Multinational Production and Innovation in Tandem*

Jin Liu†
New York University

September 26, 2023
Updated regularly, please click here for the latest version.

Abstract

Multinational firms colocate production and innovation by offshoring them to the same host country or region. In this paper, I examine the determinants of multinational firms' production and innovation locations. I find complementarities between production and innovation within host countries and regions, exploiting plausibly exogenous variation in tariffs. In order to evaluate manufacturing reshoring policies, I develop a quantifiable multicountry offshoring location choice model. I allow for rich colocation benefits and cross-country interdependence and prove supermodularity of the model to solve this otherwise NP-hard problem. I find that the effects of manufacturing reshoring policies are nonlinear, contingent upon firm heterogeneity, and they accumulate dynamically.

*I am grateful to my advisors, Martin Rotemberg, Daniel Xu, Boyan Jovanovic, and Sharon Traiberman for their constant guidance and support. I also thank Costas Arkolakis, Pierre Bodéré, Fabian Eckert, Myrto Katsikopoulou, Ezra Oberfield, Michael Peters, Andrés Rodríguez-Clare, Ana Maria Santacreu, Daniel Waldinger, and seminar participants at NYU for useful discussions that helped me improve the paper.

Any views expressed are those of the authors and not those of the U.S. Census Bureau. The Census Bureau has reviewed this data product to ensure appropriate access, use, and disclosure avoidance protection of the confidential source data used to produce this product. This research was performed at a Federal Statistical Research Data Center under FSRDC Project Number 2210. (CBDRB-FY23-P2210-R10532).

†Department of Economics, New York University. E-mail: jl7767@nyu.edu

I Introduction

Multinational firms account for a large share of global production and innovation.[1] These firms produce and innovate in various countries, choosing the optimal locations for each activity. Whether a firm should produce and innovate in the same host country or region turns on the size of their colocation benefits relative to the force separating them. This separating force arises from the fact that countries with low production costs are usually not countries with high returns to innovation (Arkolakis et al., 2018; Antras et al., 2017). In the data, the former force dominates, and firms tend to colocate production and innovation. As shown in Figure 1, large destinations of offshore production are often also large destinations of offshore innovation for U.S. firms.

The colocation benefits are twofold. First, there is synergy between production and innovation as direct interactions reduce communication and coordination costs, spur new ideas, and increase innovation efficiencies (Bahar, 2020; Fort et al., 2020). For instance, biotechnology companies' product designers must understand feasible manufacturing processes in order to invent new medicines. Hewlett-Packard and Texas Instruments both operate laboratories in Singapore, close to their manufacturing facilities, to promote close interactions between the plant and product development engineers during trial runs of new products (Kuemmerle, 1997). Second, having local production can reduce innovation costs. For example, firms often locate their innovation lab and manufacturing plant at the same place to save overhead expenses in terms of rent, utilities, insurance, and supportive infrastructure. Some of the overhead cost sharing, such as for management team and legal and accounting services, happens not only within the country but also across borders within the region (ASEAN, 2017).[2]

[1]Multinational firms constitute nearly 80% of U.S. imports and exports in the year 2000 (Bernard et al., 2009). Sales from foreign affiliates of U.S. manufacturing multinationals exceeded double the value of total U.S. exports. Furthermore, multinational firms are among the most innovation-intensive firms and account for the majority of innovation investment worldwide (see UNCTAD, 2015; Criscuolo et al., 2010). They account for 91% of the innovation investment performed by firms in the U.S. (National Science Board, 2014).

[2]For instance, the accounting firm Deloitte audits and provides accounting services to multiple affiliates of Samsung, including Samsung Electronics in South Korea and several China-based Samsung subsidiaries. Nissan established a regional R&D hub in Thailand that will also serve Indonesia, Philippines, Malaysia and Vietnam.

To delve into the empirical link between production and innovation at the micro level, I show that multinational firms offshore more innovation to a host country if they have more production there and if they produces more in other countries within the region. This pattern could partially be explained by exogenous country characteristics and regional shocks. In order to isolate the *endogenous* factors, I exploit plausibly exogenous variation in tariffs and further establish a causal impact of production on innovation by showing that an increase in a host country's tariff results in reduced production and innovation both within that country and in the surrounding region.

I develop a framework of offshoring location choices to answer several key questions. On a micro level, when multinational firms relocate production, do they reoptimize their innovation locations and how?[3] On an aggregate level, how do recent reshoring policies aiming to bring back U.S. manufacturing affect the global geography of innovation?[4] In particular, will innovation continue to stay in the host country, return to the U.S., or flow to third-party countries when production is reshored?

Allowing for cross-country interdependence is necessary for examining third-country effects, but it can create a hard permutation problem as the optimal choices in one country affects the payoffs to choices in other countries. Firm simultaneously choose the set of countries in which to produce and another set of countries in which to innovate, creating $2^{2\mathcal{L}}$ possible country combinations that grows exponentially with the number of countries $\mathcal{L}$. Furthermore, firms need to solve this choice problem for each possible state (characterized by past locations and productivity shocks) that might be reached in a dynamic setting. The full dynamic optimization problem is NP hard in general; however, the colocation and interdependence forces in my model are estimated to guarantee that the firm's lifetime objective function is supermodular. I employ an innovative algorithm to find the maximum of the supermodular

[3]The location of innovation holds significant importance due to the social benefits and local spillovers of R&D that countries often seek to retain domestically. A concerning trend for the U.S. highlighted in Figure 2b is the increasing amount of foreign R&D, implying a transfer of U.S. innovation's social spillovers to foreign shores.

[4]The CHIPS and Science Act, signed by Trump on August 9, 2022, allocates 280 billion dollars of funding to enhance domestic research and manufacturing of semiconductors, with the objective of cutting reliance on foreign sourcing, particularly from China (White House, 2022). In March 2021, the Biden administration proposed a tax plan titled the "Made in America Tax Plan," which seeks to eliminate incentives for offshore investment and discourage the offshoring of jobs and profits. One of the plan's key provisions involves ending the tax exemption for the first 10 percent return on foreign assets, thereby removing the incentive to offshore tangible assets (U.S. Treasury, 2021).

function in polynomial time.

To measure firms' offshore production and innovation in each foreign country, I use administrative data from the U.S. Census Bureau and explore a previously unused survey module in their Business R&D Survey. This module collects information on firms' annual R&D expenditure and a comprehensive breakdown of that expenditure by foreign countries. I combine this rich data with two identification strategies to establish a causal link between production and innovation. In the first strategy, a shift-share style, firm-specific tariff rate is constructed and used as an instrument for the firm's offshore production. In the second strategy, I exploit plausibly exogenous origin-by-product tariff changes during the Trump Tariffs.

The underlying identification argument for both strategies is that tariff shocks in the host country can affect firms' offshored production by affecting the cost of shipping goods. However, these tariff shocks only affect firms' innovation efforts in that country if there is an interaction between production and innovation. For both identification strategies, I find that increasing tariffs in a host country leads to declining production and innovation both within that country and the region, indicating a positive causal impact that production has on innovation.

To separately identify the synergy and cost-sharing mechanisms that drive colocation, I first pin down the synergy effect by examining how firm productivity is influenced by offshoring location choices. The colocation pattern that remains unexplained by the synergy effect is then attributed to cost sharing. I find that firms offshoring production and innovation to the same host country have significantly greater productivity growth, while those offshoring innovation to countries with no production see minimal productivity growth. The magnitude of the synergy effect is large enough to explain a significant portion of the colocation pattern observed in the data, and thus the cost-sharing mechanism has relatively smaller impact.

Through counterfactual analysis, I first evaluate the relative importance of colocation mechanisms and show that the synergy effect is the most important reason for firms to colocate production and innovation. I also show that my model can generate effects of the Trump Tariffs that match reduced-form estimates. Next, I apply the model to examine the impact of counterfactual trade policies that adversely impact U.S. firms' production offshoring to China. I find significant third-country effects of bilateral trade policies. When the U.S.

increases tariffs on China such that China's production offshoring potential drops by 25%, the likelihood of firms offshoring R&D to China falls by 0.2 percentage point while the corresponding probability for the rest of the world drops by 0.6 percentage points.

Furthermore, I find nonlinear effects of these counterfactual policies on innovation shares, contingent upon firm heterogeneity and the magnitude of the shock. As China has the highest production offshoring potential but modest synergy between production and innovation, many firms with relatively low productivity and capital stock choose to produce in China but innovate elsewhere. Under moderate shocks to production offshoring in China, such firms are predominantly affected, leading to a reduction in the innovation shares of third-party countries and an *increase* in the innovation share of China. However, under large shocks, even the most productive and capitalized firms that also innovate in China are affected. These firms relocate innovations shares away from China towards both the U.S. and third-party countries, and so the net effect is a decline in China's innovation share.

I also show dynamic effects of trade policies that are present in my framework but not captured by previous static models of global production and sourcing. When China's production offshoring potential drops, firms immediately reduce their offshore production and innovation activities, resulting in higher intermediate prices, higher marginal costs, and lower profits. In addition to these standard static losses, firms' productivities are also lowered due to less offshore production and innovation. Lower productivity hinders firms' ability to overcome the fixed and sunk costs associated with offshoring, further exacerbating the reduction in offshore production and innovation. These dynamic losses accumulate over time, with the initial average productivity loss equal to zero and gradually increasing to approximately 0.4% over the course of a decade.

This paper brings together four strands of literatures. First, it examines firms' endogenous innovation choices, which ties into research on R&D and firm performance. Previous studies in this field have empirically estimated the general returns of R&D in terms of productivity gains (Doraszelski and Jaumandreu, 2013), the returns when R&D is concurrently chosen alongside exporting (Aw et al., 2011), and the returns when domestic R&D is coupled with immigrant researchers and imported R&D services (Fan et al., 2022). This paper contributes to this literature in two ways: by incorporating multi-country R&D choices and by considering the direct colocation benefits between production and innovation.

A second contribution lies in the relatively nascent empirical literature on the colocation of production and innovation. Works by Tecu (2013), Lan (2019), Delgado (2020), and Fort et al. (2020) present evidence that there are benefits of locating production and innovation in the same localized area. In particular, Fort et al. (2020) found that patent grants increase by 12 log points when manufacturing and innovation plants are within 5 miles of each other, compared to being more than 60 miles apart. While these studies focus on localized colocation, this paper establishes the colocation pattern between production and innovation in the international landscape, examining multicountry location choices. My second contribution to this literature is to leverage plausibly exogenous tariff variations across a broad range of countries to provide evidence not only for within-location complementarities but also for cross-location complementarities between production and innovation. My findings that increased tariffs in the host country lead to decreased innovation in that country align with Branstetter et al. (2021)'s observation, where a policy shock allowing Taiwanese firms to offshore production to China resulted in a relative decrease in patenting of these firms. In addition to that, I find that increased tariffs in a country lead to decreased innovation in other countries within the same region, suggesting cross-country colocation benefits of production and innovation. Finally, I take a step further to develop a quantitative framework that embeds different sources of colocation benefits and examine their relative importance.

This work engages with a vibrant area of ongoing research that centers on the gains from multinational production, sourcing, and innovation (Rodríguez-Clare, 2010; Arkolakis et al., 2018; Bilir and Morales, 2020; Fan, 2019). Antras et al. (2017) developed a framework for global sourcing and established conditions under which sourcing from different countries becomes complementary. I integrate their structure with elements from the R&D literature, enabling a simultaneous treatment of foreign production and innovation in my model. Bøler et al. (2015) investigate the complementarity between R&D and imports through a scale effect. In comparison to their work, while my model also accounts for the scale effect, its main focus lies in introducing direct interactions between production and innovation through the synergy and cost-sharing mechanisms. Furthermore, my study distinguishes itself from previous works in this literature by solving firms' dynamic location choices. This aspect, which has often been overlooked due to technical complexities, is important to generating rich implications for offshoring and onshoring policies.

Since in my model firms choose a set of production and innovation locations instead of making independent decisions about whether to offshore for each country, this paper also joins a literature that studies large interdependent discrete choice problems at the firm level. Three main approaches have been employed to handle models with this type of problems. The first approach is using the Euler method or moment inequalities to estimate parameters without fully solving the model (Aguirregabiria and Magesan, 2013, 2016; Hsiao, 2021; Holmes, 2011; Morales et al., 2019). Estimates obtained using this approach are often set identified and permit only in-sample counterfactuals that don't involve unobserved states. In contrast, the parameters in my model are point-identified, enabling me to conduct out-of-sample counterfactual analysis. The second approach combines value function approximation and decision process restrictions to handle large state spaces and action spaces, respectively (Sweeting, 2013; Aguirregabiria and Vicentini, 2016). The approximation errors in this approach are often hard to evaluate since we have no accurate solution to compare against.

Most closely related is the third approach, which utilizes complementarities and lattice theory. Jia (2008) pioneered this method in studying the expansion game between two chain stores. Arkolakis and Eckert (2022) formalized a general solution method for combinatorial discrete choice problems with supermodularity or submodularity. However, these papers focus solely on static scenarios, while solving dynamic problems poses significantly more challenges due to the large state space. Recently, Alfaro-Urena et al. (2022) proposed the first algorithm to solve dynamic combinatorial discrete choice with supermodularity. I identify complementarities in my model, prove its supermodularity, and extend their algorithm in several ways to attain the solution. Specifically, my setting incorporates two interrelated dynamic choices with rich complementarities, accommodates a more general context where the static profit function isn't additively separable across countries but only supermodular, and allows the evolution of the unobserved state to be endogenously affected by choices.

The rest of the paper proceeds as follows. Section II describes the data sources and descriptive facts. Section III outlines two empirical identification strategies. Section IV discuss empirical results. Section V outlines the model and proves its supermodularity property. Section VI describes the solution algorithm, estimation steps, and estimation results. Section VII conducts counterfactual exercises. Section VIII concludes.

II Data and Descriptive Facts

II.A Data Sources

One data challenge faced when studying multinational innovation has been the scarcity of comprehensive firm-level data regarding R&D investment by foreign countries. To address this obstacle, I work with administrative records from the U.S. Census Bureau, exploring a previously unused survey module in their Business Research and Development and Innovation Survey (BRDIS).[5] This particular module—with its questionnaire depicted in Figure 3—gathers information on firms' annual R&D expenditure and a comprehensive breakdown of that expenditure by foreign countries.[6]

I use several additional restricted-use micro datasets from the Census Bureau to obtain information on trade transactions and other pertinent firm characteristics for the manufacturing sector. Specifically, I access import transactions at U.S. customs through the Longitudinal Firm Trade Transactions Database (LFTTD).[7] This database furnishes detailed information for each custom transaction, including a firm identifier, product categories based on the 10-digit Harmonized System (HS10) code, value and quantity of goods, origin and destination countries, duties collected for imports, and whether the transaction occurs at arm's length or with related parties.[8] With this dataset, I generate a measure of firm imports by country and calculate firm-specific tariff rates based on the products they import. Furthermore, I draw on the Census of Manufacturing (CMF) and the Annual Survey of Manufactures (ASM) to acquire supplementary data on firms' location, employment, shipments, materials,

[5] The BRDIS is an annual survey conducted by the Census Bureau and sponsored by the National Science Foundation and the National Center for Science and Engineering Statistics. It employs a representative sample of for-profit, nonfarm firms in the United States. The survey focuses on firms with five or more paid employees and at least one establishment (see Foster et al., 2020 for more details).

[6] In this survey module, 40 countries and regions are individually identified and included. Additionally, there is a category labeled "others," which encompasses countries grouped together due to their relatively smaller contribution, representing less than 5% of the total foreign R&D expenses.

[7] The LFTTD is a comprehensive firm-level dataset on export and import transactions, constructed through a collaborative effort between the US Customs and the Census Bureau (see Kamal and Ouyang, 2020 for more details). It encompasses the universe of US trade transactions in goods.

[8] Exporting parties are defined to be related when they own 10 percent or more of the other party. For imports, 19 CFR §152.102(g) defines related persons as (i) members of the same family, (ii) shared officers or directors, (iii) partners, (iv) employers and employees, and (v) a party having a 5% controlling interest in the other. A similar definition of multinationals is used in Bernard and Fort (2015), Antras et al. (2017), and Boehm et al. (2020).

and energy usage.[9]

To augment the analysis, I collect additional data at the country level. Specifically, wage data is obtained from the International Labour Organization (ILO) database. Information regarding language and contiguity is sourced from the Centre d'Etudes Prospectives et d'Informations (CEPII). The Control of Corruption Index is acquired from the Worldwide Governance Indicators database, which is maintained by the World Bank. Finally, population, GDP per capita, human capital index, capital services, and exchange rates are obtained from the Penn World Tables.

Before delving into the analysis, it is important to acknowledge several data caveats about the measurement of offshore production. Due to the lack of available information on the operation of foreign affiliates, except for their trade with U.S. plants, I resort to using a firm's imports from a specific country as a proxy for its offshore production in that location. Although imports represent only approximately half of offshore production,[10] Figure 4 demonstrates a strong correlation between imports and offshore production trends, both in terms of absolute value and growth rate. This suggests that using imports as a proxy for offshore production can still provide accurate insights, particularly when considering relative changes for analysis. Another point to be mindful of is that, although the CMF and ASM sales variables capture all shipments from U.S. plants, there remains a possibility of omitting shipments from foreign affiliates to local foreign customers if they are not rerouted through the U.S. Finally, I do not distinguish between imports from related parties or at arm's length.[11] This choice allows me to capture all types of offshore production, including offshored production to foreign affiliates within the same firm and outsourced production to

[9] The CMF is conducted in years ending in 2 and 7, covering the entire population of manufacturing establishments. The ASM is conducted annually, excluding years ending in 2 or 7, and covers a representative sample of manufacturing establishments with at least one paid employee. Appendix A.1 provides more details on how I use the raw data to construct the firm-level variables necessary for production function estimation.

[10] In 2019, the total import value of U.S. multinational firms was \$2.5T, while the total value of offshore production was \$5.3T. Therefore, imports amounted to approximately 47 percent of the total offshore production in that year. These numbers are calculated based on BEA data.

[11] This approach differs from other studies (e.g. Bilir and Morales, 2020) that use data from the Bureau of Economic Analysis to study multinational production. Although those studies have comprehensive information on the foreign affiliates of U.S. multinational firms, they do not have information on non-offshoring domestic firms to benchmark with and fail to consider production outsourced to foreign firms, a significant component of U.S. offshore production. According to Lakatos and Ohnsorge (2017), a substantial 57% of total U.S. trade occurs at arm's-length between unrelated firms. Specifically , the arm's-length trade constitutes 50% of U.S. imports and 70% of exports. These numbers underscore the importance of including outsourced production in the analysis.

foreign firms.

II.B Firm Sample and Descriptives

I compiled a sample of approximately 36,000 manufacturing firms spanning from 2008 to 2019. As shown in Appendix Table A1, the panel is unbalanced, but large firms, accounting for a substantial portion of the total sales, are consistently surveyed nearly every year. The summary statistics presented in Table 1 reveal that within the subset of firms that invest in R&D, offshore innovation holds considerable significance, representing 23% of total R&D expenditure.

The consideration of multiple locations, as opposed to assuming a single location, is crucial when studying production and innovation offshoring of manufacturing firms. While firms conducting R&D in more than five foreign countries make up only 3% of the observations, they have a significant impact, accounting for 36% of total sales, 70% of worldwide R&D, and 87% of the total offshore R&D, as showed in Appendix Table A2. Similarly, firms importing from more than ten countries represent 95% of the total import value.

I now present several descriptive facts that suggest a positive linkage between offshore production and innovation. Table 2 lists the top five offshore production and innovation destinations for U.S. firms. Germany stands out as the largest offshore R&D destination, accounting for 15% of U.S. firms' foreign R&D expenditure. Mexico leads as the largest origin country for imports, representing 20% of the total import value. Three countries—Germany, China, and Canada—appear in both lists.

The global distributions of U.S. imports and offshore R&D by country are visualized through the world maps in Figure 5. The similarity in the geographical patterns across these two world maps reveals that large destinations for offshore innovation also serve as significant destinations for offshore production. This pattern is demonstrated more clearly in Figure 1, which plots offshore R&D against imports at the host country level and shows a positive correlation.

Firm-country-year observations are grouped into four categories in Table 3, based on whether they are associated with positive R&D and imports. Interestingly, 94% of foreign R&D is conducted in countries where the firm has offshore production. This observation suggests that the return of offshoring only R&D without production to a foreign country is minimal, while that of offshoring both activities can be substantial.

II.C Firm-Level Facts on Offshoring Activities

To better establish a micro-level relationship between production and innovation, I present two novel facts through firm-country-year level regressions. These facts are suggestive of within-country and cross-country complementarities between production and innovation.

Fact 1. *(Within-Country Colocation) Firms engage in more offshore R&D activities in countries from which they import more, and vice versa.*

Fact 2. *(Cross-Country Interdependence) Production and R&D offshoring decisions are interdependent across countries. Specifically, firms engage in more offshore R&D activities in a host country if they produce more in other countries within the surrounding region, and vice versa.*

I regress imports and R&D on each other using a cross-sectional sample from the year 2017. The regression equation is specified as follows:

$$y_{il} = \beta_1 \cdot x_{il} + \beta_2 \cdot x_{iR} + \gamma_i + \gamma_{jl} + \varepsilon_{il}. \tag{1}$$

Here, x_{iR} represents the total value of the independent variable for firm i in region R surrounding country l, but with country l itself excluded. To provide an illustrative example, if we consider l as China and use x and y to represent imports and R&D, respectively, the regression investigates whether there is a correlation between firm i's R&D activities in China and its imports from China, as well as its imports from other East-Asian countries. Furthermore, the regression specification includes firm and country-industry fixed effects, denoted by γ_i and γ_{jl}.[12]

[12] In addition to the regression analysis conducted using the 2017 sample, I have also performed a separate set of regressions using data from all years in the sample. This extended analysis controls for firm-year and country-industry-year fixed effects, and the regression equation is specified as follows: $y_{ilt} = \beta_1 \cdot x_{ilt} + \beta_2 \cdot x_{iRt} + \gamma^{it} + \gamma^{jlt} + \varepsilon_{ilt}$. The results obtained from this extended analysis are highly similar to those from the 2017 sample. For detailed findings, please refer to the Appendix Table A5.

Regression results for various combinations of the extensive and intensive margins of imports and R&D are reported in Table 4. Panel A focuses on regressing R&D on imports, while Panel B investigates the opposite direction. In both panels, the coefficient estimate, $\hat{\beta}_1$, is found to be significant and positive, indicating that a firm is more likely to have R&D activities in countries where it also engages in more production, and vice versa. Specifically, the coefficient estimate for x_{ilt} is 0.0195 in Column (1) of Panel A, where the extensive margins of both activities are considered. This suggests that the probability of a firm conducting R&D in a host country increases by 1.95 percentage points if the firm also engages in production there. Since the baseline probability of conducting R&D in a host country is only 1.3 percentage points (see Table 1), the presence of offshore production more than doubles the likelihood of a firm engaging in R&D in a foreign country. This stylized fact is highly suggestive of the presence of colocation benefits between production and innovation.

Another observation from Table 4 is the robust and positive estimate of β_2, which establishes the second fact. Firm's offshore innovation in a particular host country is not only positively correlated with its offshore production in that same country but also with its offshore production in other neighboring countries within the region. Specifically, when a firm engages in production in the neighboring countries, the probability of it conducting R&D in the focal host country increases by 0.15 percentage points. This represents a 12% increase in the probability of offshoring R&D, relative to its baseline value. This fact suggests that firms' offshoring decisions for production and innovation are correlated across host countries.

III Empirical Strategy

The two descriptive facts highlight positive within-country and cross-country correlations between offshore production and innovation. These correlations may be attributed to unobserved affiliate traits—such as management skills—and correlation in country characteristics (Manski, 1993), as well as inherent connections between production and innovation. To disentangle the latter, I will now elaborate on two identification strategies. The first strategy leverages a firm-specific tariff rate—constructed based on the firm's import product bundle in a shift-share style—as an instrument for production offshoring. The second strategy exploits unexpected tariff line changes that occurred during the Trump Tariffs as exogenous shifters of offshore production.

III.A Firm-Specific Tariff Rates: IV Strategy

Firms import different goods from different origin countries, subjecting them to distinct tariff rates that react differentially to shifts in tariff lines. The import tariff rate a U.S. firm faces directly affects its cost of shipping goods from the host country, and thereby influencing production offshoring. However, import tariffs should not directly affect foreign R&D expenditures, unless an interaction between production and innovation exists. This rationale supports the idea of using the firm-specific import tariff rate as an instrumental variable for offshore production in investigating its impact on offshore innovation.

This firm-specific tariff rate, denoted as T_{ilt}, is designed to capture the effective tariff rate a firm would face in each country if it had continued importing the same bundle of goods from a fixed prior period. By holding the firm's import product bundle constant across origin countries and time periods, I can rule out selection biases at the country level, as well as any potential endogenous response of the import product bundle to changes in tariff rates.

