: This figure presents the estimated coefficient β_h from the brownfield FDI version of the specification in Eq. (5). The y-axis in each panel represents the change in the Asinh (inverse hyperbolic sine) transformation of cumulative citations from period $t-1$. The sample used in this figure is the "US sample" described in Subsections 2.3 and 4.3. 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 4.3 for the left and right columns, respectively. The bottom bars represent the number of treated samples.

4.4 Robustness

Dropping Citations From Target Firms. To address the concern that increased citations from host countries to investors might mechanically result from self-citations by newly acquired subsidiaries, rather than genuine knowledge spillovers, we assess the robustness of our findings by excluding citations originating from firms acquired in brownfield FDI, where target firm information is available.[19] We do this by applying our name matching strategies to exclude citations coming from any company that has a name similar to the acquired affiliate or the acquiring firm.

We carry out this exercise for both our world sample and the U.S. sample discussed earlier.

[19]Since the greenfield FDI data does not include the names of the newly established entities, this robustness exercise is limited to brownfield investment episodes.

The greater reliability of ownership structures provided by the U.S. sample is particularly valuable here, as we are specifically concerned about potential self-citations by acquiring firms. The resulting coefficients are reported in Figure 10. We estimate an increase in host country citations to brownfield investors of 8.9% in the world sample and 7.3% in the U.S. samples. In both cases, the confidence intervals for these estimates overlap with the baseline estimates of 7.8% and 8.4% for the respective samples presented in the previous section. Regarding investors' citations to host countries, the corresponding increases are 10.8% and 4.5%, respectively. Among these estimates, only the latter shows a significant difference from the baseline. Its smaller size for the U.S. sample suggests that about two thirds of the effect observed in Figure 3c is due to citations made by U.S. investors to their acquired affiliates in destination countries. For the world sample, the results remain the same.

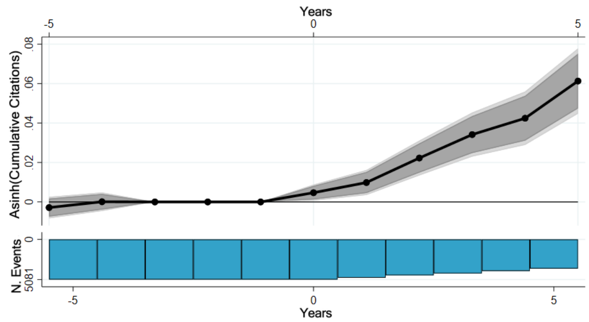
Other Robustness. The results of the baseline analysis remain robust to various alternative specifications: (1) expanding the set of destination countries to all those available in PATSTAT; (2) limiting forward citations to five years after the application of a patent to address truncation issues for 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 $\text{asinh}(x)$; (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 together as a single "FDI investment" variable that does not distinguish different investment types; (8) checking the robustness of our results for the US sample of firms to considering only patents originally assigned to US firms. 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. The details and figures corresponding to each specification, the rationale behind these exercises and the estimated effects are reported in Appendix B.

5 Conclusion

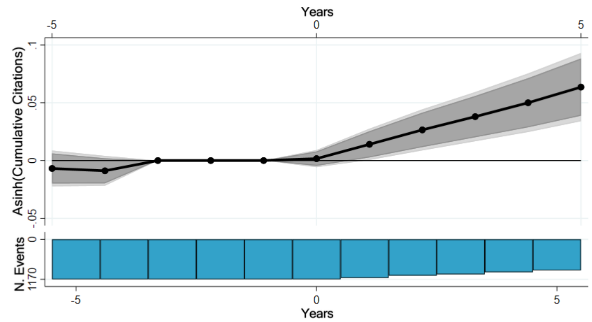
In this paper, we explore the effects of brownfield and greenfield FDI on knowledge diffusion, measured by citation flows between investing firms and FDI destinations. Our results show that, following the first entry of an investor into an FDI destination, citations made by the destination to investors 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

Figure 10: Impacts of Brownfield FDI on Citations, Excluding Targeted Firms

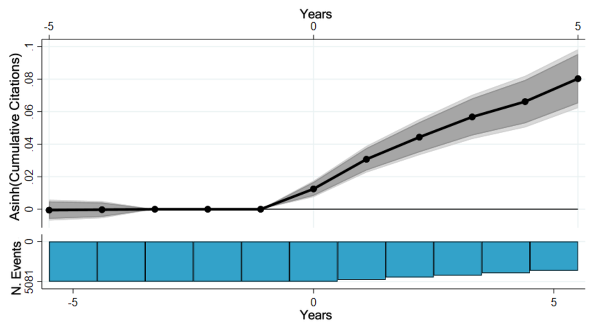
(a) Countries' Citations to Investors, World Sample