T_{ilt} is computed as a weighted average of product-country-level tariff rates (T_{glt}), with the weights being the firm's initial import value shares across goods (s_{igt_0}) from all origin countries during a prior time period (t_0):[13]

$$T_{ilt} = \sum_g s_{igt_0} T_{glt}.$$

Goods g are defined at the 10-digit HS code level. The study period for this exercise is from 2013 to 2019, and I use a prior period from 2008 and 2012 to calculate s_{igt_0}.

Aside from the firm-specific tariff rate in the host country, another useful variable is the firm-specific average tariff rate within the region, excluding the host country itself. This variable can serve as an instrument for the firm's total offshore production to other countries

[13] T_{glt} is calculated by averaging transaction-level import data from the LFTTD.

within the host region. It is denoted as T_{iRt} and calculated as follows:

$$T_{iR(l)t} = \frac{1}{\sum_{l' \neq l} c_{ll'} M_{l'}} \sum_{l' \neq l} c_{ll'} M_{l'} T_{il't}, \tag{2}$$

where $c_{ll'}$ is a dummy variable that equals one if countries l and l' are in the same region, and $M_{l'}$ is the aggregate import value from country l' over all sample years. For instance, if we consider l to represent China, then T_{ilt} would represent the firm i's specific tariff rate in China, while T_{iRt} would represent the weighted average tariff rate the firm faces in other East-Asian countries, with China excluded from the calculation.

To implement this IV strategy, I consider both a reduced-form regression,

$$y_{ilt} = \beta_1 \cdot T_{ilt} + \gamma_{it} + \gamma_{lt} + \nu_{ilt}, \tag{3}$$

where I directly regress offshore production and innovation on the instrument, as well as the following two-stage regressions:

$$\begin{aligned}
\text{R\&D}_{ilt} &= \beta \cdot \widehat{\text{Imp}}_{ilt} + \gamma_{it} + \gamma_{lt} + \varepsilon_{ilt}, \\
\text{Imp}_{ilt} &= \rho \cdot T_{ilt} + \gamma_{it} + \gamma_{lt} + \nu_{ilt},
\end{aligned}$$

where I regress imports on the tariff rate in the first stage, and R&D on the predicted imports in the second stage. I include firm-year fixed effects (γ_{it}) and country-industry fixed effects (γ_{lt}) in all of these regressions.

III.B Trump Tariffs as A Quasi-Experiment

Next, I will present the second identification strategy, which exploits the product-country-specific tariff shocks generated by the Trump Tariffs policy. Let me defer the comparison of these two identification strategies to the end of this section.

In an effort to tackle the trade deficit, the U.S. implemented a series of tariff increases on specific goods and countries in the years 2018 and 2019.[14] Consequently, major trading

[14] Please refer to Fajgelbaum et al. (2020, 2022) for information on various stages of Trump tariffs until

partners of the U.S. retaliated with their own tariffs, escalating trade tensions. As estimated by Fajgelbaum et al. (2020), these U.S. tariff changes led to an overall increase in the average tariff rate from 2.6% to 16.6% for a total of 12,043 goods. These goods accounted for approximately $303 billion (12.7%) of the annual imports into the U.S.

The Trump Tariffs present a compelling quasi-experiment to examine the influence of offshore production on offshore innovation because the unexpected tariff shocks at the product-country level should be uncorrelated with firm-country-year-specific unobservables, such as the management skill of foreign affiliates. To leverage this quasi-experiment, I obtain detailed data on tariff changes at the 10-digit HS code and country level, which has been compiled by Fajgelbaum et al. (2020, 2022) using public schedules from the U.S. International Trade Commission (USITC). As depicted in Figure 6, the Trump Tariffs impacted many countries, displaying significant variation in the number of affected goods and the average effective tariff increase across these countries. Notably, while China had the highest number of affected products, its effective tariff increase in the affected goods does not rank among the highest.

I now define the treated firm-country pairs during the Trump Tariffs. For each combination of firm i and country l, I compile a list of goods that firm i had imported from country l during the five-year period prior to the enactment of the Trump Tariffs. If any of these imported goods were affected by the tariff changes in 2018 and 2019, I designate the firm-country pair as treated and assign a value of one to the dummy variable Treat_{il}. Conversely, if none of the previously imported goods from country l were impacted by the tariff changes, the firm-country pair is classified as untreated, and Treat_{il} is assigned a value of zero. For robustness checks, I also explore alternative measures of the firm's degree of treatment, including the fraction and value share of products affected, as well as the effective amount of tariff increases for affected products. It is worth noting that the analysis around the Trump Tariffs is confined to firm-country pairs where firm i had engaged in imports from country l during the five-year prior period.

I will conduct event study and difference-in-differences regressions to assess the relative impact of the Trump Tariffs on the treated group. The event study regression is specified as

2019, and Bown (2020) for an up-to-date chart of US-China Trade War tariffs.

follows:

$$y_{ilt} = \sum_{t=2014:2019} \beta_t \cdot \text{Treat}_{il} \cdot \text{Year}_t + \gamma_{il} + \gamma_{lt} + \varepsilon_{ilt}. \tag{4}$$

This regression examines the relationship between the outcome variable y_{ilt} and the treatment indicator interacted with year indicators from 2014 to 2019. The difference-in-differences regression is specified as below:

$$y_{ilt} = \beta \cdot \text{Treat}_{il} \cdot \text{Post}_t + \gamma_{il} + \gamma_{lt} + \varepsilon_{ilt}. \tag{5}$$

This regression instead focuses on comparing the baseline (pre treatment) and endline (post treatment) outcomes. The Post dummy is set to one in 2019 and zero between 2014 and 2017. Excluding the year 2018 from the sample in the DID regression ensures a clean comparison between the baseline and endline outcomes. This choice is motivated by the fact that R&D decisions typically require time to respond to shocks and helps abstract away from various middle stages of tariff increases. In both regressions, I include firm-country and country-year fixed effects (γ^{il} and γ^{lt}).[15] Standard errors are always clustered at the firm level.

III.C Comparison of Two Strategies

Due to the sample restrictions in defining the treatment dummy for the Trump Tariffs quasi-experiment, the sample used in the event study is limited to 0.2 million observations at the firm-country-year level. In contrast, the first instrumental strategy does not face such limitations and therefore utilizes a much larger sample of 1.5 million observations at the firm-country-year level.

However, there are two potential validity concerns associated with the first instrumental strategy. Firstly, firms may anticipate tariff changes several years in advance and endogenously respond, for instance, with different investment levels. Additionally, changes in tariff schedules, particularly those resulting from Free Trade Agreements, may include Intellectual Property (IP) terms that could directly affect innovation incentives (Santacreu, 2021). The latter concern challenges the validity of the instrumental strategy only if the confounding

[15] I do not include firm-year fixed effects to retain observations for one-country firms, and because tariff shocks at the product-country level are uncorrelated with firm-year characteristics.

IP factors operate at the product level. This is because the instrument is constructed using product-level tariff variations, and the regressions have already controlled for country-year fixed effects, which account for country-level IP changes. These two concerns are of lesser relevance to the second strategy since the product-country-level tariff changes during the Trump Tariffs were unanticipated, and IP issues were not a part of the trade war.

Finally, the first instrumental strategy incorporates a broader spectrum of tariff variations dating back to 2013. This mainly includes tariff changes originating from previously built-in tariff reduction schedules, such as those within the U.S. Free Trade Agreements with Chile, Dominican Republic, Morocco, Peru, and Singapore. Conversely, the second strategy exclusively focuses on tariff variations arising from the Trump Tariffs policy. It would provide greater reassurance if the two identification strategies, which leverage different sources of variation, produce consistent and robust estimation results for the impact of production on innovation.

IV Empirical Results

IV.A Evidence from Using Firm-Specific Tariff Rates as IV

I find that higher tariffs lead to a decrease in both a firm's production and innovation within the host country. In Table 5, I report the reduced-form regression results in Panel A and the two-stage least squares results in Panel B. The estimates in Columns (1) and (4) of Panel A reveal that a 10 percentage point decrease in the tariff rate faced by firm i in country l (i.e., $\Delta T_{ilt} = 0.1$) is associated with a 19% increase in firm i's imports from country l and a 2.4% increase in its R&D expenditure in the same country.[16] The estimate in Column (3) of Panel B suggests that when firm i doubles its imports from country l, its R&D expenditure in the same country increases by 12.5%.[17] This positive effect of production on innovation indicates the presence of colocation benefits between these two activities.

[16]The estimated coefficient for imports corresponds to a trade elasticity of approximately 2, which falls within the plausible range of estimates found in the literature, albeit slightly on the lower end (Imbs and Mejean, 2015; Marquez, 2002).

[17]The first-stage F statistic is 61.93, indicating strong instrument relevance (Stock et al., 2002).

To further explore the empirical evidence for cross-country interdependence in offshoring activities, I ask the following question: Do import tariffs for neighboring countries within the region also impact innovation offshored to the host country? To answer this question, I extend the previous reduced-form regression in Equation (3) by introducing an additional independent variable, T_{iRt}, which is the regional tariff rate constructed in Subsection III.A. To illustrate this generalized regression, consider China as the host country. The new regression then investigates whether offshored activities in China are influenced not only by the tariff rate that firm i faces in China (T_{ilt}) but also by the tariff rate in the broader East Asian region, excluding China (T_{iRt}).

I find that the average regional tariff, excluding that from the host country, also has a negative effect on a firm's offshored production and innovation within the host country. Table 6 presents the results for the generalized regression. The coefficient estimate for the host country's tariff rate remains robust. Moreover, the coefficient for the regional tariff rate is estimated to be significant and negative. These results collectively indicate that offshored production and innovation are adversely affected not only by a tariff increase in the host country but also by a tariff increase in other countries within the host region. The findings that these offshoring decisions in one country are influenced by exogenous shocks occurring in other countries provide compelling causal evidence for cross-country interdependence in production and innovation offshoring.

IV.B Evidence from the Trump Tariffs Quasi-Experiment

As introduced in Section III.B, the Trump Tariffs implemented during 2018 and 2019 offer exogenous product-country level tariff increases that affect production offshoring without directly influencing innovation offshoring. Therefore, if innovation offshoring also responds to these tariff changes, it serves as evidence for the causal impact of production on innovation.

I find here that the Trump Tariffs have negative effects on both offshore production and innovation in the host country. Figure 7 presents the results from the event study regressions specified in Equation (4). The treated firms experience a 10 percent reduction in imports from the host country. Furthermore, the probability of conducting R&D in the host country declines by 1.3 percentage points, and among firms that still conduct R&D, the amount of

R&D expenditure decreases by 15 percent in the year 2019.[18]

Table 7 presents the results from the difference-in-differences regressions. Estimates in the second to fourth rows correspond to alternative measures of the firm's degree of treatment and show robust negative coefficient estimates. These results provide compelling evidence of the negative causal impact of the tariff increase resulting from the Trump Tariffs on offshored production and innovation activities, which further indicates a positive causal effect of production on innovation.

Taking stock of the empirical evidence derived from both identification strategies in this section, I find that a firm's offshored innovation to a host country is positively impacted by its offshored production to the same country, as well as by its offshored production to neighboring countries within the host region. These findings strongly suggest the existence of colocation benefits and cross-country interdependence for production and innovation. To quantify the underlying forces driving these colocation benefits and cross-country interdependencies, and to evaluate counterfactual trade policies, I develop a quantitative framework of dynamic offshoring location choices in the next section.

V A Model of Dynamic Offshoring Location Choice

This section introduces a dynamic, partial-equilibrium model that characterizes firms' decision making process regarding their choice of countries for both production and innovation. The primary goal is to incorporate two important features observed in the empirical section: colocation benefits and cross-country interdependencies between production and innovation.

Firms import finished and intermediate products from foreign countries, which may be manufactured by the firm's foreign affiliates or sourced from other foreign suppliers. These imported goods, together with the domestically produced goods, are combined in the Constant Elasticity of Substitution (CES) style into the firm's final product, which is then sold

[18] One might think that China and the semiconductor industry are the primary targets of the Trump Tariffs, and therefore, they may be driving the majority of the results. However, regressions excluding China and the semiconductor industry (NAICS code 334) from the sample show highly robust, and in some cases, even slightly larger effects. This implies that the estimated causal effects within this quasi-experiment are not significantly driven by either China or the semiconductor industry.

in the global market.

Firms solve an infinite-horizon combinatorial discrete choice problem, wherein they choose the country bundles for production and innovation in each period, based on their past location bundles. During this process, they face sunk and fixed costs associated with these two activities. The fixed cost of innovation is directly influenced by whether the firm produces in the region. The bundle of production locations determines the price index of imported goods, and in conjunction with the bundle of innovation locations, influences the firm's future productivity. Firms take into account the cross-country interdependence and recognize how their decisions made in one country can affect their optimal choices in other countries.

My model highlights three mechanisms that generate colocation benefits between production and innovation. Firstly, to account for the spillovers from local production to innovation, the model allows for a higher return to innovation in a host country if the firm also produces there. Secondly, the model allows for the sharing of overhead costs between production and innovation by assuming that having production in the surrounding region can reduce the fixed cost of innovation in the host country. Thirdly, an additional complementarity between production and innovation arises from a scale effect. For instance, when the firm engages in R&D activities in a host country, it experiences higher productivity, leading to a greater payoff for producing in that country.[19]

My model also introduces three sources of cross-country interdependencies. To begin, imported goods from different countries substitute for each other via the cost function, reflecting the firm's flexibility in choosing among various production locations based on relative costs. Secondly, production and R&D in different countries are complementary due to the firm-level scale effect. For instance, when the firm produces in a particular country, it experiences lower costs of intermediate inputs, leading to increased demand and higher payoffs for producing and innovating in another country. These two mechanisms operate on a global scale, but I also incorporate a region-specific force by allowing for a reduced cost of innovation in a host country when the firm produces in neighboring countries within the same region. By incorporating these elements of cross-country interdependencies, the model can generate third-country effects of bilateral trade policies, setting it apart from many papers

[19] This scale effect is transmitted through the firm's productivity and import prices, thereby generating consistent impacts across all affiliates of the firm rather than being specific to one foreign country.

on multinational production that have traditionally assumed independence across countries for technical simplicity.

V.A Static Demand and Production

In the static part of my model, I establish the demand and cost structures and outline the firm's production problem. Each firm is denoted by i, location by l, industry by j, and time period by t.

Market Demand. I assume a monopolistic competition structure, where the demand of firm i is given by

$$q_{it} = Q_{jt} \cdot \left(\frac{p_{it}}{P_{jt}}\right)^{-\eta} = \Phi_{jt} \cdot (p_{it})^{-\eta}.$$

Here, η represents the elasticity of substitution between products offered by different firms, and Φ_{jt} captures the market conditions for the industry in which firm i operates.

Production Cost. Following Aw et al. (2011), Berry et al. (1995), Roberts et al. (2018), and Piveteau (2021), I assume that the short-run unit production cost is independent of output levels and specified as a function of cost shifters:

$$\ln c_{it} = \beta_0 + \beta_k \cdot \ln k_i + \beta_w \cdot \ln w_{jt} + \beta_m \cdot \ln p^m_{it} - \omega_{it}.$$

To incorporate heterogeneity arising from the production side of the firm's activities, I allow the log of marginal cost to depend on the firm's exogenous capital stock (k_i), wage rate in the industry (w_{jt}), the price index of intermediate goods (p^m_{it}), and the unobserved Hicks neutral productivity (ω_{it}). As I describe below, intermediates may potentially have a domestic and an imported component that are combined according to a CES aggregator. The intermediate price and productivity are endogenously affected by the firm's production and innovation location choices.

Based on the demand structure and the marginal cost function, the firm determines the optimal quantity of its final good, which gives rise to the following revenue function:

$$\ln R_{it} = (1-\eta)\ln\left(\frac{\eta}{\eta-1}\right) + \ln \Phi_{jt} + (1-\eta)\left(\beta_0 + \beta_k \ln k_{it} + \beta_w \ln w_{jt} + \beta_m \ln p^m_{it} - \omega_{it}\right).$$

The firm's profit is proportional to its revenue, with a constant markup equal to $\eta/(\eta-1)$:

$$\pi_{it} = \frac{1}{\eta} \cdot R_{it}\left(\omega_{it}, k_i, p_{it}^m, \Phi_{jt}\right).$$

Foreign Production. Without loss of generality, I assume that all firms obtain intermediates from the domestic market. This assumption is supported by the observation that all firms in my sample purchase domestic materials.

In the static problem, the firm takes the set of production locations as fixed and choose the amount of goods to make from each chosen location. Intermediates from different locations are aggregated via the CES structure:

$$m_{it} = \left(\sum_{l\in\mathcal{L}} y_{ilt} \cdot m_{ilt}^{\frac{\rho-1}{\rho}}\right)^{\frac{\rho}{\rho-1}},$$

where y_{ilt} is a dummy variable that equals one if firm i produces in country l in period t, and ρ is the elasticity of substitution between goods from different countries.

The unit cost of good from a specific location ($p_{m,ilt}$) is determined by the local wage level w_{lt} in host country l, the shipping cost τ_{lt} between the U.S. and country l, and the U.S. import tariff rate t_{lt} (defined as $1+T_{lt}$) for country l:

$$p_{m,ilt} = w_{lt}\tau_{lt}t_{lt}.$$

As a result, the price index for the aggregated intermediates of firm i is given by

$$p_{it}^m = \left(\sum_{l\in\mathcal{L}} y_{ilt}\left(w_{lt}\tau_{lt}t_{lt}\right)^{1-\rho}\right)^{\frac{1}{1-\rho}}.$$

It is worth noting that the aggregate price index of intermediates depends both on the set of production locations and on the unit cost of good at each location. This CES aggregation of intermediates can be microfounded by considering a continuum of good varieties and assuming a Fréchet distribution of production efficiencies across countries. The equivalence of these two approaches is demonstrated in Appendix B.1.

Finally, I define θ_{lt} to be the *production offshoring potential* of country l in period t, which represents this location's average cost advantage in manufactured goods and is calculated as $\theta_{lt} = (w_{lt}\tau_{lt}t_{lt})^{1-\rho}$. Furthermore, I define Θ_{it} to be the *production offshoring capability* of firm i in period t, capturing the firm's ability to produce more at lower-cost countries and computed as $\Theta_{it} = (p_{it}^{m})^{1-\rho}$.

V.B Dynamic Location Choices

In this subsection, I will introduce the dynamic component of my model, with a specific focus on the location choice problem.

Innovation Effort and Evolution of Firm Productivity. The productivity of firm i is governed by a Markov process, which takes into account its past productivity, an i.i.d. shock, and its production and innovation decisions in all locations. The equation capturing this productivity evolution is as follows:

$$\omega_{it} = \alpha_0 + \alpha_1\omega_{it-1} + \sum_{l}\left[1 + X'_{lt-1}\mu\right] \cdot [\beta_1 r_{ilt-1} + \beta_2 y_{ilt-1} r_{ilt-1}] + \xi_{it}. \tag{6}$$

In this equation, r_{ilt} represents whether firm i conducts R&D in country l at time t, and y_{ilt} denotes whether firm i produces in country l at time t. The coefficient β_2 captures the synergy effect that local production can increase innovation efficiency. Moreover, the return to R&D varies by country and is contingent upon the specific country characteristics denoted by X_{lt}. The component ξ_{it} follows a normal distribution with mean zero and variance σ_{ξ}^2, thereby capturing the randomness in innovation.

Sunk and Fixed Costs. Firms incur initial sunk costs, denoted as ϕ_s^p for production and ϕ_s^r for innovation, when they start offshoring these activities to a foreign country for the first time. If they have previously undertaken such offshoring, they instead pay the fixed costs, ϕ_f^p for production and $\phi_{f,ilt}^r$ for innovation. The fixed cost for innovating in country l is given by:

$$\phi_{f,ilt}^r = \phi_f^r - \lambda_1 \max_{l'}\{c_{ll'} y_{il't}\}.$$

Here, $c_{ll'}$ is a dummy variable that takes the value of one if country l' and country l are in the same region. The fixed cost for innovation is reduced when the firm has nearby production plants, and the degree of reduction is determined by the parameter λ_1.

Timing Assumption. The timing assumptions for the model are as follows: (1) At the beginning of period t, the firm observes its state vector, which includes the location bundles for production and innovation from the previous period, the current-period productivity, and other exogenous demand and cost shifters:

$$\boldsymbol{s}_{it} = \left(\{y_{ilt-1}\}_l, \{r_{ilt-1}\}_l, \omega_{it};\ k_{it}, \Phi_{jt}\right).$$

The value function $V_{it}(\boldsymbol{s}_{it})$ is defined at this stage. (3) The firm is aware of the fixed and sunk costs associated with each choice and selects its optimal location bundles $\boldsymbol{y}_{it}$ and $\boldsymbol{r}_{it}$. (4) Static profits $\pi_{it}(\boldsymbol{y}_{it}, \omega_{it})$ are realized, and costs dependent on $\boldsymbol{y}_{it}, \boldsymbol{y}_{it-1}, \boldsymbol{r}_{it}, \boldsymbol{r}_{it-1}$ are paid. (5) Productivity shocks are realized, and a new state is formed:

$$\boldsymbol{s}_{it+1} = \left(\boldsymbol{y}_{it}; \boldsymbol{r}_{it}; \omega_{it+1}|\omega_{it}, \boldsymbol{y}_{it}, \boldsymbol{r}_{it}; k_{it+1}, \Phi_{jt+1}\right).$$

Dynamic Problem. The Bellman equation for the dynamic problem is given by

$$\begin{aligned} V_{it}\left(\boldsymbol{s}_{it}\right) = \max_{\boldsymbol{y}_{it}, \boldsymbol{r}_{it}} \Bigg\{ & \pi_{it}(\boldsymbol{y}_{it}, \omega_{it}) - \sum_l \left[(1 - y_{ilt-1}) \cdot y_{ilt} \cdot \phi_s^p + y_{ilt-1} \cdot y_{ilt} \cdot \phi_f^p\right] \\ & - \sum_l \left[(1 - r_{ilt-1}) \cdot r_{ilt} \cdot \phi_s^r + r_{ilt-1} \cdot r_{ilt} \cdot \phi_{f,ilt}^r\left(\boldsymbol{y}_{it}\right)\right] + \beta \mathbb{E}_\xi V_{it+1}\left(\boldsymbol{s}_{it+1}|\omega_{it}, \boldsymbol{y}_{it}, \boldsymbol{r}_{it}\right) \Bigg\}. \end{aligned} \tag{7}$$

The firm's objective is to maximize the expression inside the curly braces, which consists of the current-period profit, costs associated with location choices, and the discounted expected value for future periods.

The state transition rule for the dynamic model is as follows. Regarding productivity, the next-period productivity ω_{it+1} is determined based on the current-period productivity ω_{it}, the firm's R&D location choices $\boldsymbol{r}_{it}$, and the production location choices $\boldsymbol{y}_{it}$. For location portfolios, the optimal choices of location bundles, $\boldsymbol{y}_{it}$ and $\boldsymbol{r}_{it}$, serve as the new states in the next period.

V.C Supermodularity Property

The dynamic model characterized by the Bellman equation (7) is an NP-hard problem with huge state and action spaces. However, in this section, I will derive a condition under which the complementarity forces in my model outweigh the substitutability forces, resulting in the firm's lifetime objective function being supermodular. This condition holds empirically, as indicated by my parameter estimates in Section VI.B. Since maximizing a supermodular function can be achieved in polynomial time, I will be able to utilize the algorithm outlined in Section VI.A to effectively solve this otherwise unsolvable dynamic location choice problem.

To discuss the supermodularity property of the model, it is helpful to reframe the firm's dynamic problem as a lifetime planning problem. Let us define the net-of-cost static profit function in period $t \geq 1$, Π_t, as the variable profit net of fixed and sunk costs:

$$\begin{aligned}\Pi_t\left(\omega_{it}, \boldsymbol{y}_{it}, \boldsymbol{y}_{it-1}, \boldsymbol{r}_{it}, \boldsymbol{r}_{it-1}\right) = \pi_{it}(\boldsymbol{y}_{it}, \omega_{it}|\Phi_{jt}) - \sum_l \left[(1 - y_{ilt-1}) \cdot y_{ilt} \cdot \phi_s^p + y_{ilt-1} \cdot y_{ilt} \cdot \phi_f^p\right] \\ - \sum_l \left[(1 - r_{ilt-1}) \cdot r_{ilt} \cdot \phi_s^r + r_{ilt-1} \cdot r_{ilt} \cdot \phi_{f,ilt}^r\left(\boldsymbol{y}_{it}\right)\right].\end{aligned}$$

I further define the firm's expected lifetime payoff function as a function of its decision rule $\boldsymbol{o}_i$:

$$\Pi_0\left(\boldsymbol{o}_i|\boldsymbol{y}_{i,-1}, \boldsymbol{r}_{i,-1}, \omega_{i,-1}\right) = \mathbb{E}_{\boldsymbol{z}} \sum_{t=0}^{\infty} \Pi_t\left(\omega_{it}\left(z^t, \{\boldsymbol{o}_i\left(z^\tau\right)\}_{\tau=0}^{t-1}\right), \boldsymbol{o}_i\left(z^t\right), \boldsymbol{o}_i\left(z^{t-1}\right)\right).$$

Here, $\boldsymbol{z} = \{z^t\}_{t=0}^{\infty}$ represents a full history of productivity shocks, where $z^t = (\xi_1, \xi_2, \ldots, \xi_t)$. Ω is the space of $\boldsymbol{z}$ that encompasses all possible histories of productivity shocks. $\boldsymbol{y}$ is a point in $\{0,1\}^{\mathcal{L}\mathcal{T}\Omega}$ that specifies the production location choice for all countries and all periods, under all possible shock histories. Similarly, $\boldsymbol{r}$ is a point in $\{0,1\}^{\mathcal{L}\mathcal{T}\Omega}$ that specifies the innovation location choice. $\boldsymbol{o}_i = (\boldsymbol{y}_i, \boldsymbol{r}_i)$ is a point in $\{0,1\}^{2\mathcal{L}\mathcal{T}\Omega}$ that compactly represents a full decision rule. In this section, I consider a fixed initial state, occasionally omitting the notations $\boldsymbol{y}_{i,-1}, \boldsymbol{r}_{i,-1}, \omega_{i,-1}$. The firm's lifetime problem is to select an optimal $\boldsymbol{o}_i$ that maximizes $\Pi_0\left(\boldsymbol{o}_i\right)$.

In the following proposition, I establish that the firm's lifetime payoff function exhibits supermodularity with respect to the production and innovation location choices, thereby ensuring the effectiveness of the squeezing algorithm to be discussed in Section VI.A. For a detailed explanation of how supermodularity validates the squeezing algorithm, please refer to Theorem 1 in Alfaro-Urena et al. (2022).

Proposition 1. *Let $\mathcal{L}$ denote the set of locations, $\mathcal{T}$ the collection of time periods, and Ω the set of all possible paths of productivity shocks $\boldsymbol{z}$. Assume that sunk costs are larger than or equal to fixed costs, and β_1, β_2 and λ_1 are non-negative. If $(\eta - 1)\beta_m > \rho - 1$, then $\Pi_0 \left(\boldsymbol{o}_i | \boldsymbol{y}_{i,-1}, \boldsymbol{r}_{i,-1}, \omega_{i,-1}\right)$ is supermodular in $\boldsymbol{o}_i$ on $\{0,1\}^{2\mathcal{L}\mathcal{T}\Omega}$.*

The proof of Proposition 1 is provided in Appendix B.3. In simple words, supermodularity requires that if an item adds value to a decision set, it continues to add value in any subset of the original decision set. Intuitively, it corresponds to rich static and dynamic complementarities in the model, which I now elaborate on.

To start, under the condition that $(\eta - 1)\beta_m > \rho - 1$, y_{ilt} and $y_{il't}$ are complementary. This condition ensures that the static profit function is supermodular in y_{ilt} and $y_{il't}$. A larger value of $(\eta - 1)\beta_m$ means a greater response in revenue to a decrease in the marginal production cost (with elasticity η) and a stronger reaction of the marginal production cost to changes in intermediate prices (with elasticity β_m). Additionally, it requires the substitution effect between goods from different countries (with elasticity ρ) to be small, which is a condition similar to that of Antras et al. (2017).

Furthermore, y_{ilt} is complementary to y_{ilt+1} if sunk costs are larger than or equal to fixed costs. This is because offshoring production in the current period makes it cheaper to offshore production in the subsequent period. Similarly, r_{ilt} and r_{ilt+1} exhibit complementarity for the same reason. Thirdly, there is complementarity between y_{ilt} and r_{ilt} because local production increases the return to innovation (β_2) and decreases its cost (γ_1). Lastly, complementarity exists between y_{ilt} and $r_{il't}$ because having a production plant in neighboring countries within the region can also reduce the cost of innovation in the host country (γ_1). These complementarities together imply the supermodular nature of the model.

VI Solution Algorithm and Model Estimation

Building upon the supermodularity property of the model, I present a solution algorithm in Subsection VI.A for the dynamic combinatorial location choice problem. Then, I discuss the estimation steps and results in Subsection VI.B.

VI.A Solution Algorithm

Consistent with several other studies (Alfaro-Urena et al., 2022; Eaton et al., 2016; Caliendo et al., 2019; Kehoe et al., 2018; Igami, 2017, 2018; Igami and Uetake, 2020), I begin by assuming that the model is non-stationary until a terminal period T. Beyond this period, all exogenous determinants of payoffs, such as market demand and countries' production offshoring potentials, remain constant. Consequently, the value and policy functions become stationary for $t \geq T$. Additionally, let $\underline{t}_I$ represent the initial sample period.

Considering the computational challenges in solving the Bellman equation (7), we encounter a large state space coupled with the non-stationarity of the model. As a result, we need to solve $2^{2\mathcal{L}N_\omega\mathcal{T}}$ distinct choice problems, and for each of these problems, there are $2^{2\mathcal{L}}$ options to be evaluated.

The algorithm in this subsection addresses the challenge posed by the large state space by selectively computing the policy function $o_{it}(\cdot)_{t_I}^{t_F}$ at specific states rather than all possible states. In particular, it determines the values of the policy function at states $\{\check{\boldsymbol{y}}_{it}, \check{\boldsymbol{r}}_{it}, \check{\omega}_{it}\}_{t_I}^{t_F}$ that the firm would reach if it chooses the optimal location bundles at each period, and all exogenous determinants follow specific paths of interest $\{\check{\xi}_{it}, \check{\Phi}_{jt}\}_{t_I}^{t_F}$ (e.g., observed or simulated paths). This idea significantly limits the number of problems that need to be solved.

The second insight is that solving the optimization problem for firm i at period t and a given state does not require full knowledge of the firm's optimal choices in all states that may be subsequently reached (Alfaro-Urena et al., 2022). For instance, if a firm's potential return to R&D in the current state is sufficiently high, its optimal decision may be to innovate regardless of its optimal choices at any other states that may be reached in the future.

Therefore, the algorithm can only track the upper and lower bounds on the optimal choices, sparing the need to store the entire policy function. The last idea of the algorithm (similary to Jia, 2008; Arkolakis and Eckert, 2022) is to break down a hard problem with a large choice set into many simple problems by solving single-country problems each at a time while fixing choices in other countries at their bounds.

The algorithm consists of two steps. In the first step, upper and lower bounds on the firm's optimal choices along specific paths of interest are computed. If the upper and lower bounds coincide, they must coincide with the solution as well. However, if they do not, the algorithm then proceeds to the second step, where the bounds on the optimal choices are further tightened.

Step 1. Let's consider a specific firm i, and without loss of generality, assume that it is born in period 1.

1.1 To initiate the algorithm, I set an initial constant upper bound to be a vector of ones,

$$\bar{b}_i^{[0]} = \{\bar{y}_{ilt}, \bar{r}_{ilt}\}_{l,t} = \mathbf{1}^{2TL},$$

so that $\bar{b}_{ilt} \geq o_{ilt}(y_{it-1}, r_{it-1}, \omega_{it})$ for all $(y_{it-1}, r_{it-1}, \omega_{it})$. In other words, this is an upper bound of the firm's optimal production and innovation choices in each country l and period $t \geq 1$, regardless of the path of productivity shocks and comtemporary choices in other countries.

1.2 Next, I solve single-country problems each at a time to derive an upper bound policy function $\bar{o}_i^{[0]} \in \{0,1\}^{2TL\times 4N_\omega}$, where $\bar{o}_{ilt}^{[0]} \geq o_{ilt}(y_{it-1}, r_{it-1}, \omega_{it})$ for all $(y_{it-1}, r_{it-1}, \omega_{it})$. It is important to note that this upper bound policy function does not represent the actual policy function for the entire problem, which would be defined on the full state space $(y_{it-1}, r_{it-1}, \omega_{it}) \in \mathbb{R}^{2^L\times 2^L\times N_\omega}$, mapping to the full action space $(y_{it}, r_{it}) \in \mathbb{R}^{2^L\times 2^L}$, and requiring storage of size $\mathbb{R}^{T\times\left(2^L\times 2^L\times N_\omega\right)\times\left(2^L\times 2^L\right)}$. In contrast, the derived upper bound policy function corresponds to the problem for a specific country l, where the actions in other countries are fixed at the initial constant bounds. Therefore, the upper bound policy function bounds the firm's optimal choices regardless of contemporary choices in other countries but still dependent on the path of productivity shocks. This upper bound policy function is obtained through the following sub-steps.

1.2.1 Consider the final-period problem for a single country l.

$$\bar{V}_{ilT}\left(y_{ilT-1}, r_{ilT-1}, \omega_{iT}\right) = \max_{y_{ilT}\in\{0,1\}, r_{ilT}\in\{0,1\}} \Bigg\{ \pi_{iT}\left(\omega_{iT}, y_{iT}|\bar{b}^{[0]}_{i,-l,T}\right)$$

$$- \left[\left(1-y_{ilT-1}\right)\cdot y_{ilT}\cdot\phi^{p}_{s} + y_{ilT-1}\cdot y_{ilT}\cdot\phi^{p}_{f}\right]$$

$$- \left[\left(1-r_{ilT-1}\right)\cdot r_{ilT}\cdot\phi^{r}_{s} + r_{ilT-1}\cdot r_{ilT}\cdot\phi^{r}_{f,ilT}\left(y_{ilT}|\bar{b}^{[0]}_{i,-l,T}\right)\right]$$

$$+ \beta\mathbb{E}_{\xi}\bar{V}_{ilT}\left(s_{iT+1}|y_{ilT}, r_{ilT}, \omega_{iT}, \bar{b}^{[0]}_{i,-l,T}\right)\Bigg\}.$$

This single-country problem has a small state space and a small action space, making it easily solvable. Therefore, I obtain $\bar{V}_{ilT}(\cdot)\in\mathbb{R}^{4N_\omega}$ and $\bar{o}_{ilT}(\cdot)\in\mathbb{R}^{8N_\omega}$ using standard value function iterations.

1.2.2 Then, I use backward induction to solve a similar problem for periods $T-1$ to 1, obtaining $\bar{V}_{il}(\cdot)\in\mathbb{R}^{T\times 4N_\omega}$ and $\bar{o}^{[0]}_{il}(\cdot)\in\mathbb{R}^{2T\times 4N_\omega}$. The single-country problem for period t is depicted by

$$\bar{V}_{ilt}\left(y_{ilt-1}, r_{ilt-1}, \omega_{it}\right) = \max_{y_{ilt}\in\{0,1\}, r_{ilt}\in\{0,1\}} \Bigg\{ \pi_{it}\left(\omega_{it}, y_{it}|\bar{b}^{[0]}_{i,-l,t}\right)$$

$$- \left[\left(1-y_{ilt-1}\right)\cdot y_{ilt}\cdot\phi^{p}_{s} + y_{ilt-1}\cdot y_{ilt}\cdot\phi^{p}_{f}\right]$$

$$- \left[\left(1-r_{ilt-1}\right)\cdot r_{ilt}\cdot\phi^{r}_{s} + r_{ilt-1}\cdot r_{ilt}\cdot\phi^{r}_{f,ilt}\left(y_{ilt}\right)|\bar{b}^{[0]}_{i,-l,t}\right]$$

$$+ \beta\mathbb{E}_{\xi}\bar{V}_{ilt+1}\left(s_{it+1}|y_{ilt}, r_{ilt}, \omega_{it}, \bar{b}^{[0]}_{i,-l,t}\right)\Bigg\}.$$

1.2.3 I repeat Steps 1.2.1 and 1.2.2 for every country l and obtain the upper bound value functions $\bar{V}_{i}(\cdot)\in\mathbb{R}^{TL\times 4N_\omega}$ and the upper bound policy functions $\bar{o}^{[0]}_{i}(\cdot)\in\mathbb{R}^{2TL\times 4N_\omega}$.

1.2.4 At the end of Step 1.2, I save $\bar{o}^{[0]}_{i}(\cdot)\in\mathbb{R}^{2TL\times 4N_\omega}$ to be the updated upper bound policy function given the previous constant upper bound. The supermodularity property of the model guarantees $\bar{o}^{[0]}_{i}(\cdot)$ to be an upper bound on the firm's optimal choices.

1.3 To update the constant upper bound, I then evaluate the upper bound policy function, $\bar{o}^{[0]}_{i}$, at the most favorable path of productivity shocks. By doing this, I essentially obtain the highest choices among all possible histories. Specifically, I set

$$\bar{b}_{it'}^{[n]} = \bar{o}_{it'}^{[n-1]}\left(\bar{b}_{it'-1}^{[n]}, \omega_{it}\left(\omega_{i0}, \bar{\xi}\right)\right), t' = 1, \ldots, T$$

with the initial condition equal to $\mathbf{0}_{2J}$ and ω_{i0}.

1.4 I iterate over Steps 1.1 to 1.3 until the constant upper bounds converge, i.e. $\bar{b}_{it}^{[n]} = \bar{b}_{it}^{[n-1]}, \forall t \in [\underline{t}_i, T]$. The resulting upper bound policy function is denoted as $\bar{o}_i^*(\cdot) \in \{0,1\}^{2TL \times 4N_\omega}$. Moreover, using similar steps with the initial constant lower bound equal to $\underline{b}_i^{[0]} = \mathbf{0}^{2TL}$ and policy function bounds evaluated at the least favorable path of shocks, I obtain the converged lower-bound policy function $\underline{o}_i^*(\cdot)$.

1.5 Next, $\bar{o}_i^*(\cdot)$ and $\underline{o}_i^*(\cdot)$ are used to derive bounds on the firm's optimal choices along the path of interest, $\{\check{\xi}_{it}\}_t$.

$$\bar{\check{y}}_{it'} = \bar{o}_{it'}^*\left(\bar{\check{y}}_{it'-1}, \check{\xi}_{it'}\right), t' = \underline{t}_i, \ldots, t$$

$$\underline{\check{y}}_{it'} = \underline{o}_{it'}^*\left(\underline{\check{y}}_{it'-1}, \check{\xi}_{it'}\right), t' = \underline{t}_i, \ldots, t$$

with initial condition equal to $\mathbf{0}_{2J}$ and ω_{i0}.

1.6 If $\bar{\check{y}}_{it}$ and $\underline{\check{y}}_{it}$ coincide for all periods of interest (that is, from t_I to t_F), then they represent the firm's optimal choices along the path of interest. However, if there is at least one period where they differ, I proceed to step 2 to refine the bounds on the firm's optimal choices.

Step 2. In this step, I focus on period τ where the upper and lower bounds on the firm's optimal choices along the path of interest differ for the first time, i.e. $\tau = \min\{t \in [t_I, T] : \bar{\check{y}}_{it} > \underline{\check{y}}_{it}\}$. Using the knowledge about the firm's optimal choices along the path of interest up to period $\tau - 1$, I refine the bounds at period τ by solving for the same problem in Step 1 for periods $[\tau, T]$ with initial state $(\check{\boldsymbol{y}}_{i\tau-1}, \check{\boldsymbol{r}}_{i\tau-1}, \check{\omega}_{i\tau-1})$.

2.1 I initialize the constant upper bound in Step 2 by evaluating converged upper bound policy function from Step 1, $\bar{o}_{it}^*$, at the state firm i reaches if (1) it starts from $(\check{\boldsymbol{y}}_{i\tau-1}, \check{\boldsymbol{r}}_{i\tau-1}, \check{\omega}_{i\tau-1})$ at period τ, (2) for all $t' \in [\tau, T]$, the shock is at its highest level, and (3) the firm chooses the location bundle indicated by $\bar{o}^*_{it}$. Formally,

$$\bar{b}^{[0]}_{i\tau|\tau} = \bar{o}^*_{i\tau}\left(\check{y}_{i\tau-1}, \check{\omega}_{i\tau}\right)$$

$$\bar{b}^{[0]}_{it'|\tau} = \bar{o}^*_{it'}\left(\bar{b}^{[0]}_{it'-1|\tau}, \omega_{it}\left(\check{\omega}_{i\tau}, \bar{\xi}\right)\right), t' = \tau+1, \ldots, T,$$

where $\bar{b}^{[n]}_{i\cdot|\tau}$ denote the constant upper bound in step 2 for a given τ.

2.2 I repeat the procedure in Step 1 to solve the similar problem with a different initial state $(\check{\boldsymbol{y}}_{i\tau-1}, \check{\boldsymbol{r}}_{i\tau-1}, \check{\omega}_{i\tau-1})$ for periods $[\tau, T]$ until the constant upper bounds converge, i.e. $\bar{b}^{[n]}_{i\cdot|\tau} = \bar{b}^{[n-1]}_{i\cdot|\tau}$. Eventually, I obtain the resulting upper bound policy function denoted as $\bar{o}^*_{i\cdot|\tau}$.

2.3 I use a similar procedure as in Step 2.2 to obtain the converged lower bound policy function $\underline{o}^*_{i\cdot|\tau}$. Using the converged upper and lower bound policy functions, $\bar{o}^*_{i\cdot|\tau}$ and $\underline{o}^*_{i\cdot|\tau}$, I obtain bounds on the firm's optimal choices at period τ along the path of interest:

$$\bar{\check{y}}_{i\tau|\tau} = \bar{o}^*_{i\tau|\tau}\left(\check{y}_{i\tau-1}, \check{\omega}_{i\tau}\right),$$

$$\underline{\check{y}}_{i\tau|\tau} = \underline{o}^*_{i\tau|\tau}\left(\check{y}_{i\tau-1}, \check{\omega}_{i\tau}\right).$$

2.4 If the upper and lower bounds, $\bar{\check{y}}_{i\tau|\tau}$ and $\underline{\check{y}}_{i\tau|\tau}$, coincide at period τ, they provide the firm's optimal choice at that particular period along the path of interest. In this case, I proceed to the next $\tau' > \tau$ where the bounds from Step 1 differ and apply the Step 2 procedure again to tighten the bounds at τ'.

2.5 If the bounds at period τ from Step 2 do not coincide, I assume that the firm behaves according to the lower bound policy function. Fortunately, more than 99% of the individual problems are solved accurately, and therefore, this behavioral assumption has minimal impact on estimation results and counterfactuals.

VI.B Estimation Steps and Results

Having the solution algorithm at hand, I now estimate the model using a combination of micro firm-level data and country-level data. Table 8 provides a summary of the parameters to be estimated, along with the respective sources of identification for each parameter. Further details regarding the identification sources for each parameter will be discussed in the corresponding estimation steps below.

The estimation strategy consists of three steps. In the first step, I estimate the countries' production offshoring potentials θ_{lt}, firms' production offshoring capabilities Θ_{it}, and the elasticity of substitution ρ based on the firms' production shares across countries. Additionally, I estimate η in the first step using the average markup. In the second step, I use the control function approach to estimate the units cost function and the productivity evolution process, recovering the values for parameters $\beta_k, \beta_m, \alpha_0, \alpha_1, \mu, \beta_1,\ \beta_2$, and σ_ξ. In the third step, I use the method of simulated moments (MSM) to estimate the fixed and sunk cost parameters, i.e. $\phi_s^p, \phi_s^r, \phi_f^p, \phi_f^r$ and λ_1. Below, I will discuss each estimation step in detail.

VI.B.1 Step 1 – production offshoring potentials

In this step, I take the firm's production locations as given and focus on the variation in product value shares across countries. Given the CES aggregation structure of intermediates, firm i's value share of imported goods from country l in period t is given by

$$\chi_{ilt} = \left(\frac{w_{lt}\tau_{lt}T_{lt}}{p_{it}^m}\right)^{1-\rho} = \frac{\theta_{lt}}{\Theta_{it}},$$

which is simply the contribution of the country's production offshoring potential to the firm's production offshoring capability.

After taking logs of this equation and normalizing the foreign product shares by the firm's domestic product share (i.e. setting $\theta_{0t} = 1$ where 0 represents the US), I obtain the following expression:

$$\ln \chi_{ilt} - \ln \chi_{i0t} = \ln \theta_{lt} + \ln \epsilon_{ilt}. \tag{8}$$

A firm-country-year-level measurement error ϵ_{ilt} is added in this expression, as in Antras et al. (2017), to turn the model's equilibrium condition into an empirical specification. The dependent variable in equation (8) is the difference between a firm's share of goods imported from country l and its share of goods made domestically. Intuitively, a country's production offshoring potential is identified by how much a firm imports from the country relative to the domestic market. I estimate equation (8) via Ordinary Least Squares (OLS) and employ country-year fixed effects to capture the $\ln \theta_{lt}$ terms.

I cross-validate the estimates of production offshoring potentials, $\hat{\theta}_{lt}$, by comparing them to the number of firms importing from each country. Panel A of Figure 8 presents their overtime correlation for China, and Panel B plots the cross-sectional correlation for all countries in the year 2017. Both graphs show that a host country's production offshoring potential is highly consistent with the number of U.S. firms importing from that country. In 2017, China exhibited the highest production offshoring potential, followed by Canada and Germany. However, despite China's higher production offshoring potential, more firms imported from Canada than from China. This disparity, also observed in Antras et al. (2017), suggests that the fixed cost of production offshoring is likely lower for Canada, which provides the variation needed to separately identify the fixed cost from the production offshoring potentials.

Using the estimated production offshoring potentials, $\hat{\theta}_{lt}$, I calculate each firm's production offshoring capability as $\hat{\Theta}_{it} = \sum_{l \in \mathcal{L}} y_{ilt} \hat{\theta}_{lt}$. The estimates of $\hat{\Theta}_{it}$ imply that a firm importing from all countries in the sample has a production offshoring capability that is 10.3 percent larger than a firm obtaining intermediates only from the domestic market.

The effect of a firm's production offshoring capability on its marginal cost is determined by the parameter ρ. To estimate ρ, I run the following regression derived from the definition equation of production offshoring potentials:

$$\widehat{\ln \theta_{lt}} = -(\rho - 1) \cdot \ln (1 + T_{lt}) + \nu_{lt}.$$

Here, T_{lt} represents the ad valorem tariff rates, and ν_{lt} captures other determinants of offshore production costs, such as wage and shipping costs. I project the estimated production offshoring potentials on changes in tariffs instead of alternative cost shifters (e.g. wages and shipping costs) because it is more likely to be exogenous to firm characteristics.

The regression coefficients are reported in Table 9. Column (1) serves as the baseline without any controls, while column (2) includes controls for population, common language, and colonial relationships. In column (3), I further control for human capital and the level of corruption. The coefficient of tariff remains consistent across different sets of controls. Taking column (1) as the baseline, I estimate ρ to be 3.739.

Using the estimates for θ_{lt} and ρ, I construct the intermediate price index for each firm as

$$p_{it}^{m} = \left(\sum_{l} y_{ilt} \cdot \hat{\theta}_{lt} \right)^{\frac{1}{1-\hat{\rho}}}.$$

The intermediate price that a firm faces when it imports from all countries is approximately 3.52 percent lower compared to a firm that only obtain intermediates domestically.

Finally, I recover the demand elasticity parameter η from markups. With CES preferences and monopolistic competition, the ratio of sales to total variable cost is $\eta/(\eta-1)$, implying the following relationship:

$$\eta = \frac{R_{it}/\mathrm{tvc}_{it}}{R_{it}/\mathrm{tvc}_{it} - 1},$$

where R_{it} and tvc_{it} are the firm's revenue and total variable cost in year t, respectively. The median markup in the sample is 1.237, suggesting an estimate of $\hat{\eta} = 5.217$. For further information on the construction of total variable cost using census data variables, please refer to Appendix A.1.

VI.B.2 Step 2 – Cost Function and Productivity Evolution

At the beginning of this step, I augment the revenue function with an independently and identically distributed (i.i.d.) error term:

$$\begin{aligned} \ln R_{it} = {} & (1-\eta) \ln \left(\frac{\eta}{\eta - 1} \right) + \ln \Phi_{jt} \\ & + (1-\eta) \left(\beta_0 + \beta_k \ln k_{it} + \beta_w \ln w_{jt} + \beta_m \ln p_{it}^{m} - \omega_{it} \right) + u_{it}. \end{aligned}$$

Note that the composite error term $u_{it} - (1-\eta)\omega_{it}$ correlates with firm's input choices due to its inclusion of the firm's productivity. Consequently, a simple OLS regression of this equation would yield biased estimates for the coefficients of input factors.