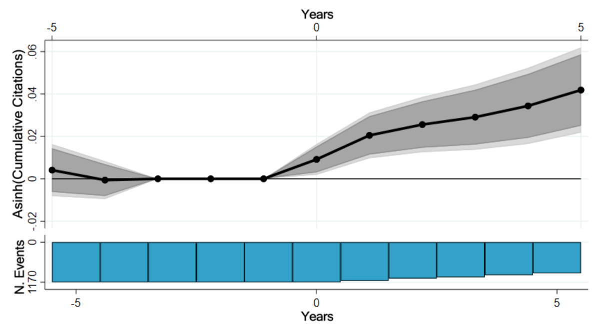
(b) Countries' Citations to Investors, US Sample



(c) Investors' Citations to Countries, World Sample



(d) Investors' Citations to Countries, US Sample



Note: This figure presents the estimated coefficients β_h in Eq. (2). The y-axis in each panel represents the change in the Asinh (inverse hyperbolic sine) transformation of cumulative citations from period $t-1$. "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.

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. Our findings underscore the critical role of absorptive capacity in determining the extent of FDI-driven knowledge spillovers. Host countries with more robust innovation systems—measured by patent stocks and technological proximity to investing firms—are able to capitalize on these spillovers, while countries with lower absorptive capacity struggle to do so. This suggests that FDI alone is not a panacea for fostering convergence; rather, it may amplify existing technological disparities unless accompanied by policies that enhance domestic capacity for innovation and learning.

Also, 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 citations, which are available only for countries that produce and report sufficient numbers of patents. However, 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. While our paper focuses on patent citations, one remaining question that warrants further exploration is whether new or acquired firms increase their overall patenting activities following investments. This is particularly promising for brownfield FDI thanks to the unique feature of the brownfield FDI data where information about acquired firm is available and thus can be matched with patent application data to evaluate the effect of acquisition on target firm's innovation. Moreover, it is conceptually relevant because the literature suggests that while M&A might be particularly motivated by the desire to enhance the innovation activity of target firms, it could also reduce overall patenting if the primary goal is to eliminate potential future competitors. Further research in this area would provide valuable insights.

References

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

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.

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.

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.

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.

Bloom, N., M. Schankerman, and J. Van Reenen (2013): "Identifying technology spillovers and product market rivalry," *Econometrica*, 81, 1347–1393.

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.

Cai, J., N. Li, and A. M. Santacreu (2022): "Knowledge diffusion, trade, and innovation across countries and sectors," *American Economic Journal: Macroeconomics*, 14, 104–145.

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.

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.

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.

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

Jaffe, A. B. (1986): "Technological opportunity and spillovers of R&D: evidence from firms' patents, profits and market value," .

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.

Jiang, K., W. Keller, L. D. Qiu, and W. Ridley (2024): "International joint ventures and internal technology transfer vs. external technology spillovers: Evidence from China," *Journal of International Economics*, 150, 103939.

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

——— (2022): "Knowledge spillovers, trade, and foreign direct investment," in *Oxford Research Encyclopedia of Economics and Finance*, Oxford.

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.

——— (2013): "The gravity of knowledge," *American Economic Review*, 103, 1414–1444.

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.

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.

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.

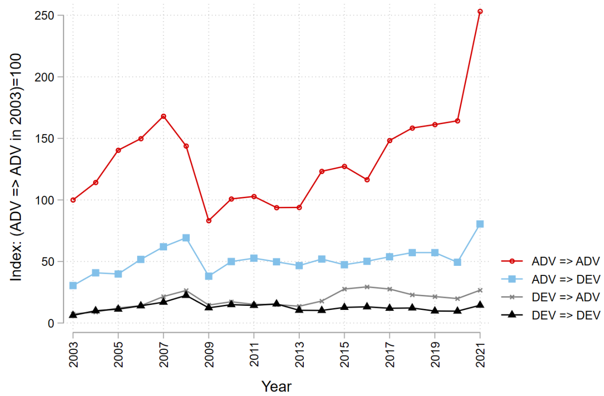
Toews, G. and P.-L. Vézina (2022): “Resource discoveries, FDI bonanzas, and local multipliers: Evidence from Mozambique,” *Review of Economics and Statistics*, 104, 1046–1058.

A Appendix A: Time-series evolution of FDI and patent citations

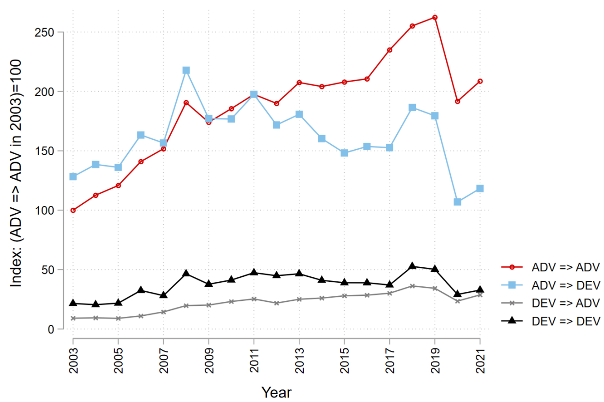
To illustrate overall patterns of FDI and patent citations over time, we group the counties into advanced and developing countries, and trace investment and citation flows between and across groups, leading to four separate series in each panel in Figure A.1: from advanced to advanced countries (ADV⇒ADV); 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 the number of ADV⇒ADV transactions or citations in 2003.

The top panel (A.1a) 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 (A.1b) 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 (A.1c) 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.

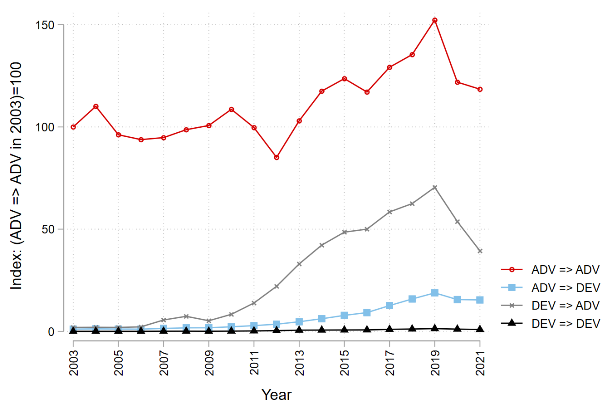
Figure A.1: FDI and Patent Citations over Time by Country-pair Income Levels



(a) Brownfield FDI



(b) Greenfield FDI



(c) Patent citations

Note: This figure plots the time-series evolution pattern of brownfield FDI (A.1a), greenfield FDI (A.1b), and patent citations (A.1c), 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 level where the number ADV⇒ADV transactions or citations in 2003 corresponds to 100.

B Appendix B: 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 2, 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 2 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 to Including All Available Countries. In our baseline, we limited our analysis to 60 top destinations and origin countries, representing 99.7% of global patenting activity. Figure B.2 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 to 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 B.3 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 B.4 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 B.5, 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 B.6 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 B.7 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%.

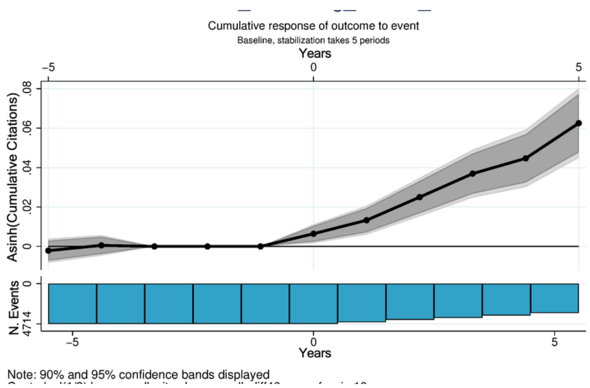
Robustness to Grouping Greenfield and Brownfield FDI Together. The LP-DiD methodology rests on the assumption that treatment effects stabilize after L periods. In our baseline, we considered greenfield and brownfield FDI as separate events. This may create a bias in our results, as we have shown that 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 B.8 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 B.7 $(30,845)$. 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.

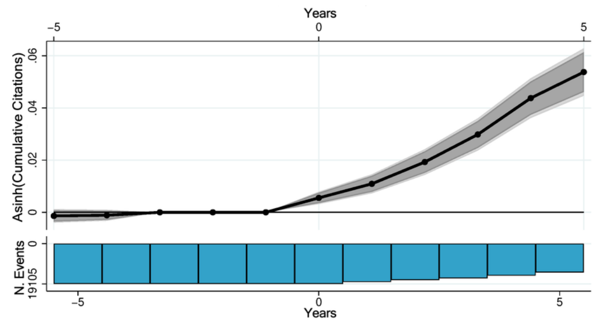
Robustness to Using Only Patents Originally Assigned to US Firms in the US Sample. We further consider the robustness of our results for the US sample to limiting our citation measure to only include patents that are originally assigned to US firms. Indeed, as noted in the main text, the Arora et al. (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 B.9 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 3 (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 B.2: Impacts of FDI on Citations, World Sample, Using All Available Destinations