Building on insights from Olley and Pakes (1996) and Ackerberg et al. (2015), I express productivity, conditional on k_{it} and p_{it}^{m}, as a function of the variable energy input; i.e. $\omega_{it}(k_{it}, p_{it}^{m}, n_{it})$. After combining the constant and industry-year-specific terms, I obtain the following equation:

$$\ln R_{it} = \psi_0 + \sum_j \sum_t \psi_{jt} + \underbrace{h\left(k_{it}, p_{it}^m, n_{it}\right)}_{\phi_{it}} + \nu_{it}, \tag{9}$$

from which I can estimate $\hat{\phi}_{it}$ using a second-order polynomial function $h(\cdot)$. By plugging the estimated equation (9) into the productivity evolution process, namely equation (6), I obtain a nonlinear equation that can be estimated using regressions:

$$\begin{aligned} \hat{\phi}_{it} =& \beta_k^* \cdot \ln k_{it} + \beta_m^* \cdot \ln p_{it}^m - \alpha_0^* + \alpha_1 \cdot \left(\hat{\phi}_{it-1} - \beta_k^* \cdot \ln k_{it-1} - \beta_m^* \cdot \ln p_{it-1}^m\right) \\ & - \sum_l \left[1 + X_{lt-1}\rho\right] \cdot \left[\beta_1^* r_{ilt-1} + \beta_2^* r_{ilt-1} y_{ilt-1}\right] - \xi_{it}^*, \end{aligned} \tag{10}$$

where the transformation $x^* = (1-\eta)x$. I estimate this equation using Nonlinear Least Squares. Once the coefficients are estimated, a firm's productivity can be computed as follows:

$$\omega_{it} = -\frac{\hat{\phi}_{it}}{1-\hat{\eta}} + \hat{\beta}_k \cdot \ln k_{it} + \hat{\beta}_m \cdot \ln p_{it}^m.$$

Table 10 presents the coefficients estimated through the NLS regressions. Columns (1) to (3) include all countries in the sample, while columns (4) to (6) exclude countries often considered as tax havens for robustness. The known tax havens in the sample are Hong Kong, Ireland, Luxembourg, the Netherlands, Switzerland, and Singapore according to Gravelle (2015).[20] Column (3) is the preferred specification for subsequent model estimation steps. The estimated coefficient for log capital, $\hat{\beta}_k$, is -0.165, implying that if a firm's capital stock doubles, its marginal cost of production will decrease by 16.5%. Additionally, the positive coefficient of intermediate price confirms that the unit production cost increases when intermediates become more expensive.

The most notable observation from Table 10 is that the estimate of β_1 is not significantly different from zero, while the estimate of β_2 is significant and positive. This indicates that having only innovation in a foreign country without local production does not significantly contribute to a firm's future productivity. However, when innovation is combined with local

[20] This list of tax havens was prepared by the U.S. Congressional Research Service and is similar to lists prepared by the Organization for Economic Cooperation and Development (OECD) and the U.S. Government Accountability Office (GAO).

production, it has a significant positive effect on productivity. This finding emphasizes the importance of the synergy effect, suggesting that the proximity between production and innovation activities is crucial for enhancing innovation efficiency.

Table 10 also reports the mean and standard deviation of the implied elasticity of R&D across countries, represented by $(1 - X_{lt}\hat{\mu}) \cdot (\hat{\beta}_1 + \hat{\beta}_2)$. The estimated elasticity indicates that, on average, conducting R&D in a foreign country leads to a 0.13% to 0.17% increase in firm productivity in the following year. The term $X\mu$ in the productivity evolution process essentially captures heterogeneous synergy effect between production and innovation across countries since β_1 is estimated to be close to zero. Figure 9 displays the estimates of $1 + X\hat{\mu}$ for all countries in the sample, highlighting that having production and R&D in countries like Canada and India yields larger productivity gains. On the other hand, countries like Australia and Spain exhibit smaller synergy between production and R&D.[21]

VI.B.3 Step 3 – Dynamic Parameters

Taking stock, it is important to note that the estimates obtained in steps 1 and 2 confirm that $(\eta - 1)\beta_m > \rho - 1$, which indicates that the dynamic model satisfies the supermodularity property as stated in Proposition 1. Consequently, it can be solved using the algorithm outlined in Section VI.A. However, the dynamic model does not yield a closed form solution to firms' location choices conditioning on market observables and a given vector of parameter values; therefore, I use simulation methods in this step to estimate the cost parameters. Popular simulation methods include the method of simulated log likelihood (MSL) and the method of simulated moments (MSM). Implementing MSL is difficult because the cross-

[21] The estimates in Table 10 reveal that U.S. firms experience stronger synergy between production and innovation in countries with lower income levels. That is, the increase in the return to R&D resulting from local production is higher in poorer countries. There are multiple explanations for this result. Firstly, the initial obstacles of innovating without producing may be higher in poorer countries for U.S. firms. For instance, new product testing that requires communication with the headquarter teams is much more difficult in a stand-alone lab in South Africa than that in Canada. Secondly, immersion to a more exotic culture often sparks new idea and requires more product customization. Finally, it can also be due to factors ignored in the model—such as the intensive margin of offshoring and industry heterogeneity. For example, the higher fixed costs associated with offshoring production to poorer countries may imply that, although firms are more likely to offshore production to richer countries, conditional on already producing in both locations, they produce relatively more in poorer countries to compensate the higher fixed costs. The greater amount of offshore production in poorer countries thus leads to a larger boost in the return to R&D stemming from local production. This paper do not provide further evidence for which explanations are more relevant.

country and cross-period dependence in the location choices imply that the log-likelihood of the sample is no longer the sum of the log-likelihood of each country and period, and one needs an exceptionally large number of simulations to get a reasonable estimate of the sample's likelihood. Hence, I employ the MSM method to estimate five dynamic parameters, $\phi_s^p, \phi_s^r, \phi_f^r, \phi_f^p$ and λ_1.

As demonstrated in Table 11, six moments are used to identify five cost parameters. The first two moments, $\mathbb{E}[y_{ilt}]$ and $\mathbb{E}[r_{ilt}]$, reflect the fraction of firms that offshore production and R&D to foreign countries, respectively, and help identify the fixed costs ϕ_f^r and ϕ_f^p. The third and fourth moments, $\mathbb{E}[y_{ilt}\,(1 - y_{ilt-1})]$ and $\mathbb{E}[r_{ilt}\,(1 - r_{ilt-1})]$, capture the frequency at which non-offshoring firms start to offshore production and innovation, providing information about the sunk costs ϕ_s^p and ϕ_s^r. The last two moments, $\mathbb{E}[y_{ilt}y_{il't}|c_{ll'} = 1] - \mathbb{E}[y_{ilt}y_{il't}|c_{ll'} = 0]$ and $\mathbb{E}[r_{ilt}r_{il't}|c_{ll'} = 1] - \mathbb{E}[r_{ilt}r_{il't}|c_{ll'} = 0]$, measure the disparity in the frequency of firms offshoring production and innovation to both country l and l' based on whether they are in the same region or not. A higher value of the parameter λ_1 results in a larger difference in this frequency.

The estimated values of the cost parameters are reported in Panel A of Table 11. The sunk and fixed costs associated with offshoring production to a foreign country are estimated to be around $1.1 million. The sunk and fixed costs associated with offshoring R&D activities to a foreign country are estimated to be around $44 million. The latter estimates are comparable to the conditional mean R&D expenditure of the large multinational firms in my sample, as showed in Table 1. The cost-sharing parameter λ_1 is estimated to be $0.22 million, which accounts for only 0.5% of the fixed cost of R&D. This implies that firms primarily colocate production and innovation to improve innovation efficiency rather than to share overhead costs.

VII Counterfactual Exercises

I proceed by conducting four counterfactual exercises. The first exercise quantifies the relative importance of colocation mechanisms in my model, implemented by individually shutting down the synergy effect and the cost-sharing mechanism. Next, I demonstrate the alignment between model-predicted effects of the Trump Tariffs on China and their reduced-form counterparts. Then, I simulate counterfactual shocks to U.S. firms' production offshoring in China, analyzing how they impact the global geography of production and innovation. I find nontrivial third-country effects as well as nonlinear effects on innovations shares that are contingent upon firm heterogeneity and the size of shocks. Lastly, I highlight my model's prediction of dynamic losses from tariff increases, distinct from static models of global production and sourcing (e.g. Antras et al., 2017).

VII.A Quantifying Importance of Colocation Mechanisms

I perform two analyses in this subsection to shed light on the relative importance of two key model mechanisms. In the first analysis, I gradually weaken the synergy effect between production and innovation by reducing the value of parameter β_2 from its baseline estimate to zero. Simulation outcomes are reported in Panel A of Table 12. Under the baseline $\hat{\beta}_2$, the probability of a firm offshoring production and innovation to a foreign country are 16.05 and 1.28 percentage points, respectively, as shown in the first column. When β_2 is reduced by a half in the third column, the production offshoring probability decreases by less than 1%, while the R&D offshoring probability drops by more than 90%. Additionally, the probability of a firm offshoring R&D to a foreign country, conditional on it having offshore production in that country, is reduced by more than 90%. This result indicates that the synergy effect is a crucial reason for why firms want to conduct R&D in general and why they want to do so particularly in countries where they have production sites.

The second analysis involves reducing the value of the cost sharing parameter λ_1 from its baseline estimate to zero. By doing so, I completely shut down the mechanism whereby local production can reduce the fixed cost for firms to conduct R&D in a foreign country. The results for this analysis are reported in Panel B of Table 12. As a consequence of the higher effective costs for offshoring, the first two rows of the table show that the probability of a firm offshoring production and innovation to a foreign country decreases by 0.16% and 7.55%, respectively. The within-country colocation pattern, as shown in the third row, gets weakened: the probability of a firm performing R&D in a foreign country, conditional on it having offshore production in that country, decreases by 7.41%. Furthermore, the cross-country colocation within the same region is also negatively affected. In the fourth row, given that a firm has production in country l, the probability that it conducts R&D in other countries within the same region as l reduces from 7.12 to 6.59 percentage points.

The fact that these colocation measures are only reduced by less than 8% when the cost sharing mechanism is shut down suggests that over 90% of the observed colocation pattern is attributable to the synergy effect (along with the implicit scale effects). Combined with the large effects observed in the first analysis, this subsection demonstrates that the synergy effect is the primary factor behind firms' incentives to colocate production and innovation.

VII.B Policy Implications

In recent years, the trade relationship between the world's two largest economies, the U.S. and China, has become increasingly contentious. China has been a major target of the U.S. in the trade war since 2018 and recent trade policies of both the Trump and Biden administrations. The increased tariffs and U.S. government's reshoring efforts together made it more costly for U.S. firms to offshore production to China. My framework is particularly suitable for studying such policies for three reasons. Firstly, given the incentives of multinational firms to colocate production and innovation, policies that shift their production locations will inherently impact their R&D locations. Secondly, third-country effects emerge due to cross-country interdependence. For instance, a U.S. multinational firm may want to have only one R&D lab in East Asia. As the cost of offshoring production to China gets higher, this firm might stop innovating in China and move its lab to South Korea. Such outcomes would be missed by models that do not allow for interdependent choices across countries. Thirdly, trade policies that increase tariffs generate not only static losses stemming from higher intermediate prices and lower profits, but also creating dynamic losses from diminished offshore R&D investments and thus slower productivity growth. The dynamic aspect will be missed in static models of global production and sourcing that lack an endogenous R&D process.

VII.B.1 Model Validation Based On Trump Tariffs

Between 2017 and 2019, the U.S. tariff rate on Chinese goods increased by 3.8 percentage point, scaling from 4.07 to 7.87 (calculated from the Trade Analysis Information System data). The estimates in Table 5 indicate that this tariff change would lead to a 7.2% drop in imports and a 0.11 percentage point reduction in R&D likelihood.

In my model, implementing the identical tariff increase effectively equates to a 9.4% decrement in China's production offshoring potential. Following such a policy shock, the model predicts a 6.6% reduction in imports and a 0.09 percentage point decline R&D likelihood in China. This exercise highlights the model's capacity to generate effects of the right magnitude, aligning with the reduced-form estimates.

VII.B.2 Third-Country Effects and Nonlinear Effects in Counterfactual Policies

In this exercise, I simulate two sets of policy shocks adversely affecting U.S. firms' production offshoring to China after 2017. The first set of shocks consist of tariff increases equivalent to a reduction of China's production offshoring potential by 25% to 100%. The second set of shocks pertains to increased fixed and sunk costs for production offshoring in China, escalating by 2 to 100 thousand dollars. I examine how the global distributions of production and innovation respond to shocks of different magnitudes.

The first observation underscores the importance of third-country effects of such bilateral trade policies. When the U.S implements tariff increases on China to the extent that China's production offshoring potential plummets by 25%, the likelihood of a firm offshoring production to China over the subsequent two years declines by 9.4 percentage points, whereas the corresponding decrease for other regions of the world (ROW) is 0.8 percentage points. Additionally, the probability of a firm offshoring R&D to China is reduced by 0.2 percentage point, compared to a 0.61 percentage point decrease for ROW. These nontrivial ROW effects consistently manifest across all counterfactual policy shocks. They emerge from the cross-country interdependence built in my framework, a feature typically missing in previous models that assume independent decisions in each host country.

Next, the changes in the relative shares of production and innovation among China, the U.S., and ROW (as illustrated in Figure 10) unveil more interesting patterns. The production shares shift from China to the U.S. and ROW under all policy shocks. However, the direction of changes in innovation shares hinges on the magnitude of the shock. For moderate shocks (e.g. production costs increasing by less than 30 thousand dollars), innovation shares rise for

China and the U.S., while decreasing for ROW. In contrast, under large shocks, innovation shares drop in China but increase in the U.S. and ROW.

Firm heterogeneity plays a big role in driving the nonlinear effects in innovation shares. Figure 11 depicts the fraction of firms that offshore production and innovation to China and ROW, categorized by the deciles of the firm's productivity and capital stock. A comparison between the first and second columns of plots reveals that many firms with relatively low productivity and capital stock tend to produce in China without conducting innovation there. Instead, they carry out R&D activities in other countries. This is because China is estimated to have the highest production offshoring potential (see Figure 8) but relatively low synergy effect between production and innovation (see Figure 9).

When a moderate shock hits, these firms are among the first to be affected, and due to the scale effect at the firm level, they reduce offshoring activities worldwide. In particular, they reduce production in China and innovation in ROW, leading to a relative increase in China's innovation share. However, as the shock gets larger, even firms in the top productivity and capital stock decile (i.e., the upper right block) that also innovate in China start to be impacted. This is when innovation shares shift from China to the U.S. and ROW.

In summary, I find sizable third-country effects of counterfactual policy shocks that adversely affect U.S. firms' production offshoring to China, as well as nonlinear effects in innovation shares that are contingent upon firm heterogeneity and the magnitude of the shock.

VII.B.3 Dynamic Effects of Trade Policies

One important difference between my framework and previous static models of global production and sourcing is that it builds in endogenous productivity that is affected by firms' R&D investments. As a result, my framework can evaluate not just the static losses from adverse trade shocks that are standard in traditional models, but also the dynamic losses that emerge as offshoring decisions and endogenous productivity influence each other. The last counterfactual exercise, which simulates a permanent 50% decrease in China's production offshoring potential, demonstrates my framework's ability to capture such dynamic effects.

The simulation results show a 1.77 percentage point decline in the probability of offshoring production and a 0.546 percentage point reduction in the probability of offshoring innovation to a host country, immediately following the 50% reduction in China's production offshoring potential. Panel A of Figure 12 presents additional outcome measures regarding static losses: both the distribution of log intermediate price and log marginal production cost shift rightward, indicating cost increases. Correspondingly, the distribution of log profits shifts leftward. These changes capture the static losses from deteriorated production offshoring opportunities, aligning with the findings of earlier static global production and sourcing models.

Furthermore, this adverse trade shock also generates dynamic losses depicted in Panel B of Figure 12. With the decline in China's production offshoring potential, firms opt to reduce their offshore production and innovation. This reduction subsequently diminishes their future productivity, as specified in Equation (6). As firms experience lowered productivity levels, they are less able to overcome the sunk and fixed costs associated with offshoring, further reducing their likelihood to conduct production and innovation in foreign countries. During this process, firms' intermediate prices and marginal production costs rise due to less offshore production. These negative effects accrue dynamically: the initial average productivity loss originates from zero and gradually accumulates to around 0.4% over the span of a decade, while average firm profit experiences a 7% reduction in the first year, intensifying to 8.5% after the ten-year period.

VIII Conclusion

In this paper, I study the location choices of multinational firms regarding the offshoring of production and innovation. I show empirically the importance of colocation benefits between production and innovation, as well as their cross-country interdependence in shaping these decisions. Causal evidence reveals that an increase in a host country's tariff not only diminishes production and innovation within that country but also affects other countries within the region.

I contribute to the literature on multinational production and innovation by developing a new dynamic framework that allows for direct interaction between production and innovation and considers interdependence across countries. My model incorporates rich static and dynamic complementarities between offshoring decisions. I also establish conditions for the model's supermodularity and employ a new algorithm to effectively solve this otherwise NP hard problem. The quantification exercises demonstrate that the synergy effect between production and innovation is the main incentive for firms to colocate these two activities.

I apply the model to examine the impact of U.S. trade policies that adversely affect production offshoring to China. I find significant third-country effects, as well as nonlinear effects in innovations shares that are contingent upon firm heterogeneity and the magnitude of the policy shocks. Moreover, I demonstrate the importance of dynamic effects that were absent in earlier static frameworks of global production and sourcing. Hence, there is a need for policymakers to consider the interactions between production and innovation, and thus the complex but potentially unintended consequences of manufacturing offshoring and reshoring policies on the geography of innovation.

References

Ackerberg, Daniel A, Kevin Caves, and Garth Frazer (2015) "Identification properties of recent production function estimators," *Econometrica*, 83 (6), 2411–2451.

Aguirregabiria, Victor and Arvind Magesan (2013) "Euler equations for the estimation of dynamic discrete choice structural models," in *Structural Econometric Models*, 31, 3–44: Emerald Group Publishing Limited.

——— (2016) "Solution and estimation of dynamic discrete choice structural models using Euler equations," Available at SSRN 2860973.

Aguirregabiria, Victor and Gustavo Vicentini (2016) "Dynamic Spatial Competition Between Multi-Store Retailers," *The Journal of Industrial Economics*, 64 (4), 710–754.

Alfaro-Urena, Alonso, Juanma Castro-Vincenzi, Sebastián Fanelli, and Eduardo Morales (2022) "Firm Export Dynamics in Interdependent Markets,"Technical report, Mimeo.

Antràs, Pol, Evgenii Fadeev, Teresa C Fort, and Felix Tintelnot (2023) "Exporting, Global Sourcing, and Multinational Activity: Theory and Evidence from the United States,"Technical report, National Bureau of Economic Research.

Antras, Pol, Teresa C Fort, and Felix Tintelnot (2017) "The margins of global sourcing: Theory and evidence from us firms," *American Economic Review*, 107 (9), 2514–2564.

Arkolakis, Costas and Fabian Eckert (2022) "Combinatorial discrete choice," Available at SSRN 3455353.

Arkolakis, Costas, Natalia Ramondo, Andrés Rodríguez-Clare, and Stephen Yeaple (2018) "Innovation and production in the global economy," *American Economic Review*, 108 (8), 2128–73.

ASEAN (2017) "ASEAN Investment Report 2017: Foreign Direct Investment and Economic Zones in ASEAN."

Aw, Bee Yan, Mark J Roberts, and Daniel Yi Xu (2011) "R&D investment, exporting, and productivity dynamics," *American Economic Review*, 101 (4), 1312–1344.

Bahar, Dany (2020) "The hardships of long distance relationships: time zone proximity and the location of MNC's knowledge-intensive activities," *Journal of International Economics*, 125, 103311.

Barro, Robert J and Jong Wha Lee (2013) "A new data set of educational attainment in the world, 1950–2010," *Journal of development economics*, 104, 184–198.

Bernard, Andrew B and Teresa C Fort (2015) "Factoryless goods producing firms," *American Economic Review*, 105 (5), 518–523.

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

Berry, Steven, James Levinsohn, and Ariel Pakes (1995) "Automobile Prices in Market Equilibrium," *Econometrica*, 63 (4), 841–890.

Bilir, L Kamran and Eduardo Morales (2020) "Innovation in the global firm," *Journal of Political Economy*, 128 (4), 1566–1625.

Bils, Mark and Peter J Klenow (2000) "Does schooling cause growth?" *American economic review*, 90 (5), 1160–1183.

Boehm, Christoph E, Aaron Flaaen, and Nitya Pandalai-Nayar (2020) "Multinationals, offshoring, and the decline of US manufacturing," *Journal of International Economics*, 127, 103391.

Bøler, Esther Ann, Andreas Moxnes, and Karen Helene Ulltveit-Moe (2015) "R&D, international sourcing, and the joint impact on firm performance," *American Economic Review*, 105 (12), 3704–3739.

Bown, Chad P (2020) "US-China trade war tariffs: An up-to-date chart," *Peterson Institute for International Economics*, 14.

Branstetter, Lee G, Jong-Rong Chen, Britta Glennon, and Nikolas Zolas (2021) "Does Offshoring Production Reduce Innovation: Firm-Level Evidence from Taiwan,"Technical report, National Bureau of Economic Research.

Caliendo, Lorenzo, Maximiliano Dvorkin, and Fernando Parro (2019) "Trade and labor market dynamics: General equilibrium analysis of the china trade shock," *Econometrica*, 87 (3), 741–835.

Criscuolo, Chiara, Jonathan E Haskel, and Matthew J Slaughter (2010) "Global engagement and the innovation activities of firms," *International journal of industrial organization*, 28 (2), 191–202.

Cunningham, Cindy, Lucia Foster, Cheryl Grim, John Haltiwanger, Sabrina Wulff Pabilonia, Jay Stewart, and Zoltan Wolf (2021) "Dispersion in dispersion: Measuring establishment-level differences in productivity," IZA Discussion Paper.

Delgado, Mercedes (2020) "The co-location of innovation and production in clusters," *Industry and Innovation*, 27 (8), 842–870.

Doraszelski, Ulrich and Jordi Jaumandreu (2013) "R&D and productivity: Estimating endogenous productivity," *Review of economic studies*, 80 (4), 1338–1383.

Eaton, Jonathan and Samuel Kortum (2002) "Technology, geography, and trade," *Econometrica*, 70 (5), 1741–1779.

Eaton, Jonathan, Samuel Kortum, Brent Neiman, and John Romalis (2016) "Trade and the global recession," *American Economic Review*, 106 (11), 3401–3438.

Fajgelbaum, Pablo D, Pinelopi K Goldberg, Patrick J Kennedy, and Amit K Khandelwal (2020) "The return to protectionism," *The Quarterly Journal of Economics*, 135 (1), 1–55.

Fajgelbaum, Pablo David, Pinelopi Koujianou Goldberg, Patrick Kennedy, Amit Kumar Khandelwal, and Daria Taglioni (2022) "The US-China Trade War and Global Reallocations,"Technical report, The World Bank.

Fan, Jingting (2019) "Talent, geography, and offshore R&D," Working Paper.

Fan, Jingting, Eunhee Lee, and Valerie Smeets (2022) "High-Skill Immigration, Offshore R&D, and Firm Dynamics."

Fort, Teresa C, Wolfgang Keller, Peter K Schott, Stephen Yeaple, and Nikolas Zolas (2020) "Colocation of Production and Innovation: Evidence from the United States," Working Papaer.

Foster, Lucia, Cheryl Grim, and Nikolas Zolas (2020) "A portrait of US firms that invest in R&D," *Economics of Innovation and New Technology*, 29 (1), 89–111.

Gravelle, Jane G (2015) "Tax havens: International tax avoidance and tax evasion," *Congressional Research Service Report* (7-5700), 605–627.

Holmes, Thomas J (2011) "The diffusion of Wal-Mart and economies of density," *Econometrica*, 79 (1), 253–302.

Hsiao, Allan (2021) "Coordination and commitment in international climate action: evidence from palm oil," Working Paper.

Igami, Mitsuru (2017) “Estimating the innovator’s dilemma: Structural analysis of creative destruction in the hard disk drive industry, 1981–1998,” *Journal of Political Economy*, 125 (3), 798–847.

——— (2018) “Industry dynamics of offshoring: The case of hard disk drives,” *American Economic Journal: Microeconomics*, 10 (1), 67–101.

Igami, Mitsuru and Kosuke Uetake (2020) “Mergers, innovation, and entry-exit dynamics: Consolidation of the hard disk drive industry, 1996–2016,” *The Review of Economic Studies*, 87 (6), 2672–2702.

Imbs, Jean and Isabelle Mejean (2015) “Elasticity optimism,” *American economic journal: macroeconomics*, 7 (3), 43–83.

Jia, Panle (2008) “What happens when Wal-Mart comes to town: An empirical analysis of the discount retailing industry,” *Econometrica*, 76 (6), 1263–1316.

Kamal, Fariha and Wei Ouyang (2020) “Identifying US Merchandise Traders: Integrating Customs Transactions with Business Administrative Data,”Technical report, US Census Bureau, Center for Economic Studies.

Kehoe, Timothy J, Kim J Ruhl, and Joseph B Steinberg (2018) “Global imbalances and structural change in the United States,” *Journal of Political Economy*, 126 (2), 761–796.

Kuemmerle, Walter (1997) “Building effective R&D capabilities abroad,” *Harvard business review*, 75, 61–72.

Lakatos, Csilla and Franziska Ohnsorge (2017) “Arm’s-length trade: a source of post-crisis trade weakness,” World Bank Policy Research Working Paper 8144.

Lan, Ting (2019) “The Coagglomeration of Innovation and Production,”Technical report, mimeo, University of Michigan.

Manski, Charles F (1993) “Identification of endogenous social effects: The reflection problem,” *The review of economic studies*, 60 (3), 531–542.

Marquez, Jaime (2002) *Estimating trade elasticities*, 39: Springer Science & Business Media.