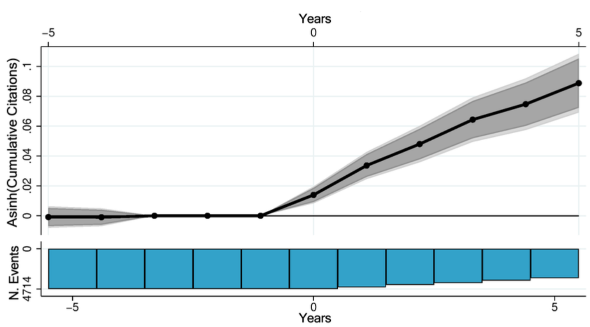
(a) Country's Citations to Brownfield Investors



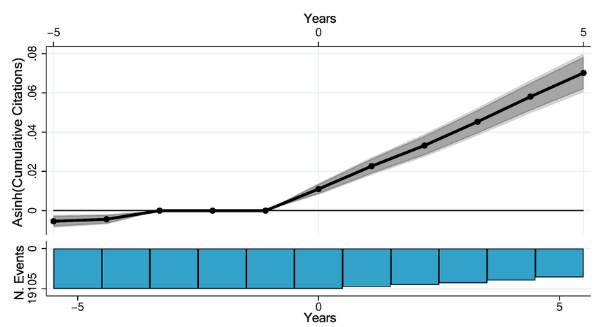
(b) Country's Citations to Greenfield Investors



(c) Brownfield Investors' Citations to Country



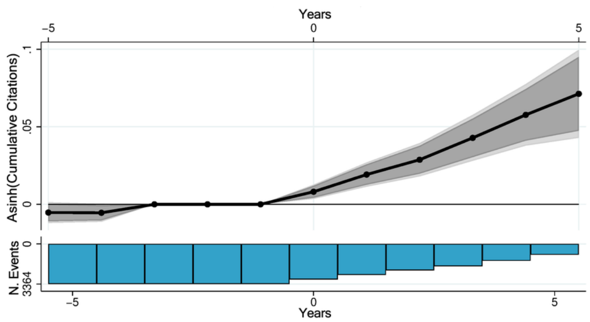
(d) Greenfield Investors' Citations to Country



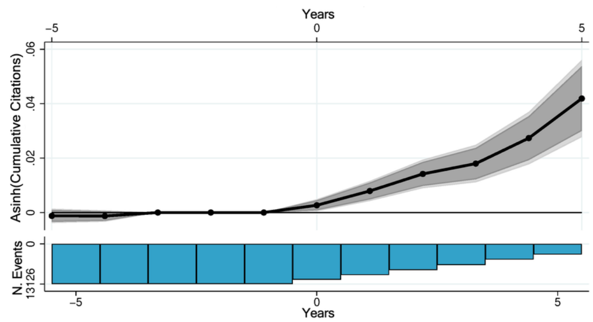
Note: This figure presents the estimated coefficients β_h in Eq. (2). The y-axis in each panel represents the change in the Asinh (inverse hyperbolic sine) transformation of cumulative citations from period $t-1$. Controls and treatment are unchanged relative to Figure 2. We extend the sample to cover all destinations available in PATSTAT.

Figure B.3: Impacts of FDI on Citations, World Sample, Using Citations Within 5 Years of Patent Registration

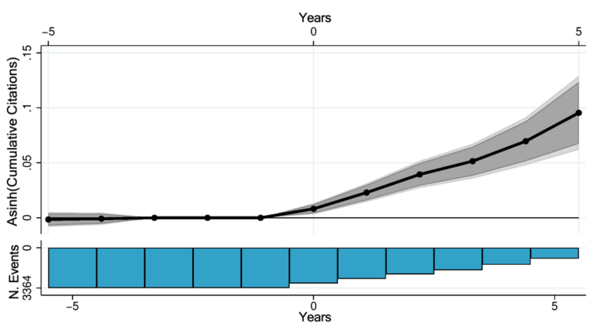
(a) Country's Citations to Brownfield Investors



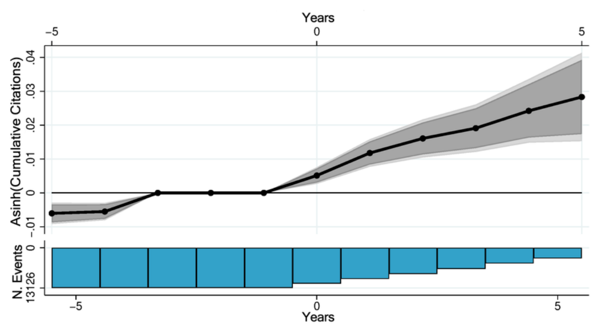
(b) Country's Citations to Greenfield Investors



(c) Brownfield Investors' Citations to Country



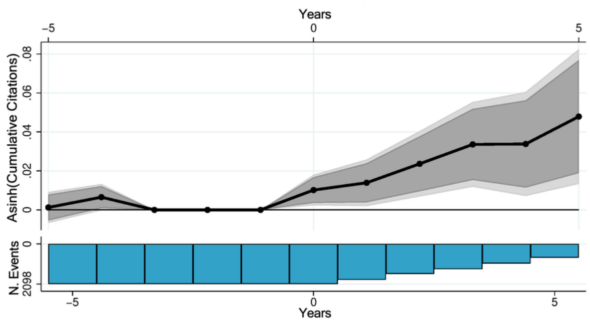
(d) Greenfield Investors' Citations to Country



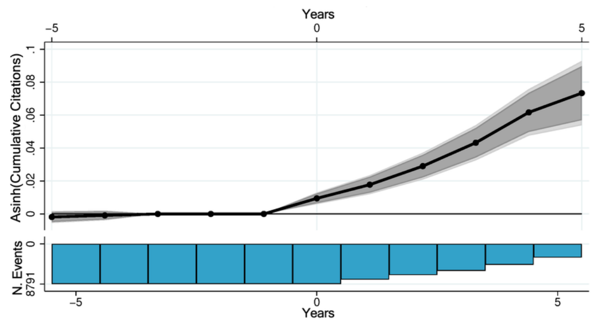
Note: This figure presents the estimated coefficients β_h in Eq. (2). The y-axis in each panel represents the change in the Asinh (inverse hyperbolic sine) transformation of cumulative citations from period $t-1$. Controls and treatment are unchanged relative to Figure 2. We now restrict only to citations received within 5 years of patent registration and shorten the sample to cover 2003-2015 to reduce truncation bias.

Figure B.4: Impacts of FDI on Citations, World Sample, Using 12 Stabilization Periods

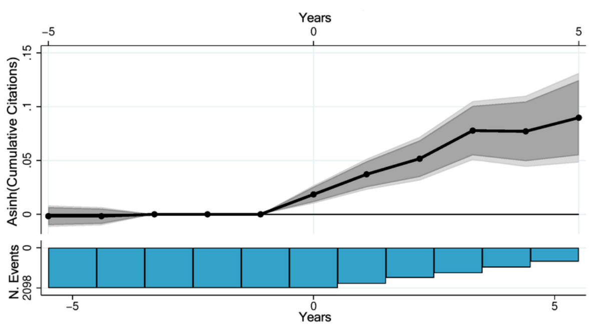
(a) Country's Citations to Brownfield Investors



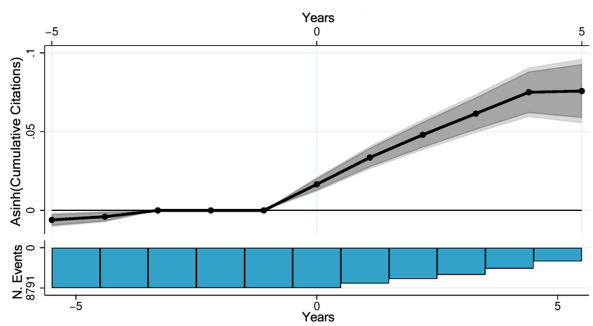
(b) Country's Citations to Greenfield Investors



(c) Brownfield Investors' Citations to Country



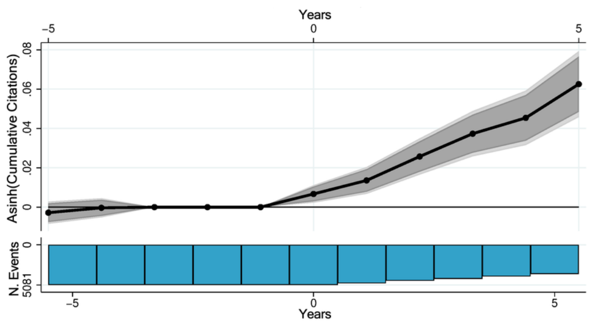
(d) Greenfield Investors' Citations to Country



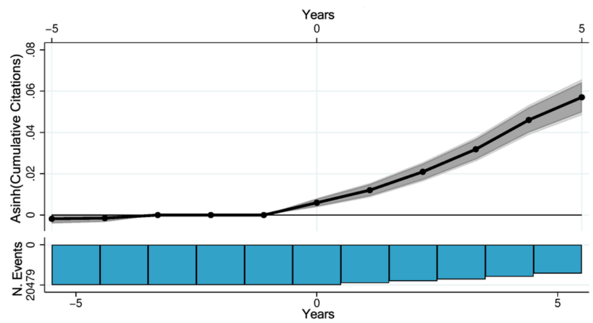
Note: This figure presents the estimated coefficients β_h in Eq. (2). The y-axis in each panel represents the change in the Asinh (inverse hyperbolic sine) transformation of cumulative citations from period $t-1$. Controls, sample selection and treatment are unchanged relative to Figure 2. We now set $L = 12$ instead of $L = 5$.