Morales, Eduardo, Gloria Sheu, and Andrés Zahler (2019) “Extended gravity,” *The Review of economic studies*, 86 (6), 2668–2712.

National Science Board (2014) "Science and Engineering Indicators 2014," Arlington, VA: National Science Foundation.

Olley, G Steven and Ariel Pakes (1996) "The Dynamics of Productivity in the Telecommunications Equipment," *Econometrica*, 64 (6), 1263–1297.

Piveteau, Paul (2021) "An empirical dynamic model of trade with consumer accumulation," *American Economic Journal: Microeconomics*, 13 (4), 23–63.

Roberts, Mark J, Daniel Yi Xu, Xiaoyan Fan, and Shengxing Zhang (2018) "The role of firm factors in demand, cost, and export market selection for Chinese footwear producers," *The Review of Economic Studies*, 85 (4), 2429–2461.

Rodríguez-Clare, Andrés (2010) "Offshoring in a ricardian world," *American Economic Journal: Macroeconomics*, 2 (2), 227–258.

Santacreu, Ana Maria (2021) "Dynamic gains from trade agreements with intellectual property provisions," FRB St. Louis Working Paper (2021-10).

Stock, James H, Jonathan H Wright, and Motohiro Yogo (2002) "A survey of weak instruments and weak identification in generalized method of moments," *Journal of Business & Economic Statistics*, 20 (4), 518–529.

Sweeting, Andrew (2013) "Dynamic product positioning in differentiated product markets: The effect of fees for musical performance rights on the commercial radio industry," *Econometrica*, 81 (5), 1763–1803.

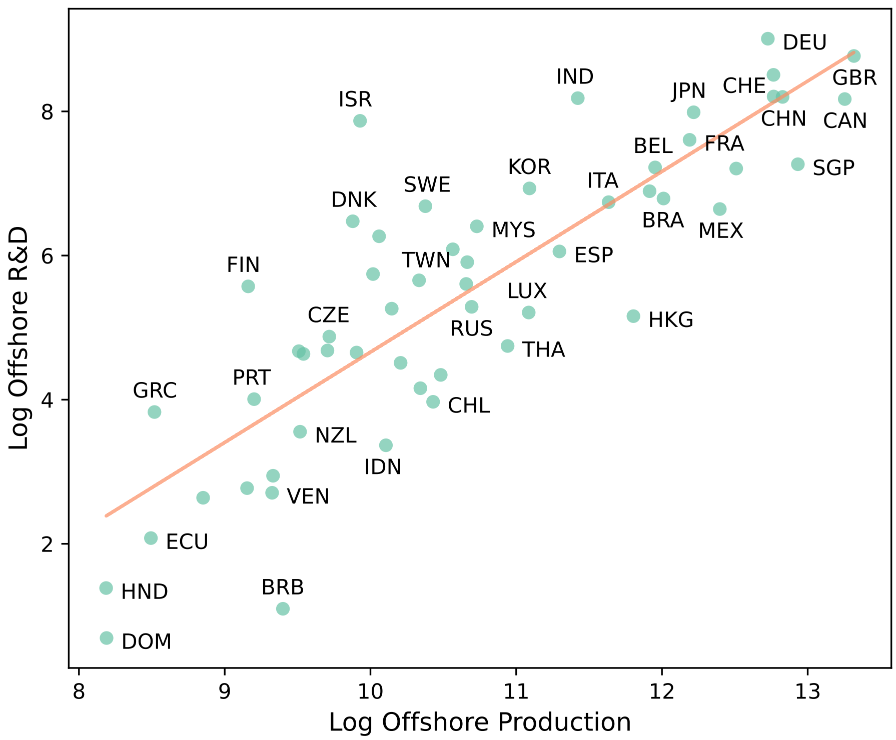
Tecu, Isabel (2013) "The location of industrial innovation: does manufacturing matter?," US Census Bureau Center for Economic Studies Paper No. CES-WP-13-09.

UNCTAD, United Nations Conference on Trade and Development (2015) "World Investment Report 2005: Transnational Corporations and the Internationalization of R&D," New York: United Nations.

U.S. Treasury (2021) "The made in America tax plan."

White House (2022) "FACT SHEET: CHIPS and Science Act Will Lower Costs, Create Jobs, Strengthen Supply Chains, and Counter China."

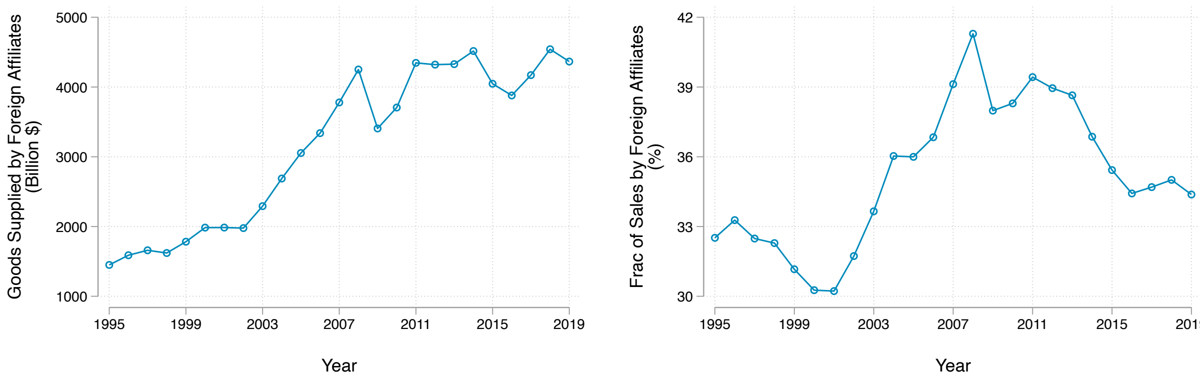
Figure 1: U.S. Offshore Production and R&D by Host Country, 2017



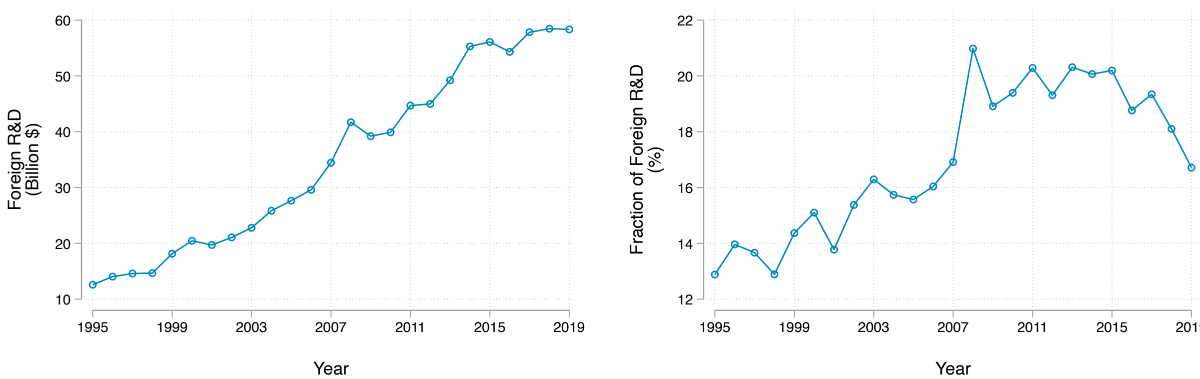
Notes: Figure plots the offshore production and R&D expenditure of U.S. multinational firms across destination countries in the year 2017. Data is obtained from the Bureau of Economic Analysis's Survey of U.S. Direct Investment Abroad (USDIA), a publicly available survey that collects information about activities of U.S. multinational parent firms and their foreign affiliates. Offshore Production is measured by the dollar value of goods supplied by foreign affiliates, and offshore R&D refers to the research and development activities performed by foreign affiliates.

Figure 2: Trend of Offshore Production and R&D, 1995-2020

(a) Production



(b) R&D



Notes: Figure plots time trends of offshored production and R&D activities between 1995 and 2019. Data is obtained from the Bureau of Economic Analysis's Survey of U.S. Direct Investment Abroad (USDIA). The survey covers activities of US multinational parent firms and their foreign affiliates. Panel (a) plots the dollar value of goods supplied by foreign affiliates of US parent firms and the fraction of sales made by foreign affiliates among the firm's total sales. Panel (b) plots the dollar value of foreign R&D expenditure of US firms and the fraction of foreign R&D expenditure in the firm's total R&D expenditure worldwide.

Figure 3: BRDIS Survey Questionnaire

2-11 Of the amount reported in Question 2-10, column 2, how much R&D was performed in the following locations? For full list of countries in each region see Question by Question Guidance at https://www.census.gov/programs-surveys/brds/information/brdshelp.html#q2-11.

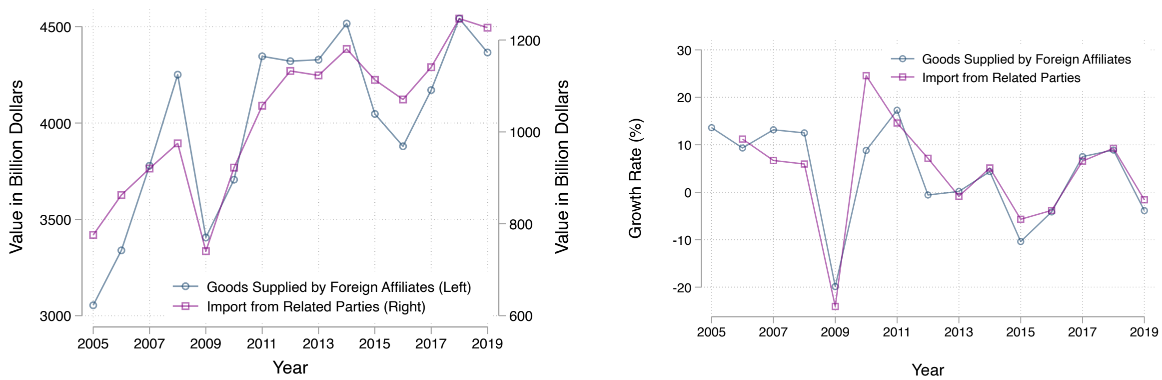
	$Bil.	Mil.	Thou.		$Bil.	Mil.	Thou.
Canada				Germany			
Puerto Rico				Hungary.			
Europe	$Bil.	Mil.	Thou.	Ireland.			
Austria.				Italy			
Belgium.				Luxembourg . . .			
Czech Rep				Netherlands. . . .			
Denmark				Norway			
Finland				Poland.			
France				Russia			

Question continues on next page

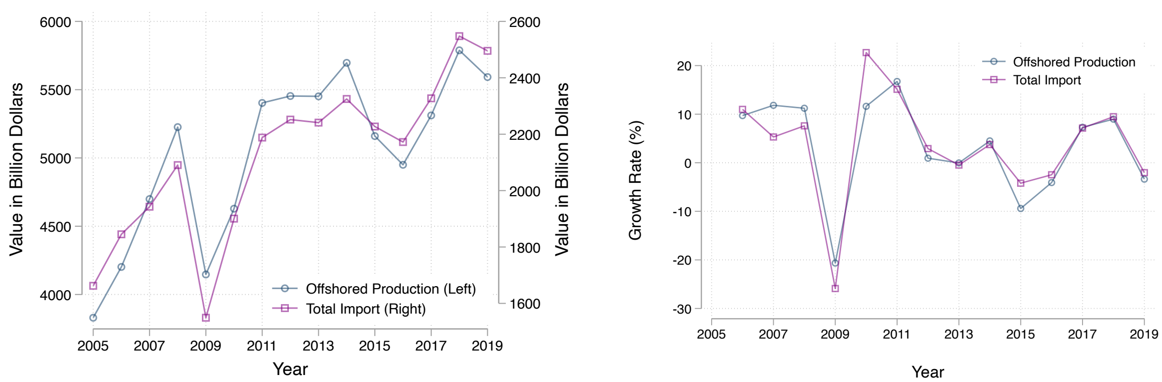
Notes: Figure presents a snapshot of Question 2-11 in "Section 2: R&D Paid For by Your Company" of the original survey form for the Business R&D and Innovation Survey. It lists 40 countries and regions together with five residual categories, namely "Other Europe", "Other Latin America/OWH", "Other Asia/Pacific", "Other Middle East", and "Other Africa".

Figure 4: Correlation between Offshored Production and Import

(a) Within-Firm Offshoring



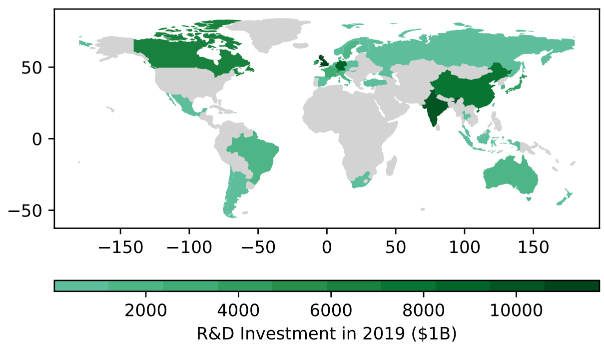
(b) Total Offshoring (Within-Firm Offshoring + Across-Firm Outsourcing)



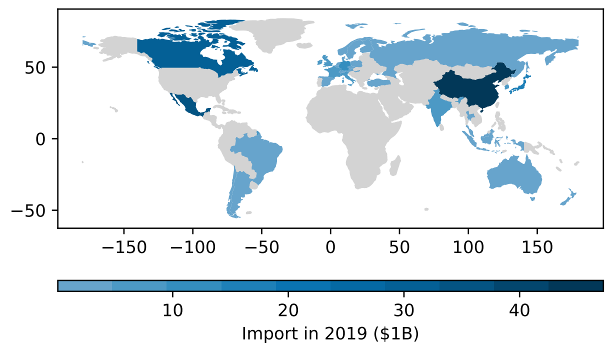
Notes: Figure compares the aggregate amount of offshored production and imports between 2005 and 2019 in order to evaluate the accuracy of using the latter as a proxy for the former. Panel (a) looks at within-firm offshoring. It compares the absolute amount and growth rate of the production offshored within the multinational firm to foreign affiliates with those of the import from related parties. Panel (b) looks at the sum of within-firm and outsourced foreign production. It compares the absolute amount and growth rate of the production offshored abroad (including those outsourced to other foreign firms) with those of the total import both from related parties and at arm's length.

Figure 5: Spatial Distribution of Offshoring Activities

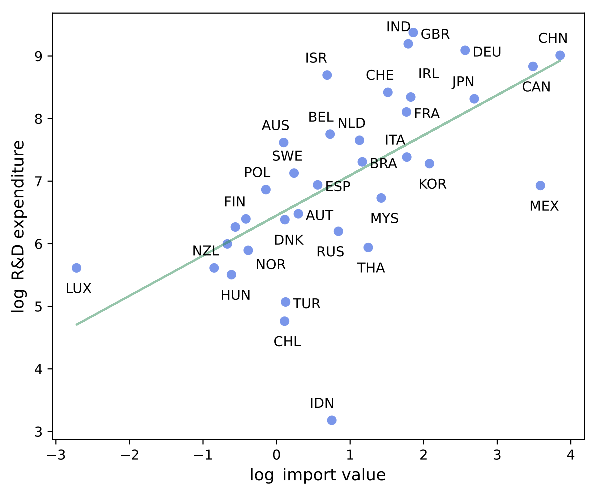
(a) Offshored R&D



(b) Imports (Proxy for Offshored Production)

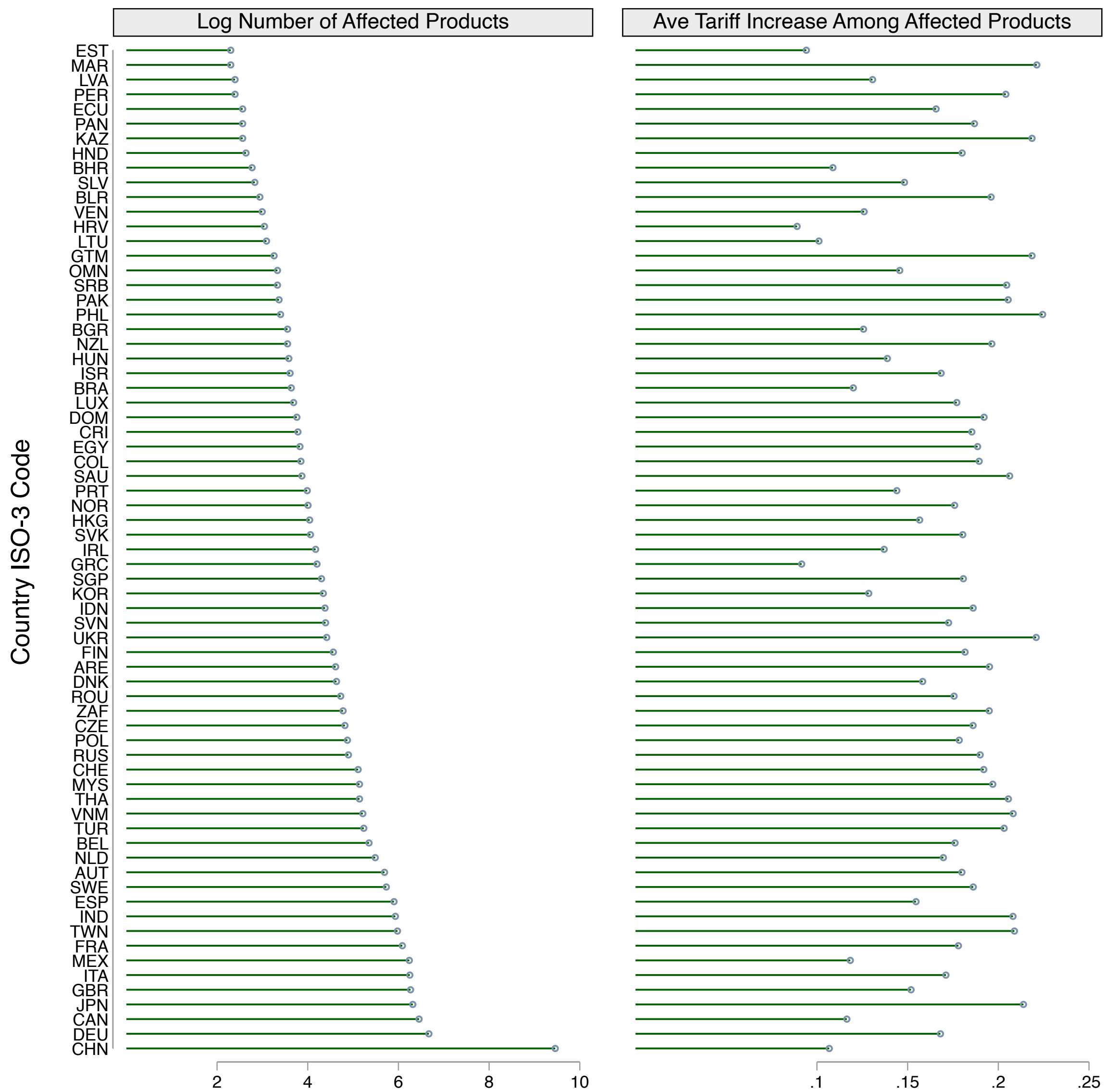


(c) Correlation Between Imports and Offshored R&D



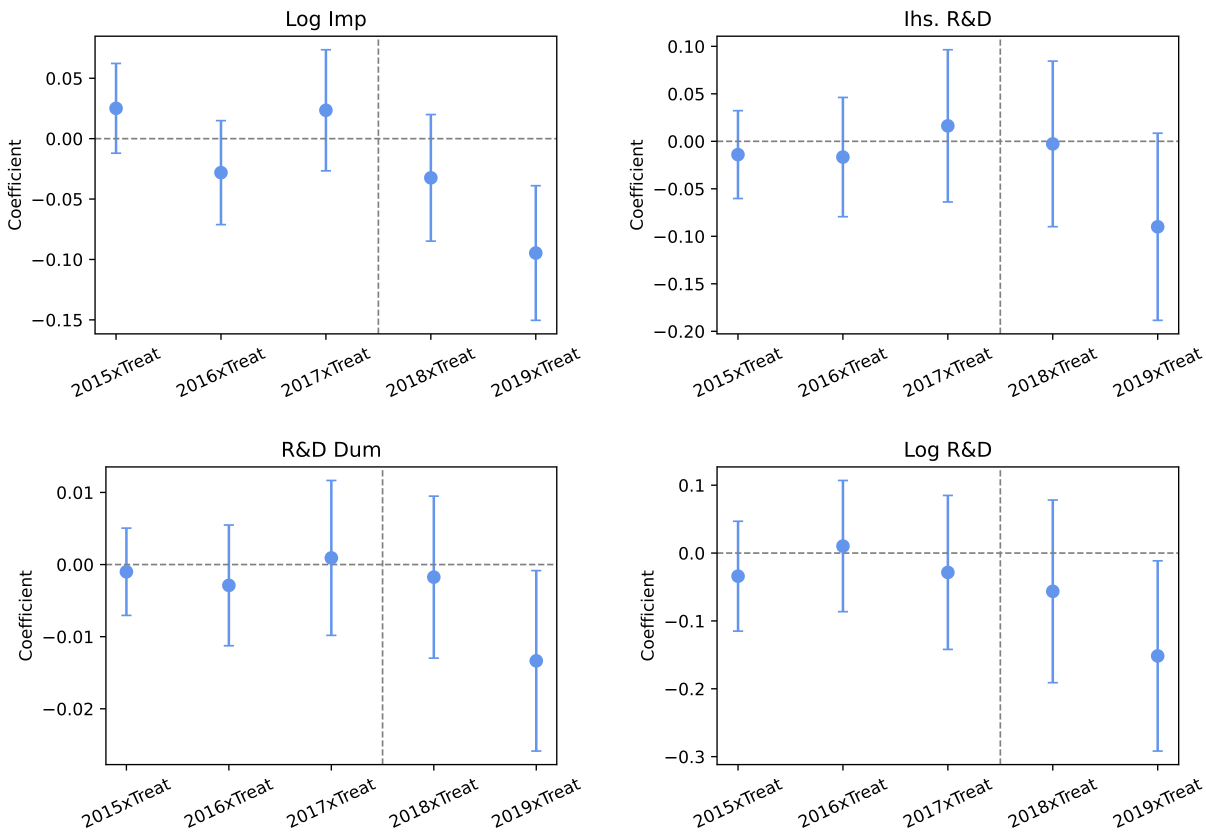
Notes: Panel A and B plot the global spatial distribution of US firms' offshoring activities in 2019. Panel C plots R&D against import value at the country level. Data on R&D offshoring is obtained from the public BRDIS statistics. Data on imports is obtained from the World Trade Organization (WTO). The gray color in the first two panels indicates missing data. Countries that are not surveyed in the BRDIS are categorized into residual groups, and they account for a small fraction of US offshored R&D in total.

Figure 6: Tariff Increases by Country During Trump Tariffs



Notes: Figure plots the log number of products that had tariff increases from Trump Tariffs during 2018 and 2019 for each country (in the left panel) and the average effective tariff rate increase in this period among affected products (in the right panel). Raw data on tariff increases is obtained from Fajgelbaum et al. (2020) and Fajgelbaum et al. (2022). The *effective* tariff rate increase for a product refers to the raw tariff increase scaled by the number of months in a year during which the increase was in effect. For example, if a 10 percentage point tariff increase were implemented for the product in July 2018 and lasted until the end of 2019, then the scaled tariff increase for this period would be 6.67 p.p., i.e. $10 * 18/24$.