Figure B.5: Impacts of FDI on Citations, World Sample, Using the $\log(1+x)$ Transformation

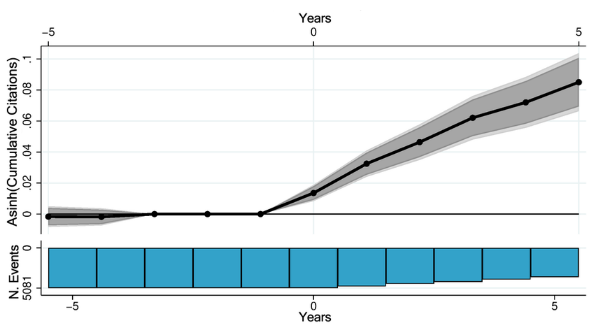
(a) Country's Citations to Brownfield Investors



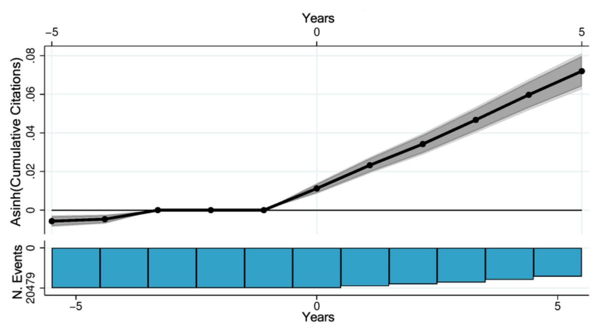
(b) Country's Citations to Greenfield Investors



(c) Brownfield Investors' Citations to Country



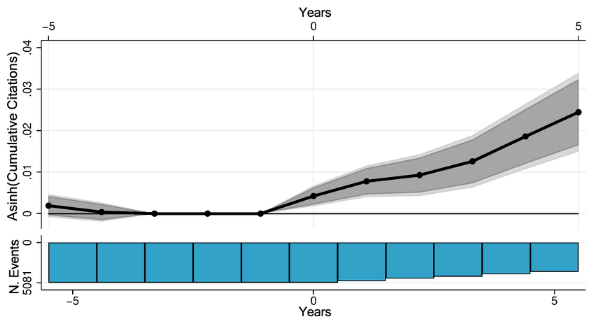
(d) Greenfield Investors' Citations to Country



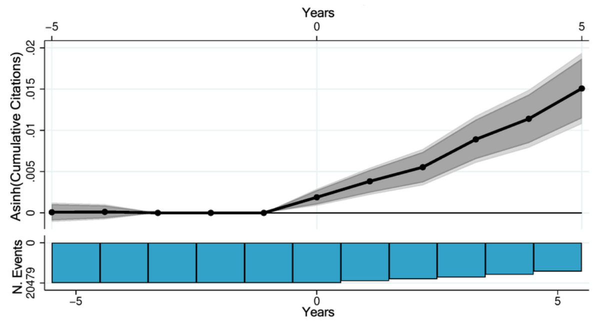
Note: This figure presents the estimated coefficients β_h in Eq. (2). The y-axis in each panel represents the change in the Asinh (inverse hyperbolic sine) transformation of cumulative citations from period $t-1$. Controls and sample selection are unchanged relative to Figure 2. We now use as outcome the the logarithm of one plus the number of cumulative citations since 1995 instead of the asinh transformation.

Figure B.6: Impacts of FDI on Citations, World Sample, Only Citations Between Triadic Patents

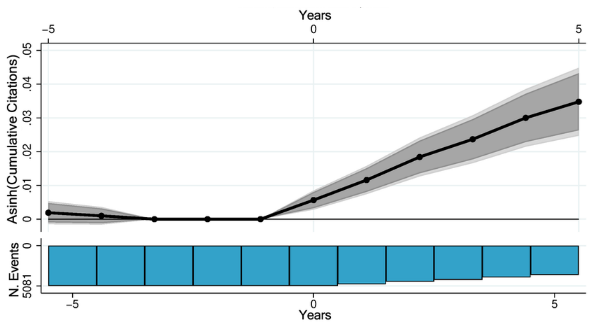
(a) Country's Citations to Brownfield Investors



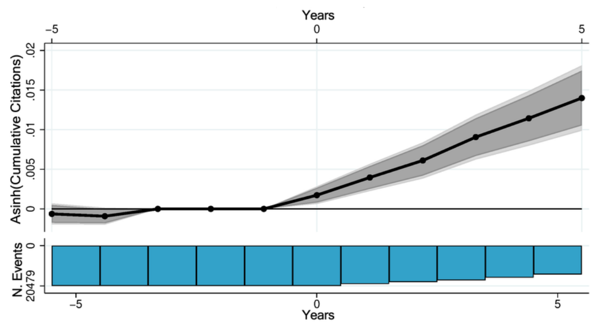
(b) Country's Citations to Greenfield Investors



(c) Brownfield Investors' Citations to Country



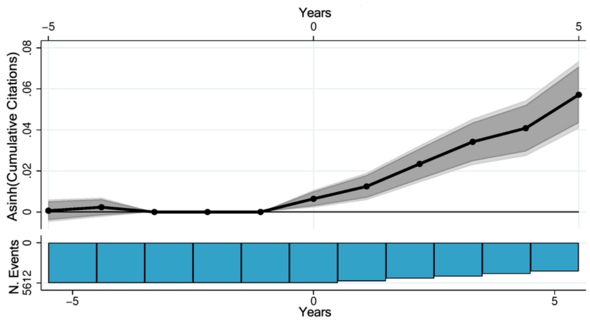
(d) Greenfield Investors' Citations to Country



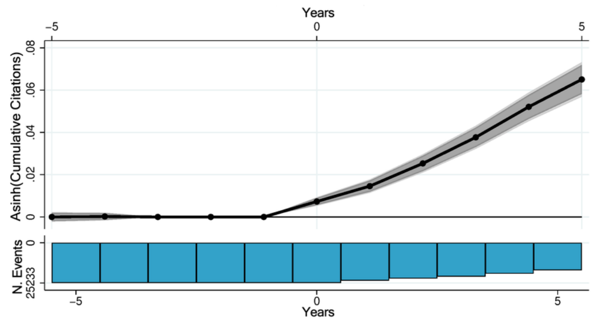
Note: This figure presents the estimated coefficients β_h in Eq. (2). The y-axis in each panel represents the change in the Asinh (inverse hyperbolic sine) transformation of cumulative citations from period $t-1$. Controls, sample selection and treatment are unchanged relative to Figure 2. We now restrict citations to those made and received by triadic patents, that is, patents simultaneously registered at the USPTO, EPO, and JPO.

Figure B.7: Impacts of FDI on Citations, World Sample, Each Investment is a Separate Event

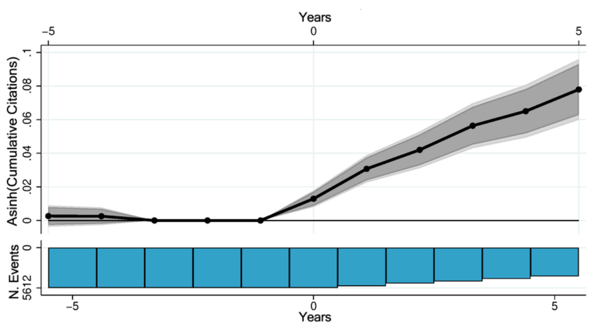
(a) Country's Citations to Brownfield Investors



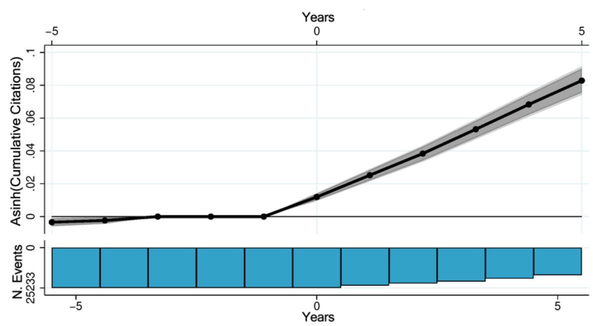
(b) Country's Citations to Greenfield Investors