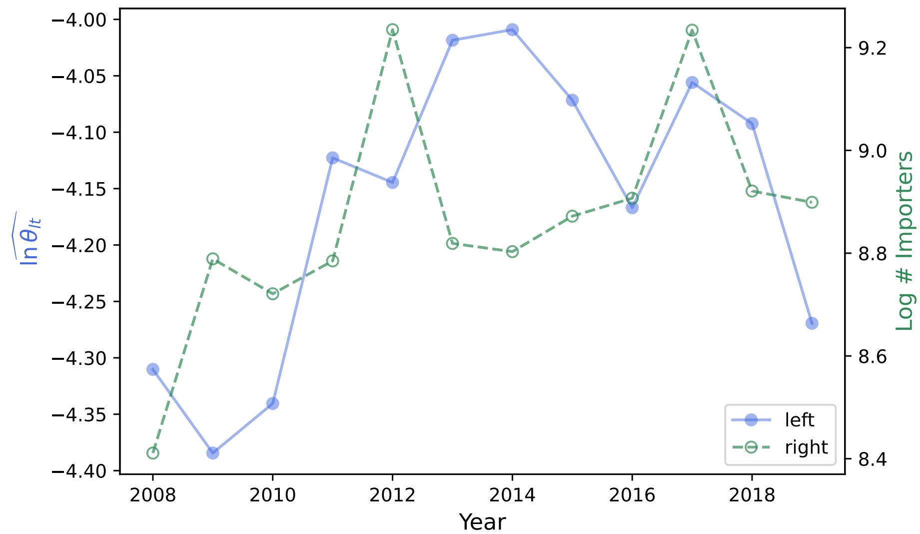
Figure 7: Effects of Trump Tariffs on Production and Innovation Offshoring



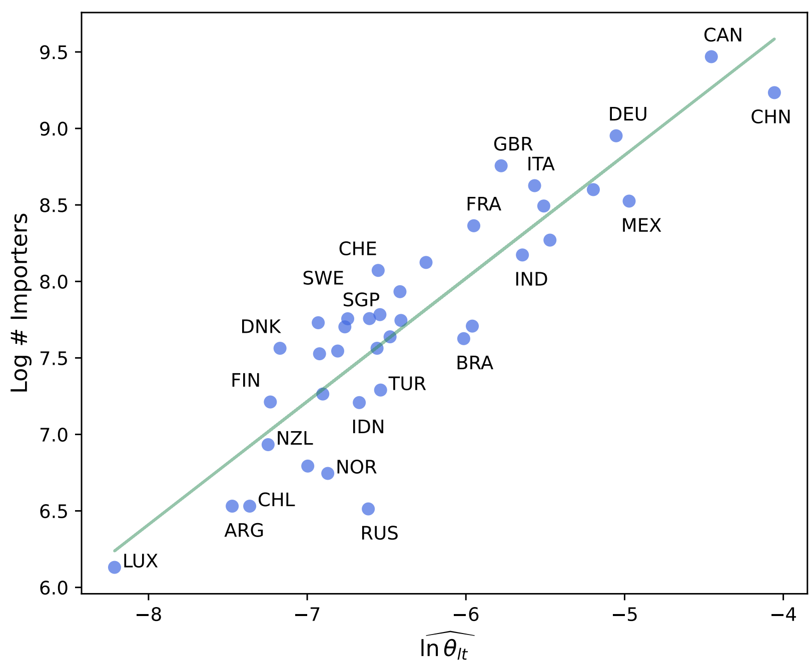
Notes: Figure plots coefficient estimates of the event study regressions as specified in Equation (4). The four panels correspond to four outcome variables: log import value, inverse hyperbolic transformation of R&D expenditure, indicator for positive R&D expenditure, and log R&D expenditure. See Subsection III.B for details on the quasi-natural experiment, i.e. Trump Tariffs. The treatment indicator equals one for firm-country pairs whose tariff rate was affected during Trump Tariffs. Firm-country and country-year fixed effects are controlled in the regressions. Standard errors are clustered at the firm level. 90% confidence intervals are plotted.

Figure 8: Estimates of Country Production Offshoring Potentials

(a) China, 2008-2019

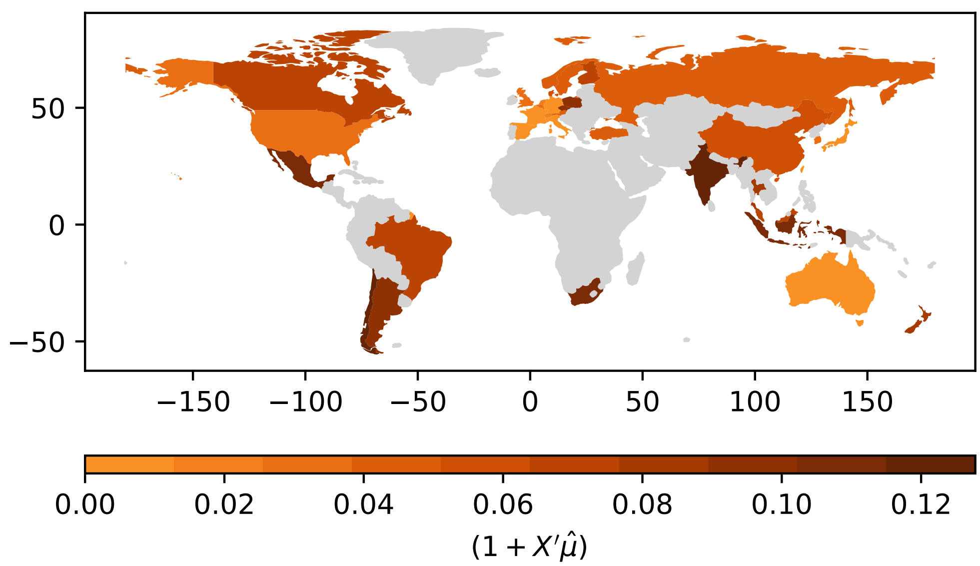


(b) All Countries, 2017



Notes: Figure plots the log of estimated production offshoring potential ($\log\hat{\theta}_{lt}$) against the number of importing firms. Panel A focuses on China's production offshoring potentials for US firms during 2008 and 2019. Panel B shows production offshoring potentials of all sample countries for US firms in the census year 2017. See Section V for details on the definition of production offshoring potentials and Section VI.B.1 for how they are estimated.

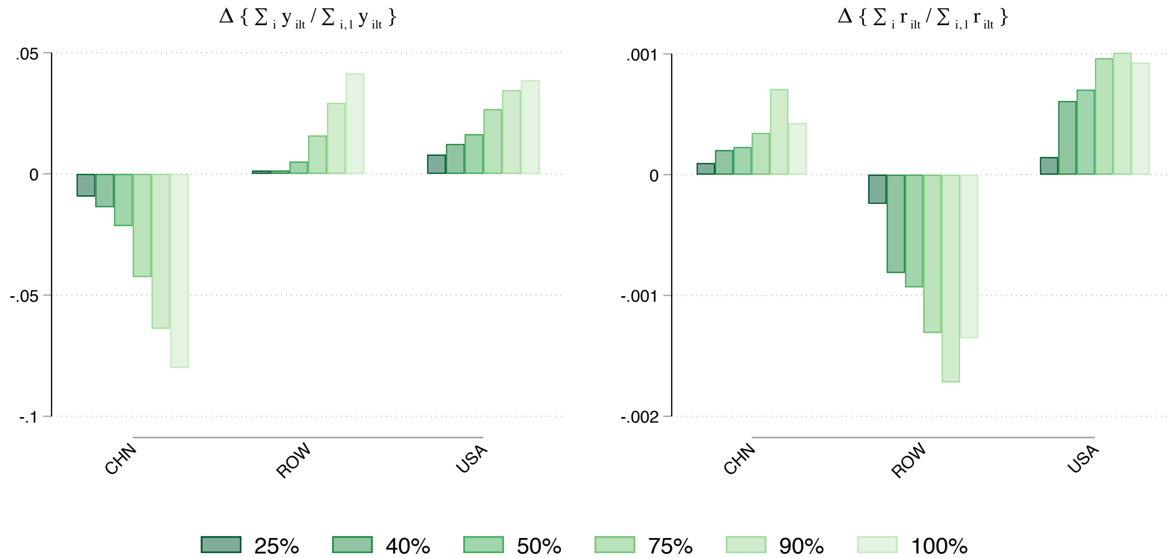
Figure 9: Estimates of Synergy Between Production and Innovation



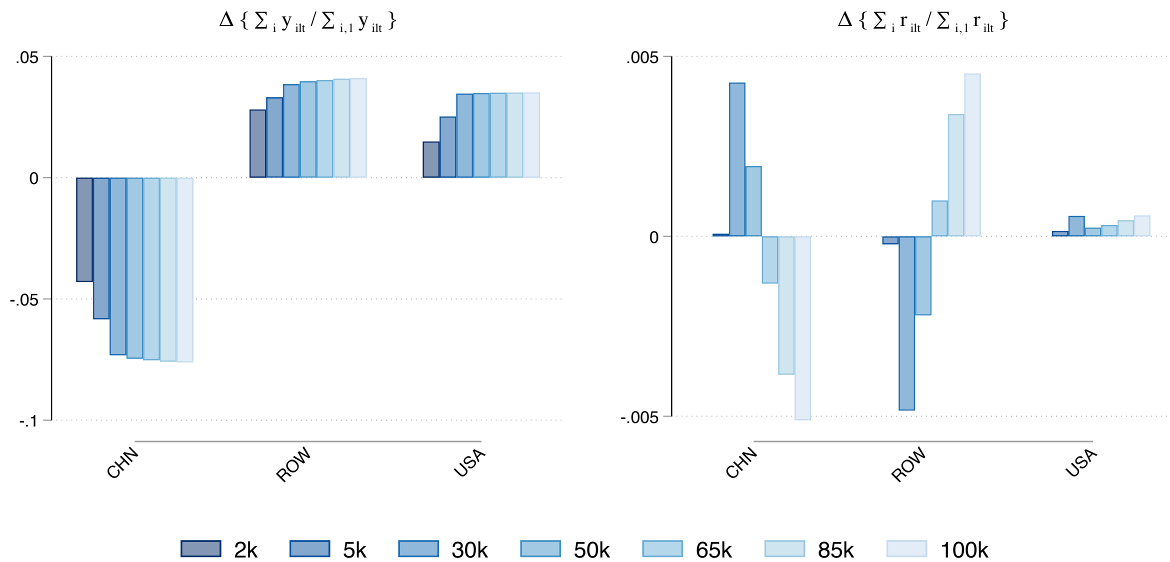
Notes: Figure plots the estimated coefficients $1+X\hat{\rho}$ in the productivity evolution process specified in Equation 6. X is a vector of country characteristics, including log income, human capital, log distance to the US, and capital services. $\hat{\rho}$ are the coefficient estimates in Column (6) of Table 10. Countries in grey are not surveyed in the BRDIS and thus not in the study sample. They are categorized into residual groups such as "Other African Countries" and together account for a small fraction of US offshored R&D.

Figure 10: Simulated Effects of Counterfactual Policy Shocks

Panel A: Increasing Import Tariffs

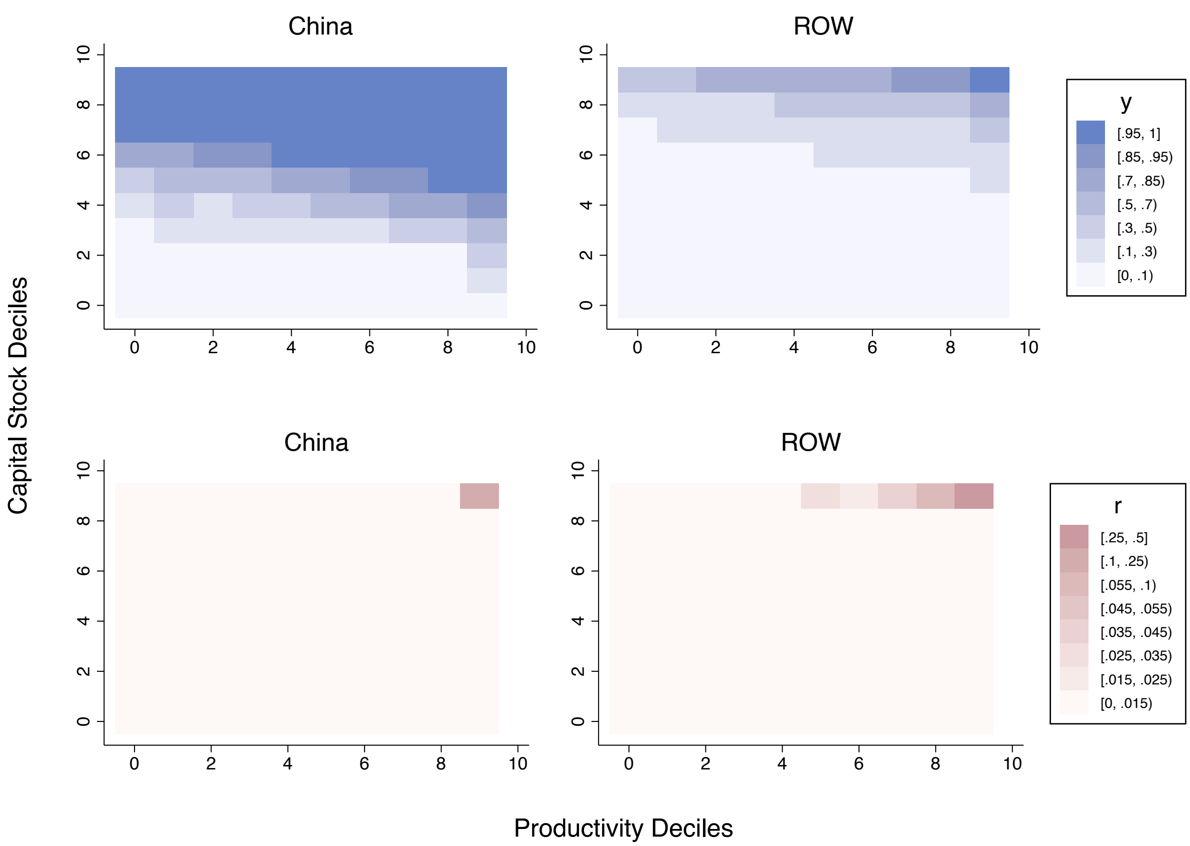


Panel B: Increasing Production Costs



Notes: Figure presents the counterfactual results depicting the changes in countries' production and innovation shares in response U.S. trade policy shocks that adversely affect production offshoring to China. In Panel A, each country is represented by six bars, corresponding to different reductions in China's production offshoring potential (25%, 40%, 50%, 75%, 90%, and 100%). In Panel B, each country is represented by seven bars, corresponding to different increases in the sunk and fixed costs of production offshoring to China (\$2K, \$5K , \$30K , \$50K , \$65K , \$85K , and \$100K). See Section VII.B.2 for further details on these two policy counterfactual exercises.

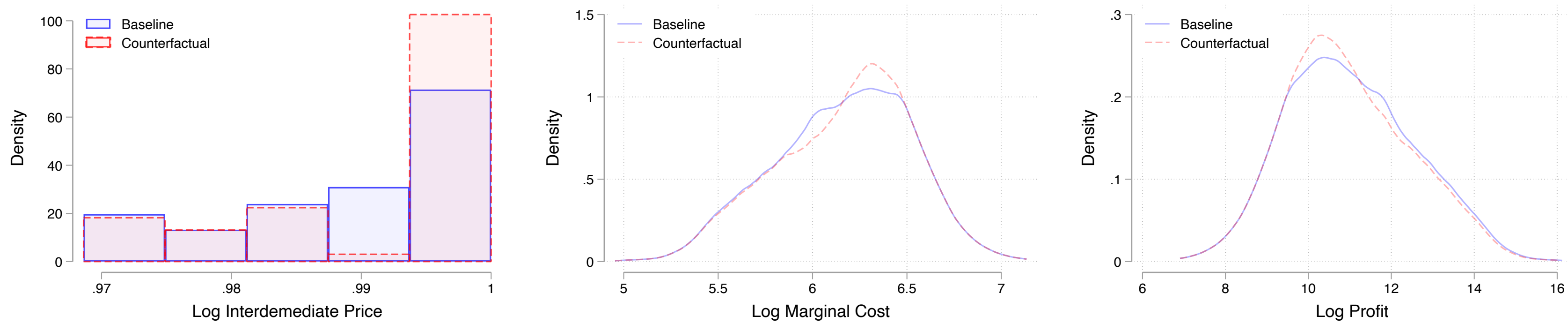
Figure 11: Firm Heterogeneity in Production and Innovation Offshoring



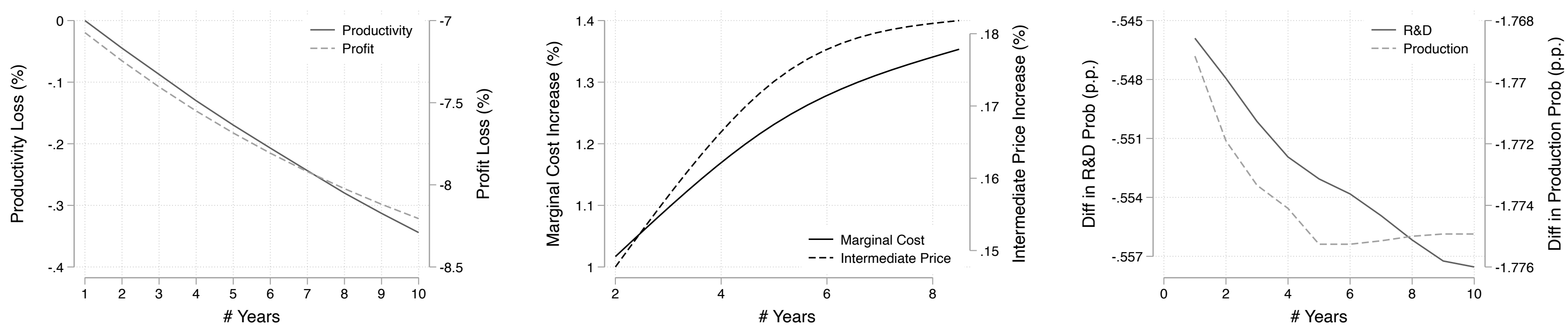
Notes: Figure depicts the fraction of firms that engage in production and innovation offshoring to China and ROW in 2017, simulated using the baseline model and categorized based on firms' productivity (on the x-axis) and capital stock (on the y-axis) deciles. The blue (upper) panels represent production offshoring, while the orange (lower) panels represent innovation offshoring.

Figure 12: Dynamic Effects of Trade Policies

Panel A: Static Losses



Panel B: Dynamic Losses



Notes: Figure depicts static losses (in Panel A) and dynamic losses (in Panel B) from a counterfactual exercise where I permanently reduce China's production offshoring potential by half. In Panel A, I plot the distributions of log marginal production cost, log intermediate price, and log profit in the first year of the shock. In Panel B, I plot the changes in productivity, log marginal production cost, and offshoring decisions over the number of years under the shock. See Section VII.B.3 for more details on this counterfactual exercise.

Table 1: Summary Statistics

Firm-Year Level	
Mean Sales ($K)	497700
Mean Emp	2945
Mean Domestic Emp	1797
Mean Foreign Emp	1145
Observations	85000
% Importing	83.40
Conditional on Importing	
# Imp Countries	8.007
Ave Imp Value ($K)	142800
% Performing R&D	57.95
Conditional on Performing R&D	
Mean R&D Expenditure ($K)	49400
Mean Domestic R&D Expenditure ($K)	38030
Mean Foreign R&D Expenditure ($K)	11380
% Performing Foreign R&D	19.65
Conditional on Doing Foreign R&D	
# Foreign R&D Countries	5.63
Mean Foreign R&D Expenditure ($1K)	61970
Firm-Country-Year Level	
% Importing	16.06
Conditional Imp Value ($1K)	17840
% Doing Foreign R&D	1.303
Conditional Foreign R&D Expenditure ($1K)	11010
Observations	3475000

Notes: Table presents summary statistics of the study sample at both the firm-year and firm-country-year levels. The study sample is constructed based on micro firm data from the BRDIS, LFTTD, and CMF/ASM. See Section II.A for more details on data sources and sample construction. Numbers are rounded to four effective digits according to Census data disclosure requirements.

Table 2: Top Offshoring Destinations for Production and Innovation

Top R&D Locations	% R&D Expenditure	Top Imp Locations	% Imp Value
Germany	14.76	Mexico	19.51
UK	11.32	Canada	17.76
China	8.25	China	12.58
India	6.78	Japan	8.18
Canada	5.38	Germany	7.16

Notes: Table presents descriptive statistics about top five R&D destination countries and import origin countries for US firms. Column (2) reports the fraction of foreign R&D expenditure of US firms in each destination country. Column (4) reports the fraction of import value of US firms for each origin country. Calculation is based on Census micro data (R&D expenditure from the BRDIS and import value from the LFTTD) during 2010-2019.

Table 3: Types of Offshoring Decisions For Production and Innovation

Type	% Obs	% Imp Value	% R&D Expenditure
None	83.75	0	0
Imp Only	14.94	62.26	0
R&D Only	0.19	0	6.17
Both	1.12	37.74	93.83
Total	100	100	100

Notes: Table presents descriptive statistics about the colocation pattern between import and offshored R&D. Observations are at the firm-country-year level and divided into four groups based on whether it is associated with positive R&D expenditure and positive import value. The fraction of observations and the shares of import value and R&D expenditure are reported for each group. Firm R&D expenditure is obtained from the BRDIS, and the import value is from the LFTTD. See Section II.A for more details on data sources and sample construction.

Table 4: Firm-Level Colocation of Offshore Production and Innovation

Panel A: R&D Offshoring on Imp.	(1)	(2)	(3)	(4)	(5)
	R&D Dum	R&D Dum	Log R&D	Log R&D	Ihs. R&D
Imp Dum	0.0195***		0.322***		
	(0.00109)		(0.119)		
Region Imp Dum	0.00147***		-0.00580		
	(0.000338)		(0.143)		
Log Imp		0.0150***		0.212***	
		(0.000761)		(0.0191)	
Log Region Imp		0.00167***		0.0105	
		(0.000626)		(0.0211)	
Ihs. Imp					0.0217***
					(0.00102)
Ihs. Region Imp					0.000936***
					(0.000233)
N	499000	41000	4100	3100	499000
R-squared	0.392	0.486	0.569	0.592	0.419
Firm FE	Yes	Yes	Yes	Yes	Yes
Country-Ind-FE	Yes	Yes	Yes	Yes	Yes

Panel B: Imp on R&D Offshoring.	(1)	(2)	(3)	(4)	(5)
	Imp Dum	Imp Dum	Log Imp	Log Imp	Ihs. Imp
R&D Dum	0.210***		1.763***		
	(0.00909)		(0.0546)		
Region R&D Dum	0.0591***		0.239***		
	(0.00634)		(0.0498)		
Log R&D		0.00498		0.325***	
		(0.00329)		(0.0308)	
Log Region R&D		0.000284		0.106***	
		(0.00428)		(0.0403)	
Ihs. R&D					0.576***
					(0.0161)
Ihs. Region R&D					0.126***
					(0.0100)
N	499000	2800	57000	2300	499000
R-squared	0.421	0.608	0.476	0.681	0.471
Firm FE	Yes	Yes	Yes	Yes	Yes
Country-Ind-FE	Yes	Yes	Yes	Yes	Yes

Notes: Table presents coefficient estimates for the regressions specified in XX for year 2017. In Panel A, the dependent variable y_{il} is R&D in country l, the independent variable x_{il} is import from country l, and the other independent variable x_{iR} is import from the same region with country l excluded. In Panel B, the independent and dependent variables are switched. For each variable, the extensive margin (captured by the dummy variable), the intensive margin (captured by the logged variable), and the combination of both margins (captured by the inverse hyperbolic transformation) are considered. Industries are identified by 3-digit NAICS codes. * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$. Standard errors are clustered at the firm level and reported in the parentheses.