(c) Brownfield Investors' Citations to Country



(d) Greenfield Investors' Citations to Country

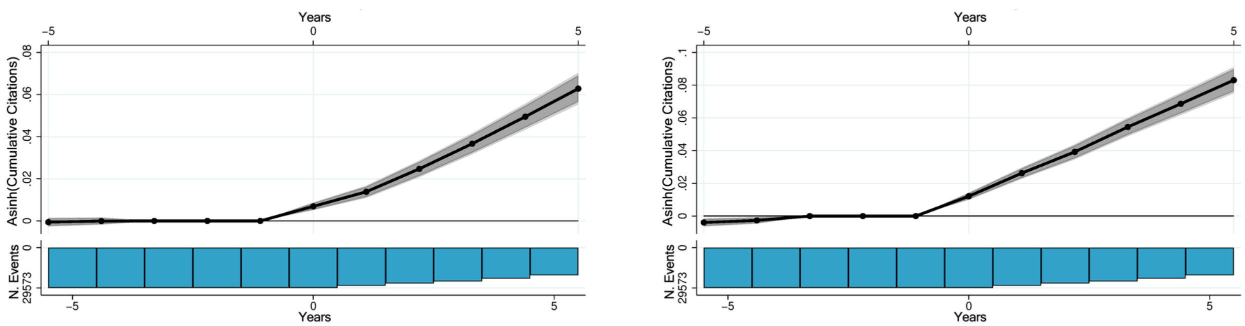


Note: This figure presents the estimated coefficients β_h in Eq. (2). The y-axis in each panel represents the change in the Asinh (inverse hyperbolic sine) transformation of cumulative citations from period $t-1$. Controls and sample selection are unchanged relative to Figure 2. We now use as treatment *any* FDI, not just the first investment observed in our sample period.

Figure B.8: 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

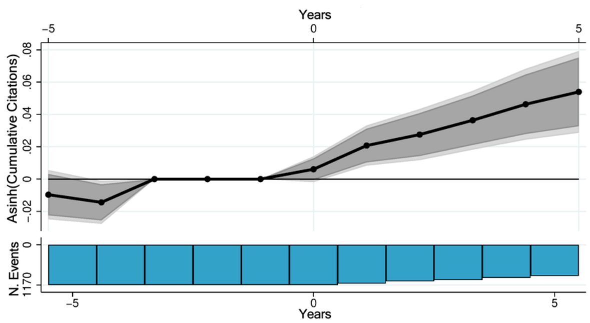
(b) FDI Investors' Citations to Country



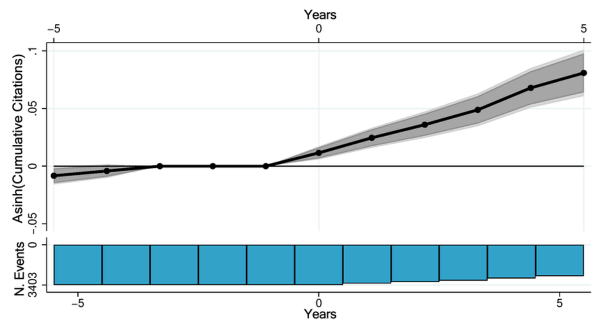
Note: This figure presents the estimated coefficients β_h in Eq. (2). The y-axis in each panel represents the change in the Asinh (inverse hyperbolic sine) transformation of cumulative citations from period $t-1$. Controls and sample selection are unchanged relative to Figure 2. We now use as treatment 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 B.9: Impacts of FDI on Citations, US Sample, Patents Originally Assigned to US Entities.

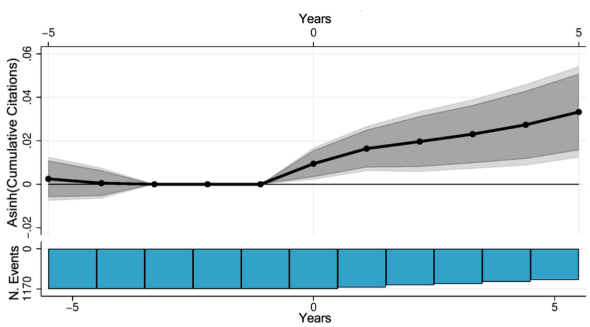
(a) Country's Citations to Brownfield Investors



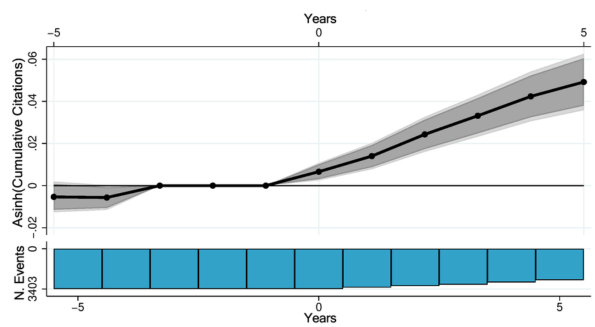
(b) Country's Citations to Greenfield Investors



(c) Brownfield Investors' Citations to Country



(d) Greenfield Investors' Citations to Country



Note: This figure presents the estimated coefficients β_h in Eq. (2). The y-axis in each panel represents the change in the Asinh (inverse hyperbolic sine) transformation of cumulative citations from period $t-1$. The specification and sample selection correspond to Figure 3. 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.