Table 5: Causal Effects of Imports on R&D Offshoring, Using Firm-Specific Tariff Rates As Instrument

Panel A: Reduced-Form Regressions

	(1) Ihs. Imp	(2) Imp Dum	(3) Log Imp	(4) Ihs. R&D	(5) R&D Dum	(6) Log R&D
T_{ilt}	-1.906***	-0.0643*	-5.163***	-0.239***	-0.0281***	-0.728
	(0.504)	(0.0357)	(0.712)	(0.0811)	(0.0104)	(1.531)
N	1516000	1516000	317000	1516000	1516000	27500
R-sq	0.491	0.440	0.396	0.401	0.384	0.475
Firm-Year-FE	Yes	Yes	Yes	Yes	Yes	Yes
Country-Year-FE	Yes	Yes	Yes	Yes	Yes	Yes

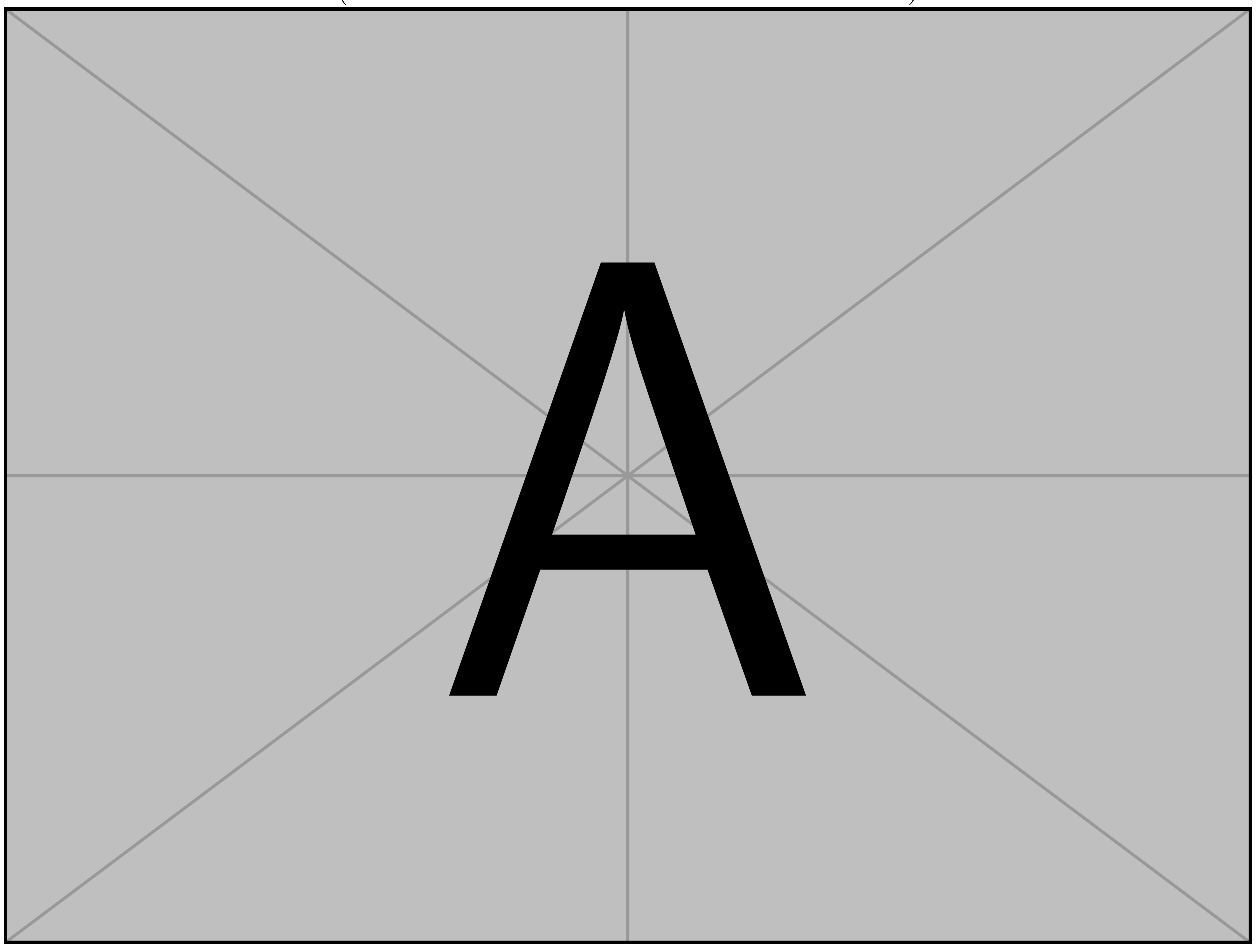
Panel B: Instrumented Regressions

	(1) Ihs. Imp	(2) Ihs. R&D	(3) Ihs. R&D	(4) R&D Dum	(5) R&D Dum
Ihs. Imp		0.0252***	0.125**	0.00315***	0.0147**
		(0.00111)	(0.0493)	(0.000131)	(0.00616)
T_{ilt}	-1.906***				
	(0.504)				
Method	OLS	OLS	IV	OLS	IV
1st-stage F	61.93				
N	1516000	1516000	1516000	1516000	1516000
Firm-Year-FE	Yes	Yes	Yes	Yes	Yes
Country-Year-FE	Yes	Yes	Yes	Yes	Yes

Notes: Table presents regression results of using the firm-specific tariff rate as an instrument for offshored production. See Subsection III.A for details on how the firm-specific tariff rate is constructed. Panel A shows the reduced-form regression results where import and offshored R&D are regressed on firm-specific tariff rate. In Panel B, Column (1) shows the first-stage result of regressing the inverse hyperbolic transformation of import value on the instrument, i.e. firm-specific tariff rate. Columns (2) and (4) show the OLS results, and Columns (3) and (5) show the second-stage results. For each variable, the extensive margin is captured by the dummy variable, the intensive margin is captured by the logged variable, and the combination of both margins is captured by the inverse hyperbolic transformation. The regressions focus on the period from 2013 to 2019. * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$. Standard errors are clustered at the firm level and reported in the parentheses.

Table 6: Cross-Country Interdependence With Firm-Specific Tariff Rates

(To be disclosed from the Census Bureau.)



Notes: Table presents regression results of using the firm-specific tariff rates in the focal country and region as instruments for offshored production in the focal country and region. See Subsection III.A for details on how the firm-specific tariff rate is constructed. Region import is the firm's total import value from the region with the focal country excluded. Panel A shows the reduced-form regression results where import and offshored R&D are regressed on firm-specific tariff rate in the focal country and region. In Panel B, Column (1) shows the first-stage result of regressing the inverse hyperbolic transformation of import value on the instruments. Columns (2) and (4) show the OLS results, and Columns (3) and (5) show the second-stage results. For each variable, the extensive margin is captured by the dummy variable, the intensive margin is captured by the logged variable, and the combination of both margins is captured by the inverse hyperbolic transformation. The regressions focus on the period from 2013 to 2019. * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$. Standard errors are clustered at the firm level and reported in the parentheses.

Table 7: Effects of Trump Tariffs, DID Regressions

	(1) Log Imp	(2) Ihs. R&D	(3) R&D Dum	(4) Log R&D
Treat $\times$ Post	-0.106***	-0.0945*	-0.0130**	-0.151**
	(0.0263)	(0.0494)	(0.00637)	(0.0659)
% Products Affected $\times$ Post	-0.258*	-0.189***	-0.0207***	-0.250
	(0.137)	(0.0567)	(0.00757)	(0.518)
% Product Value Affected $\times$ Post	-0.161*	-0.160***	-0.0175**	-0.243
	(0.0861)	(0.0596)	(0.00772)	(0.250)
Product-Count Weighted	-1.686**	-0.758**	-0.101**	3.031
Effective Tariff Increase $\times$ Post	(0.845)	(0.318)	(0.0421)	(2.958)
Product-Value Weighted	-1.098**	-0.642**	-0.0807**	0.119
Effective Tariff Increase $\times$ Post	(0.540)	(0.313)	(0.0394)	(1.760)
N	187000	187000	187000	16500
R-squared	0.889	0.877	0.838	0.893
Firm-Country-FE	Yes	Yes	Yes	Yes
Country-Year-FE	Yes	Yes	Yes	Yes

Notes: Table presents coefficient estimates of the DID regressions as specified in Equation (5). Each row represents a separate regression. See Subsection III.B for details on the quasi-natural experiment, i.e. Trump Tariffs. The treatment dummy equals one for firm-country pairs whose tariff rate was affected during Trump Tariffs. The Post dummy equals one for year 2019 and zero for years 2014-2017. Four additional measures of treatment size are considered: (1) "% Products Affected" is the fraction of products that were affected by Trump Tariffs among all products the firm imported in the prior period; (2) "% Product Value Affected" is the value share of affected products measured in the prior period; (3) "Product-Count Weighted Effective Tariff Increase" is the simple average of the effective tariff increase across the firm's imported products; (4) "Product-Value Weighted Effective Tariff Increase" is the weighted average of the effective tariff increase across the firm's imported products where the weights are the import value shares in the prior period. The four columns correspond to four outcome variables: log import, inverse hyperbolic transformation of R&D expenditure, indicator for positive R&D expenditure, and log R&d expenditure. * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$. Standard errors are clustered at the firm level and reported in the parentheses.

Table 8: Model Parameters and Sources of Identification

Parameter	**Source of Identification**
η	Average markup.
ρ	Response of country production offshoring potential to tariff change.
β_k, β_m	Relationship between output and input factors.
$\alpha_0, \alpha_1, \sigma_\xi$	Persistence and variation in firm productivity.
β_1, β_2, μ	Relationship between productivity change and innovation efforts in each country.
$\phi_s^p, \phi_s^r, \phi_f^p, \phi_f^r$	Fraction of firms that offshore production and innovation (unconditional and conditional on past choices).
λ_1	Colocation of production and innovation in and out of the region.

Notes: Table lists model parameters and their sources of identification. See Section V for how these parameters enter the model and Section VI.B for their estimated values.

Table 9: Production Offshoring Potentials and Country Characteristics

	(1) $\ln \hat{\theta}_{lt}$	(2) $\ln \hat{\theta}_{lt}$	(3) $\ln \hat{\theta}_{lt}$
$\ln t_{lt}$	-2.739*	-2.952***	-3.697***
	(1.567)	(1.123)	(1.110)
Log Population		0.358***	0.580***
		(0.0203)	(0.0252)
Common Language Dum		0.0246	-0.109*
		(0.0820)	(0.0601)
Colony Dum		0.0622	-0.210***
		(0.0712)	(0.0535)
Human Capital Index			0.657***
			(0.0840)
Control of Corruption Index			0.230***
			(0.0467)
N	450	450	450

Notes: Table presents coefficient estimates from regressing the log estimated production offshoring potential ($\ln \hat{\theta}_{lt}$) on log import tariff rate and other country characteristics. Regression is based on a country-year panel. Control variables include the country's log population, an indicator for whether the country has the same official language as the US, an indicator for whether the country had direct or indirect colony relationships with the US, an index for human capital, and an index for control of corruption. See Section II.A for sources of country characteristics data. * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$. Standard errors are in the parentheses.

Table 10: Estimation of Production Function and Productivity Evolution Process

	All Countries			No Tax Havens		
	(1)	(2)	(3)	(4)	(5)	(6)
	$\hat{\phi}_{it}$	$\hat{\phi}_{it}$	$\hat{\phi}_{it}$	$\hat{\phi}_{it}$	$\hat{\phi}_{it}$	$\hat{\phi}_{it}$
β_k	-0.165***	-0.165***	-0.165***	-0.165***	-0.165***	-0.165***
	(0.0014)	(0.0014)	(0.0014)	(0.0014)	(0.0014)	(0.0014)
β_m^F	0.418***	0.418***	0.417***	0.417***	0.418***	0.417***
	(0.0038)	(0.0038)	(0.0038)	(0.0038)	(0.0038)	(0.0038)
α_0	-0.0388***	-0.0388***	-0.0386***	-0.0389***	-0.0388***	-0.0386***
	(0.0024)	(0.0024)	(0.0024)	(0.0024)	(0.0024)	(0.0024)
α_1	0.918***	0.918***	0.917***	0.918***	0.918***	0.917***
	(0.0037)	(0.0037)	(0.0037)	(0.0037)	(0.0037)	(0.0037)
β_1	-0.0006	-0.0066	-0.0101	-0.0007	-0.0069	-0.0116
	(0.0005)	(0.0051)	(0.0075)	(0.0006)	(0.0054)	(0.0086)
β_2	0.00140**	0.0161**	0.0298***	0.0016**	0.0176**	0.0340***
	(0.0006)	(0.0069)	(0.0103)	(0.0007)	(0.0071)	(0.0115)
ρ: Log Income		-0.0881***	-0.0755***		-0.0891***	-0.0905***
		(0.0029)	(0.0202)		(0.0026)	(0.0181)
ρ: Human Capital			0.0148			0.0463
			(0.0430)			(0.0391)
ρ: Log Distance			-0.0332***			-0.0265***
			(0.0110)			(0.0095)
ρ: Capital Services			-0.0322***			-0.0279***
			(0.0093)			(0.0088)
N	28500	28500	28500	28500	28500	28500
VCE	hc2	hc2	hc2	hc2	hc2	hc2
Mean Elasticity	0.0014	0.0013	0.0016	0.0016	0.0015	0.0017
SD of Elasticity	0	0.0009	0.0012	0	0.0009	0.0013
RMSE	0.126	0.126	0.126	0.126	0.126	0.126

Notes: Table presents coefficient estimates for the NLS regression specified in Equation (10). Columns (1) to (3) include all sample countries while Columns (4) to (6) exclude countries known as tax havens. See Section VI.B.2 for details on the estimation equation and the identification of tax havens. β_m^F is the elasticity of log unit production cost with respect to log foreign intermediate price index. See Appendix B.2 for the relationship between β_m^F and β_m. Income is measured by real GDP per capita. Log Distance is the log of the country's distance to the US. Human capital and capital services are indices obtained from Penn World Tables. * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$. Heteroskedasticity robust standard errors are reported in the parentheses. The table also reports the mean and standard deviation of R&D elasticities implied by the coefficient estimates (see Section VI.B.2 for a detailed discussion). The last row reports the root mean squared errors to form an estimate of σ_ξ.

Table 11: Estimates of Offshoring Costs

Panel A: Parameter Estimates

ϕ_s^p	ϕ_f^p	ϕ_s^r	ϕ_f^r	λ_1
1171.22	1137.20	44210.76	43851.27	222.88
(49.83)	(54.78)	(8894.96)	(5808.21)	(51.97)

Panel B: Matched Moments

Moment	Data	Model
$E[y_{ilt}]$	0.16059	0.16054
$E[r_{ilt}]$	0.01303	0.01282
$E[y_{ilt}\,(1-y_{ilt-1})]$	XX	0.01933
$E[r_{ilt}\,(1-r_{ilt-1})]$	XX	0.00186
$E[y_{ilt}y_{il't}\|c_{ll'}=1]-E[y_{ilt}y_{il't}\|c_{ll'}=0]$	XX	0.02974
$E[r_{ilt}r_{il't}\|c_{ll'}=1]-E[r_{ilt}r_{il't}\|c_{ll'}=0]$	XX	0.00212

Notes: Panel A reports estimated values of dynamic cost parameters. The unit is a thousand dollars. Standard errors are reported in the parentheses. Panel B lists the six moments used in the MSM. The second and third columns show their empirical values from the data and thier simulated values based on the model, separately.

Table 12: Counterfactual - Mechanisms Behind Colocation Patterns

Panel A: Significance of Synergy Effect

	(1) $\beta_2 = \hat{\beta}_2$	(2) $\beta_2 = \frac{3}{4}\hat{\beta}_2$	(3) $\beta_2 = \frac{1}{2}\hat{\beta}_2$	(4) $\beta_2 = \frac{1}{4}\hat{\beta}_2$	(5) $\beta_2 \approx 0$
$\mathbb{E}[y_{ilt}]$	0.1605	0.1586	0.1595	0.1595	0.1595
	(100)	(98.82)	(99.38)	(99.38)	(99.38)
$\mathbb{E}[r_{ilt}]$	0.0128	0.0044	0.0012	0.0	0.0
	(100)	(34.38)	(9.37)	(0.0)	(0.0)
$\mathbb{E}[r_{ilt} \| y_{ilt} = 1]$	0.0798	0.0276	0.0078	0.0	0.0
	(100)	(34.59)	(9.77)	(0.0)	(0.0)

Panel B: Significance of Cost Sharing Effect

	(1) $\lambda_1 = \hat{\lambda}_1$	(2) $\lambda_1 = 0$	(3) Δ
$\mathbb{E}[y_{ilt}]$	0.1605	0.1603	0.0003
	(100)	(99.84)	(0.16)
$\mathbb{E}[r_{ilt}]$	0.0128	0.0118	0.001
	(100)	(92.45)	(7.55)
$\mathbb{E}[r_{ilt} \| y_{ilt} = 1]$	0.0798	0.0739	0.0059
	(100)	(92.59)	(7.41)
$\mathbb{E}[r_{il't} \| y_{ilt} = 1, c_{ll'} = 1, l' \neq l]$	0.0712	0.0659	0.0053
	(100)	(92.54)	(7.46)
$\mathbb{E}[r_{iRt} \| y_{iRt} = 1]$	0.0629	0.058	0.0049
	(100)	(92.16)	(7.84)

Notes: This table presents outcomes for the counterfactual exercises described in Section VII.A. In Panel A, the knowledge spillover parameter β_2 (see Section V.B for the definition of β_2) is reduced from its baseline estimate gradually to zero. In Panel B, the cost-sharing parameter λ_1 (see Section V.B for the definition of λ_1) is reduced from its baseline estimate to zero. Column (3) of Panel B reports the difference between the first two columns. In the last row of Panel B, y_{iRt} and r_{iRt} are dummy variables that equal one if the firm has production or innovation in region R. For both panels, simulated values for corresponding model moments are reported, and the relative changes in percentage are calculated in parentheses.

Appendices

A Data Construction

A.1 Firm Level Variables

Total output

$$Y_{it} = \begin{cases} (\mathrm{TVS}_{it} + \mathrm{FIE}_{it} - \mathrm{FIB}_{it} + \mathrm{WIE}_{it} - \mathrm{WIB}_{it} - \mathrm{CR}_{it}) / \mathrm{PISHIP}_{jt} & \text{if } Y_{it} > 0 \\ \mathrm{TVS}_{it} / \mathrm{PISHIP}_{jt} & \text{otherwise} \end{cases}$$

where TVS is the total value of shipments, FIE and FIB are the total value of finished goods inventories at the end and beginning of the year respectively, WIE and WIB are the work-in-progress inventories at the end and beginning of the year respectively, and CR is the cost of resales, all in nominal dollars. They are deflated by PISHIP, which is the four-digit industry level shipments deflator from the NBER-CES Manufacturing Database.

Labor input is defined as total hours worked. The ASM surveys the total number of employees, the number of production workers, and the total number of hours worked by production workers.

$$\mathrm{TH}_{it} = \begin{cases} \frac{\mathrm{PH}_{it} \cdot \mathrm{SW}_{it}}{\mathrm{WW}_{it}} & \text{if } \mathrm{SW}_{it} > 0, \mathrm{WW}_{it} > 0 \\ \mathrm{PH}_{it} & \text{otherwise} \end{cases}$$

where PH is the production worker hours, SW is the total wages including supplementary labor costs, and WW is the wages of production workers. Alternatively, we can use the same methodology as BLS to estimate nonproduction worker hours and compute the total annual hours worked using the following equation

$$\mathrm{TH}_{it}^{\mathrm{BLS}} = \mathrm{PH}_{it} + (\mathrm{TE}_{it} - \mathrm{PW}_{it}) \cdot \frac{\mathrm{PH}_{it}}{\mathrm{PW}_{it}} \cdot \left(\frac{\mathrm{AWH}_{\mathrm{NP}}^{\mathrm{CPS}}}{\mathrm{AWH}_{\mathrm{P}}^{\mathrm{CPS}}} \right)_{jt}$$

where TE is total employment, PW is the number of production workers, and $\frac{\mathrm{AWH}_{\mathrm{NP}}^{\mathrm{CPS}}}{\mathrm{AWH}_{\mathrm{P}}^{\mathrm{CPS}}}$ is the ratio of non-production to production average weekly hours for the 4-digit NAICS industry.

Materials (excluding energy) is defined as the real value of non-energy materials inputs:

$$\mathrm{M}_{it} = \left(\mathrm{CP}_{it} + \mathrm{CR}_{it} + \mathrm{CW}_{it}\right) / \mathrm{PIMAT}_{jt}$$

where CP is the total cost of materials and parts, CW is the total cost of contract work done by others, and PIMAT is the NBER-CES 4-digit industry level materials deflator.

Energy cost is defined as

$$E_{it} = \left(\mathrm{EE}_{it} + \mathrm{CF}_{it}\right) / \mathrm{PIEN}_{jt}$$

where EE is the cost of purchased electricity, CF is the cost of fuels, and PIEN is the NBER-CES 4-digit industry level energy deflator.

Capital stock K_{it} is not directly available in the ASM and CMF and is thus constructed using the Perpetual Inventory method for equipment and structures separately. The detailed procedures involving estimating initial values and discounting are as described in Cunningham et al. (2021).

Total variable cost TVC_{it} is the sum of SW_{it}, M_{it}, E_{it}, and total capital expenditures (variable TCE).

A.2 Country Level Variables

Wage. Monthly wages are downloaded from the International Labor Organization (ILO). I use the reported monthly earnings in local currencies and convert them to US dollars using exchange rates from the Penn World Tables.[22] I further deflate nominal wages using the GDP deflator obtained from the Bureau of Economic Analysis (BEA).

[22] The ILO does provide a harmonized series in US dollars; however, it contains many missing data and would compromise the sample size. In addition, I use wages in local currencies instead of purchasing power parity because the goal is to capture the differences in cost of production rather than consumption.

As in Eaton and Kortum (2002), I adjust wages for human capital by multiplying wages in country l by $\exp^{-gH_l}$, where g is the return to education and H_l is the years of schooling in country l in year 2010.[23] g is set to 0.06, which Bils and Klenow (2000) suggest is a conservative estimate. Data on schooling comes from Barro and Lee (2013).

[23]The years of schooling measure is constructed for every five years, and 2010 is the closest to my initial year 2008.

B Model Appendix

B.1 Microfoundation of CES Input Aggregation

Consider a framework of input sourcing as in Antras et al. (2017). Each firm sources a continuum of intermediate varieties, $v \in [0, 1]$. The varieties aggregate to the firm's intermediate accroding to CES,

$$m_{it} = \left[\int_0^1 q_i (v)^{\frac{\sigma-1}{\sigma}} \, dv \right]^{\frac{\sigma}{\sigma-1}} .$$

Let v_{it} denote the optimal price for sourcing input v. The price index of intermediate is then

$$p_{it}^m = \left[\int_0^1 z_{it} (v)^{1-\sigma} \, dv \right]^{\frac{1}{1-\sigma}} .$$

The firm always sources variety v from the cheapest location, therefore,

$$z_{it} (v) = \min_{l \in \mathcal{L}_{it}} \left\{ w_{lt} \cdot \tau_{lt} \cdot t_{lt} \cdot a_{lt} (v) \right\} ,$$

where $a_{lt} (v)$ is the unit labor requirement for producing v in country l at time t.

Assume Fréchet distribution such that

$$\Pr \left(a_{lt} (v) \geq a \right) = e^{-T_{lt} \cdot a^{\theta}} ,$$

with $T_l > 0$ capturing the technology level of country l and $\theta > 0$ capturing the dispersion in productivity.

Then it can be shown that the price for intermediate is

$$p_{it}^m = \left[c_0 \cdot \underbrace{\sum_{l \in \mathcal{L}_{it}} T_{lt} \left(w_{lt} \tau_{lt} t_{lt} \right)^{-\theta}}_{\Theta_{it} \text{:sourcing capability}} \right]^{-\frac{1}{\theta}}$$

where

$$c_0 = \left[\Gamma \left(\frac{\theta + 1 - \rho}{\theta} \right) \right]^{\frac{\theta}{1-\rho}} .$$

The share of sourcing for each country is

$$\chi_{il}(\varphi) = \frac{T_{lt}\left(w_{lt}\tau_{lt}t_{lt}\right)^{-\theta}}{\Theta_{it}}.$$

This is equivalent to a CES aggregation with elasticity of substitution being $1+\theta$ and unit labor productivity varying by country as $(c_0 T_j)^{-\frac{1}{\theta}}$.

B.2 Deriving β_m^F from β_m

The overall price index for intermediate goods is defined to be

$$p_{it}^m = \left(1 + \sum_{l>0} y_{ilt}\theta_{lt}\right)^{1/(1-\rho)},$$

and that for foreign intermediate goods is

$$p_{it}^{m,F} = \left(\sum_{l>0} y_{ilt}\theta_{lt}\right)^{1/(1-\rho)}.$$

Combining these two equations, I derive the following relationship between two price indices:

$$\left(p_{it}^m\right)^{1-\rho} = 1 + \left(p_{it}^{m,F}\right)^{1-\rho},$$

or equivalently,

$$\ln p_{it}^m = \frac{\ln\left(1 + e^{(1-\rho)\cdot\ln p_{it}^{m,F}}\right)}{1-\rho}.$$

Next, let's define a function

$$y = f(x) = \ln\left(1 + e^{(1-\rho)x}\right).$$

Taking the first-order approximation of $f(x)$ around x_0 implies

$$f(x) \approx \frac{1}{1-\rho}\left[\ln\left(1 + e^{(1-\rho)x_0}\right) - \frac{e^{(1-\rho)x_0}\cdot(1-\rho)}{1+e^{(1-\rho)x_0}}x_0\right] + \frac{e^{(1-\rho)x_0}}{1+e^{(1-\rho)x_0}}\cdot x.$$

Plugging in $y = \ln p_{it}^{m}$ and $x = \ln p_{it}^{m,F}$ to achive the first-order approximation of the relationship between two price indices:

$$\ln p_{it}^{m} \approx C + \frac{1}{1 + e^{(\rho-1)x_0}} \cdot \ln p_{it}^{m,F},$$

where

$$C = \frac{1}{1-\rho}\left[\ln\left(1 + e^{(1-\rho)x_0}\right) - \frac{e^{(1-\rho)x_0} \cdot (1-\rho)}{1 + e^{(1-\rho)x_0}} x_0\right].$$

It follows that

$$\partial \ln p_{it}^{m} = \frac{1}{1 + e^{(\rho-1)x_0}} \cdot \partial \ln p_{it}^{m,F}$$

and thus

$$\beta_m^F \equiv \frac{\partial \ln c_{it}}{\partial \ln p_{it}^{m,F}} = \frac{1}{1 + e^{(\rho-1)x_0}} \frac{\partial \ln c_{it}}{\partial \ln p_{it}^{m}} = \frac{1}{1 + e^{(\rho-1)x_0}} \beta_m$$

where

$$\beta_m \equiv \frac{\partial \ln c_{it}}{\partial \ln p_{it}^{m}}.$$

Finally, evaluating x_0 at the mean value of $\ln p_{it}^{m,F}$, 1.0274, implies

$$\hat{\beta}_m^F = \frac{\hat{\beta}_m}{1 + e^{(\hat{\rho}-1)\times 1.0274}} = 0.42.$$

B.3 Proof of Proposition 1

B.3.1 Notations

The expected lifetime payoff function Π_0 can be decomposed as

$$\Pi_0\left(\boldsymbol{o}_i\right) = \sum_{\boldsymbol{z} \in \Omega} \Pr\left(\boldsymbol{z}\right) \Pi^{\dagger}\left(\boldsymbol{o}_i | \boldsymbol{z}\right),$$

where $\Pi^{\dagger}\left(\boldsymbol{o}_i | \boldsymbol{z}\right)$ is the deterministic lifetime payoff following decision rule $\boldsymbol{o}_i$ under history $\boldsymbol{z}$:

$$\Pi^{\dagger}\left(\boldsymbol{o}_i | \boldsymbol{z}\right) = \sum_{t=0}^{\infty} \Pi_t\left(\omega_{it}\left(z^t, \{\boldsymbol{o}_i\left(z^{\tau}\right)\}_{\tau=0}^{t-1}\right), \boldsymbol{o}_i\left(z^t\right), \boldsymbol{o}_i\left(z^{t-1}\right)\right).$$

Since payoff in one history is independent of decisions rules along other histories, we can define $\tilde{\Pi}^{\dagger}\left(\boldsymbol{o}_{\boldsymbol{z}}|\boldsymbol{z}\right)=\Pi^{\dagger}\left(\boldsymbol{o}_{i}\left(\boldsymbol{z}\right)|\boldsymbol{z}\right)$, which is a function from $\{0,1\}^{2\mathcal{L}\mathcal{T}}$ to $\mathbb{R}$. Note that $\tilde{\Pi}^{\dagger}\left(\cdot|\boldsymbol{z}\right)$ is identical to $\Pi^{\dagger}\left(\cdot|\boldsymbol{z}\right)$, but written as only a function of the subvector of choices for all countries and periods in a given history $\boldsymbol{z}$.

B.3.2 Lemmas and Proofs

I use Lemma 2.6.1 in Topkis (1998), stated below.

Lemma 1. *Suppose X is a lattice. Then,*

1. *If $f(x)$ is supermodular on X and $\alpha>0$, then $\alpha f(x)$ is supermodular on X.*
2. *If $f(x)$ and $g(x)$ are supermodular on X, then $f(x)+g(x)$ is supermodular on X.*

I then state the second lemma that will be proved at the end of this section.

Lemma 2. $\tilde{\Pi}^{\dagger}\left(\boldsymbol{o}_{\boldsymbol{z}}|\boldsymbol{z}\right)$ *has increasing differences in* $\{0,1\}^{2\mathcal{L}\mathcal{T}}$.

Repeat and prove the main property here.

Proposition. $\Pi_{0}\left(\boldsymbol{o}_{i}|\boldsymbol{y}_{i,-1},\boldsymbol{r}_{i,-1},\omega_{i,-1}\right)$ *is supermodular in* $\boldsymbol{o}_{i}$ *on* $\{0,1\}^{2\mathcal{L}\mathcal{T}\Omega}$.

Proof of Proposition. From Lemma 2, $\tilde{\Pi}^{\dagger}\left(\boldsymbol{o}_{\boldsymbol{z}}|\boldsymbol{z}\right)$ has increasing differences in $\{0,1\}^{2\mathcal{L}\mathcal{T}}$. Using Corollary 2.6.1 in Topkis (1998), $\tilde{\Pi}^{\dagger}\left(\boldsymbol{o}_{\boldsymbol{z}}|\boldsymbol{z}\right)$ is supermodular in $\boldsymbol{o}_{\boldsymbol{z}}$ on $\{0,1\}^{2\mathcal{L}\mathcal{T}}$.

I then show that $\Pi^{\dagger}\left(\boldsymbol{o}_{i}|\boldsymbol{z}\right)$ is supermodular in $\boldsymbol{o}_{i}$ on $\{0,1\}^{2\mathcal{L}\mathcal{T}\Omega}$: consider two decision rules $\boldsymbol{o}_{i},\boldsymbol{o}_{i}'\in\{0,1\}^{2\mathcal{L}\mathcal{T}\Omega}$, it holds for any history $\boldsymbol{z}$ that

$$
\begin{aligned}
\Pi^{\dagger}\left(\boldsymbol{o}_{i}|\boldsymbol{z}\right)+\Pi^{\dagger}\left(\boldsymbol{o}_{i}'|\boldsymbol{z}\right) &= \tilde{\Pi}^{\dagger}\left(\boldsymbol{o}_{i}\left(\boldsymbol{z}\right)|\boldsymbol{z}\right)+\tilde{\Pi}^{\dagger}\left(\boldsymbol{o}_{i}'\left(\boldsymbol{z}\right)|\boldsymbol{z}\right)\\
&\leq \tilde{\Pi}^{\dagger}\left(\boldsymbol{o}_{i}\left(\boldsymbol{z}\right)\vee\boldsymbol{o}_{i}'\left(\boldsymbol{z}\right)|\boldsymbol{z}\right)+\tilde{\Pi}^{\dagger}\left(\boldsymbol{o}_{i}\left(\boldsymbol{z}\right)\wedge\boldsymbol{o}_{i}'\left(\boldsymbol{z}\right)|\boldsymbol{z}\right)\\
&= \tilde{\Pi}^{\dagger}\left(\boldsymbol{o}_{i}\vee\boldsymbol{o}_{i}'\left(\boldsymbol{z}\right)|\boldsymbol{z}\right)+\tilde{\Pi}^{\dagger}\left(\boldsymbol{o}_{i}\wedge\boldsymbol{o}_{i}'\left(\boldsymbol{z}\right)|\boldsymbol{z}\right)\\
&= \Pi^{\dagger}\left(\boldsymbol{o}_{i}\vee\boldsymbol{o}_{i}'|\boldsymbol{z}\right)+\Pi^{\dagger}\left(\boldsymbol{o}_{i}\wedge\boldsymbol{o}_{i}'|\boldsymbol{z}\right),
\end{aligned}
$$

where the first and last equality follow from the relationship between the functions $\tilde{\Pi}^{\dagger}\left(\cdot|\boldsymbol{z}\right)$ and $\Pi^{\dagger}\left(\cdot|\boldsymbol{z}\right)$, the inequality in the second line follows from the supermodularity of the function

$\tilde{\Pi}^{\dagger}(\boldsymbol{o_z}|\boldsymbol{z})$, and the equality in the third line follows from basic linear algebra rules. The "join" $\vee$ takes the maximum element by element, and the "meet" $\wedge$ takes the minimum element by element.

Finally, recall that

$$\Pi_0(\boldsymbol{o}_i) = \sum_{\boldsymbol{z}\in\Omega} \Pr(\boldsymbol{z})\,\Pi^{\dagger}(\boldsymbol{o}_i|\boldsymbol{z}).$$

Since from Lemma 1 we know that the finite sum of supermodular functions is supermodular, $\Pi_0(\boldsymbol{o}_i)$ is supermodular in $\boldsymbol{o}_i$ on $\{0,1\}^{2\mathcal{L}\mathcal{T}\Omega}$. □

Proof of Lemma 2. For a given history $\boldsymbol{z}$, unpack the decision rule vector as

$$\boldsymbol{o_z} = \left(\{y_{ilt}\}_{l\in\mathcal{L},t\in\mathcal{T}}, \{r_{ilt}\}_{l\in\mathcal{L},t\in\mathcal{T}}\right).$$

Note that I omit the notation of $\boldsymbol{z}$ in y_{ilt} since we are looking at a fixed $\boldsymbol{z}$ throughout this proof. The goal is to show that $\tilde{\Pi}^{\dagger}(\boldsymbol{o_z}|\boldsymbol{z})$ has increasing difference along y_{ilt} and r_{ilt} for any l and any t in the given history $\boldsymbol{z}$.

Increasing difference along y_{ilt}. Consider two decision rules $\boldsymbol{o_z}, \boldsymbol{o'_z} \in \{0,1\}^{2\mathcal{L}\mathcal{T}}$ where the only difference between them is that $y_{ilt} = 0$ and $y'_{ilt} = 1$ for a specific l and t. The difference between $\tilde{\Pi}^{\dagger}(\boldsymbol{o'_z}|\boldsymbol{z})$ and $\tilde{\Pi}^{\dagger}(\boldsymbol{o_z}|\boldsymbol{z})$ has the following components:

- An increase in variable profit π_{it} due to higher sourcing capability:

$$\Delta\pi_{it} = \frac{1}{\eta}\cdot\left(\frac{\eta}{\eta-1}\right)^{1-\eta}\cdot\Phi_{jt}\cdot\left(\frac{e^{\beta_0}}{e^{\omega_{it}}}\cdot k_i^{\beta_k}\cdot w_{it}^{\beta_w}\right)^{1-\eta}$$
$$*\left[\left((w_{lt}\tau_{lt})^{1-\rho} + \sum_{l'\neq l} y_{il't}\cdot(w_{l't}\tau_{l't})^{1-\rho}\right)^{\frac{(1-\eta)\beta_m}{1-\rho}} - \left(\sum_{l'\neq l} y_{il't}\cdot(w_{l't}\tau_{l't})^{\frac{\rho-1}{\rho}}\right)^{\frac{(1-\eta)\beta_m}{1-\rho}}\right].$$

- A change in the cost paid in period t:

$$-\phi_s^p + y_{ilt-1}\left(\phi_s^p - \phi_f^p\right) + \lambda_1\sum_{l_1} r_{il_1t-1}r_{il_1t}\left(\max_{l_2}\{c_{l_1l_2}y_{il_2t}|y_{ilt}=1\} - \max_{l_2}\{c_{l_1l_2}y_{il_2t}|y_{ilt}=0\}\right).$$

- A change in the cost paid in period $t+1$: $y_{ilt+1}\cdot\left(\phi_s^p - \phi_f^p\right)$.

- All future productivities, holding $\{\xi_{it}\}_t$ fixed, change by

$$\Delta\omega_{it+\tau} = \alpha^{\tau-1}\left[1 + X_{lt}\rho\right] \cdot \left[\beta_2 r_{ilt}\right], \tau \geq 1.$$

  This leads to changes in variable profit for all periods after t, the sum of which has the following first-order approximation:

$$\sum_{\tau=1}^{\infty} \frac{1}{\eta} \cdot \left(\frac{\eta}{\eta-1}\right)^{1-\eta} \cdot \Phi_{jt+\tau} \cdot \left(e^{\beta_0} \cdot k_{it+\tau}^{\beta_k} \cdot w_{it+\tau}^{\beta_w} \cdot \left(p_{it+\tau}^{m}\right)^{\beta_m}\right)^{1-\eta}$$
$$*\alpha^{\tau-1}\left[1 + X_{lt}\rho\right] \cdot \left[\beta_2 r_{ilt}\right] \cdot \exp\left((\eta-1) \cdot \omega_{it+\tau}\left(y_{ilt}=0\right)\right).$$

Combining the four components, the first-order change from $\tilde{\Pi}^{\dagger}\left(\boldsymbol{o_z}|\boldsymbol{z}\right)$ to $\tilde{\Pi}^{\dagger}\left(\boldsymbol{o}'_{\boldsymbol{z}}|\boldsymbol{z}\right)$ is thus

$$\begin{aligned}
\tilde{\Pi}^{\dagger}\left(\boldsymbol{o}'_{\boldsymbol{z}}|\boldsymbol{z}\right) - \tilde{\Pi}^{\dagger}\left(\boldsymbol{o_z}|\boldsymbol{z}\right) = {} & \frac{1}{\eta} \cdot \left(\frac{\eta}{\eta-1}\right)^{1-\eta} \cdot \Phi_{jt} \cdot \left(\frac{e^{\beta_0}}{e^{\omega_{it}}} \cdot k_i^{\beta_k} \cdot w_{it}^{\beta_w}\right)^{1-\eta} \\
& * \left[\left(\left(w_{lt}\tau_{lt}\right)^{1-\rho} + \sum_{l'\neq l} y_{il't} \cdot \left(w_{l't}\tau_{l't}\right)^{1-\rho}\right)^{\frac{(1-\eta)\beta_m}{1-\rho}} - \left(\sum_{l'\neq l} y_{il't} \cdot \left(w_{l't}\tau_{l't}\right)^{\frac{\rho-1}{\rho}}\right)^{\frac{(1-\eta)\beta_m}{1-\rho}}\right] \\
& - \phi_s^p + y_{ilt-1} \cdot \left(\phi_s^p - \phi_f^p\right) + y_{ilt+1} \cdot \left(\phi_s^p - \phi_f^p\right) \\
& + \lambda_1 \sum_{l_1} r_{il_1t-1} \cdot r_{il_1t} \cdot \left(\max_{l_2}\left\{c_{l_1l_2} y_{il_2t} | y_{ilt}=1\right\} - \max_{l_2}\left\{c_{l_1l_2} y_{il_2t} | y_{ilt}=0\right\}\right) \\
& + \sum_{\tau=1}^{\infty} \frac{1}{\eta} \cdot \left(\frac{\eta}{\eta-1}\right)^{1-\eta} \cdot \Phi_{jt+\tau} \cdot \left(e^{\beta_0} \cdot k_{it+\tau}^{\beta_k} \cdot w_{it+\tau}^{\beta_w} \cdot \left(p_{it+\tau}^{m}\right)^{\beta_m}\right)^{1-\eta} \\
& * \alpha^{\tau-1}\left[1 + X_{lt}\rho\right] \cdot \left[\beta_2 r_{ilt}\right] \cdot \exp\left((\eta-1) \cdot \omega_{it+\tau}\left(y_{ilt}=0\right)\right).
\end{aligned}$$

If $\frac{(1-\eta)\beta_m}{1-\rho} > 1$ then the first component is increasing in $\sum_{l'\neq l} y_{il't} \cdot \left(w_{l't}\tau_{l't}\right)^{1-\rho}$ and thus $y_{il't}$; vice versa if $\frac{(1-\eta)\beta_m}{1-\rho} < 1$. When $\phi_s^p > \phi_f^p$, the second and third components are increasing in $y_{ilt-1}, y_{ilt+1}, r_{ilt}, r_{ilt-1}, r_{il't}, r_{il't-1}$. The last component is increasing in r_{ilt} when $\beta_2\left(1 + X_{lt}\rho\right) > 0$. Therefore, $\tilde{\Pi}^{\dagger}\left(\boldsymbol{o_z}|\boldsymbol{z}\right)$ has increasing differences along y_{ilt} for any l and any t.

Increasing difference along r_{ilt}. Consider two decision rules $\boldsymbol{o_z}, \boldsymbol{o}'_{\boldsymbol{z}} \in \{0,1\}^{2\mathcal{L}\mathcal{T}}$ where the only difference between them is that $r_{ilt} = 0$ and $r'_{ilt} = 1$ for a specific l and t. Switching from the first to the second decision rule doesn't affect period t's variable profit, but it affects next period's cost and all future periods' productivities.

The difference between $\tilde{\Pi}^{\dagger}\left(\boldsymbol{o}'_{\boldsymbol{z}}|\boldsymbol{z}\right)$ and $\tilde{\Pi}^{\dagger}\left(\boldsymbol{o}_{\boldsymbol{z}}|\boldsymbol{z}\right)$ has the following components:

1. A change in period $t+1$'s cost:

$$r_{ilt+1} \cdot \left(\phi_s^r - \phi_f^r + \lambda_1 \max_{l'} \left\{c_{ll'} y_{il't+1}\right\}\right).$$

2. All future productivities will be increased by

$$\Delta\omega_{it+\tau} = \alpha^{\tau-1}\left[1 + X_{lt}\rho\right] \cdot \left[\beta_1 + \beta_2 y_{ilt}\right], \tau \geq 1.$$

   This leads to a first-order increase in all future periods' profit that is equal to

$$\begin{aligned}&\sum_{\tau=1}^{\infty} \frac{1}{\eta} \cdot \left(\frac{\eta}{\eta-1}\right)^{1-\eta} \cdot \Phi_{jt+\tau} \cdot \left(e^{\beta_0} \cdot k_{it+\tau}^{\beta_k} \cdot w_{it+\tau}^{\beta_w} \cdot \left(p_{it+\tau}^m\right)^{\beta_m}\right)^{1-\eta} \\ &* \alpha^{\tau-1}\left[1 + X_{lt}\rho\right] \cdot \left[\beta_1 + \beta_2 y_{ilt}\right] \cdot \exp\left(\left(\eta-1\right) \cdot \omega_{it+\tau}\left(r_{ilt}=0\right)\right).\end{aligned}$$

Combining the two components, the first-order change from $\tilde{\Pi}^{\dagger}\left(\boldsymbol{o}_{\boldsymbol{z}}|\boldsymbol{z}\right)$ to $\tilde{\Pi}^{\dagger}\left(\boldsymbol{o}'_{\boldsymbol{z}}|\boldsymbol{z}\right)$ is thus

$$\begin{aligned}\tilde{\Pi}^{\dagger}\left(\boldsymbol{o}'_{\boldsymbol{z}}|\boldsymbol{z}\right) - \tilde{\Pi}^{\dagger}\left(\boldsymbol{o}_{\boldsymbol{z}}|\boldsymbol{z}\right) = {}& r_{ilt+1} \cdot \left(\phi_s^r - \phi_f^r + \lambda_1 \max_{l'} \left\{c_{ll'} y_{il't+1}\right\}\right) \\ &+ \sum_{\tau=1}^{\infty} \frac{1}{\eta} \cdot \left(\frac{\eta}{\eta-1}\right)^{1-\eta} \cdot \Phi_{jt+\tau} \cdot \left(e^{\beta_0} \cdot k_{it+\tau}^{\beta_k} \cdot w_{it+\tau}^{\beta_w} \cdot \left(p_{it+\tau}^m\right)^{\beta_m}\right)^{1-\eta} \\ &* \alpha^{\tau-1}\left[1 + X_{lt}\rho\right] \cdot \left[\beta_1 + \beta_2 y_{ilt}\right] \cdot \exp\left(\left(\eta-1\right) \cdot \omega_{it+\tau}\left(r_{ilt}=0\right)\right).\end{aligned}$$

The first component is increasing in r_{ilt+1}, y_{ilt+1} and $y_{il't+1}$. The second component is increasing in y_{ilt} when $\left[1 + X_{lt}\rho\right] \cdot \left[\beta_1 + \beta_2 y_{ilt}\right] \geq 0$ and decreasing in $p_{it+\tau}^m$, implying that it is also increasing in $y_{ilt+\tau}$ and $y_{il't+\tau}$ for all $\tau \geq 1$. Therefore, $\tilde{\Pi}^{\dagger}\left(\boldsymbol{o}_{\boldsymbol{z}}|\boldsymbol{z}\right)$ has increasing differences along r_{ilt} for any l and any t.

Combining the increasing differences along y_{ilit} and r_{ilt} for any l and t, I have showed that $\tilde{\Pi}^{\dagger}\left(\boldsymbol{o}_{\boldsymbol{z}}|\boldsymbol{z}\right)$ has increasing differences in $\{0,1\}^{2\mathcal{L}\mathcal{T}}$. $\square$

B.4 More Counterfactual Exercises

B.4.1 US-China Decoupling

- Increase the costs of production and innovation in China to infinity, equivalent to eliminating US firms' activities in China.
- Total offshore production drops by 23%. Worldwide R&D drops by 32%.
- The number of firms that offshore production drops by 10%. The number of firms conducting R&D drops by 26%.
- Figure A3 reports the changes in production and innovation shares by country.

C Additional Tables and Figures

1. Table A3. Regression table for stylized fact 1 using 2017 data, where I only include host-country-specific explanatory variables. That is, region-specific variables on the RHS are excluded. The exact regression specification is the following:

$$y_{il} = \beta_1 \cdot x_{il} + \gamma_i + \gamma_{jl} + \varepsilon_{il}$$

2. Table A4. An alternative version of Table A3 but using the full-year panel data. I include firm-year and country-industry-year fixed effects in this regression specification:

$$y_{ilt} = \beta_1 \cdot x_{ilt} + \gamma_{it} + \gamma_{jlt} + \varepsilon_{ilt}$$

3. Table A5. An alternative version of Table 4 but using the full-year panel data. This regression includes region-specific terms on the RHS to capture cross-country interdependence in offshoring:

$$y_{ilt} = \beta_1 \cdot x_{ilt} + \beta_2 \cdot x_{iRt} + \gamma_{it} + \gamma_{jlt} + \varepsilon_{ilt}$$

4. Table A6. I examine the causal effect of offshore production in other countries within the destination region on the offshore innovation in the focal host country. For this purpose, I run the following DID regression:

$$y_{ilt} = \beta \cdot \text{Treat}_{il} \cdot \text{Post}_t + \text{TreatNeighbor}_{il} \cdot \text{Post}_t + \gamma_{il} + \gamma_{lt} + \varepsilon_{ilt},$$

   where $\text{TreatNeighbor}_{il}$ is a dummy variable that equals one if $\text{Treat}_{il'}$ equals one for any neighboring country l' that is in the same region as country l. Results in this table shows that he effect of Treat_{il} is still significantly negative as in Table 7, and the effect of $\text{TreatNeighbor}_{il}$ is negative and of nontrivial magnitude despite its statistical insignificance.

5. Figure A1. This figure presents results for the same counterfactual exercises as in Figure 10 but all each foreign country in my study sample.

6. Figure A2. This figure presents probabilities of offshoring production and innovation by

firm types for more host countries than the main Figure 11.

7. Figure A3. This figure presents counterfactual results for the scenario where the U.S. firms cannot offshore production or innovation to China. This is implemented by setting the corresponding sunk and fixed costs to infinity.

Table A1: Sample Structure and Survey Frequency

Survey Freq.	# Firms	% Firms	% Sales	% VA
1-2	27500	76.39	3.05	3.50
3-5	5000	13.89	7.39	8.17
6-9	2500	6.94	22.42	16.24
10-12	1400	3.89	67.14	72.09
Total	36000	100	100	100

Notes: Table presents statistics about firms' survey frequencies. The sample is constructed at the intersection of the BRDIS, LFTTD, and CMF/ASM. Section II.A provides more information on data sources and sample construction. The LFTTD and CMF covers all firms while the BRDIS and ASM are based on representative samples of firms. Between 2008 and 2019, a firm in my sample appeared in at least one year and at most 12 years. The counts, fractions, shares of sales, and shares of value added are reported for firm groups by survey frequency. Firm counts are rounded to hundreds according to Census data disclosure requirements.

Table A2: Multiple Offshoring Locations for Production and Innovation

Panel A: R&D Locations

# Foreign R&D Locations	% Obs	% Sales	% Worldwide R&D	% Foreign R&D
0	90.37	38.22	13.68	0.53
1	2.83	6.39	4.40	1.84
2-5	3.74	19.37	12.14	10.93
6-10	1.66	10.86	13.35	16.14
Above 10	1.40	25.16	56.44	70.56
Total	100	100	100	100

Panel B: Import Locations

# Foreign Imp Locations	% Obs	% Sales	% Imp Value
0	16.60	0.52	0
1	13.18	0.88	0.09
2-10	48.26	14.18	4.97
11-20	14.39	29.10	19.55
Above 20	7.58	55.31	75.38
Total	100	100	100

Notes: Table presents descriptive statistics about the number of offshoring locations for firms in my study sample during 2008-2019. Observation is at the firm-year level. Panel A reports the fractions of observations, sales, worldwide R&D expenditure, and foreign R&D expenditure for firm groups based on how many foreign countries they perform R&D in. Panel B reports the fractions of observations, sales, and import value for firm groups based on how many origin countries they import from.

Table A3: Colocation of Offshore Production and Innovation - 2017

Panel A: R&D Offshoring on Imp.

	(1) R&D Dum	(2) R&D Dum	(3) Log R&D	(4) Log R&D	(5) Ihs. R&D
Imp Dum	0.0196***		0.322***		
	(0.000697)		(0.117)		
Log Imp		0.0134***		0.211***	
		(0.000448)		(0.0167)	
Ihs. Imp					0.0218***
					(0.000546)
N	499000	57000	4100	3400	499000
R-squared	0.392	0.478	0.569	0.595	0.419
Firm FE	Yes	Yes	Yes	Yes	Yes
Country-Ind-FE	Yes	Yes	Yes	Yes	Yes

Panel B: Imp on R&D Offshoring.

	(1) Imp Dum	(2) Imp Dum	(3) Log Imp	(4) Log Imp	(5) Ihs. Imp
R&D Dum	0.210***		1.755***		
	(0.00675)		(0.0529)		
Log R&D		0.00711***		0.309***	
		(0.00261)		(0.0254)	
Ihs. R&D					0.578***
					(0.0118)
N	499000	4100	57000	3400	499000
R-squared	0.420	0.612	0.475	0.661	0.470
Firm FE	Yes	Yes	Yes	Yes	Yes
Country-Ind-FE	Yes	Yes	Yes	Yes	Yes

Notes: Table presents coefficient estimates for the regressions specified in Appendix Section C for year 2017. In Panel A, the dependent variable y_{il} is R&D and the independent variable x_{il} is import. In Panel B, the opposite is true. For each variable, the extensive margin (captured by the dummy variable), the intensive margin (captured by the logged variable), and the combination of both margins (captured by the inverse hyperbolic transformation) are considered. Industries are identified by 3-digit NAICS codes. * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$. Standard errors are clustered at the firm level and reported in the parentheses.

Table A4: Colocation of Offshored Production and Innovation

Panel A: R&D Offshoring on Imp.

	(1) R&D Dum	(2) R&D Dum	(3) Log R&D	(4) Log R&D	(5) Ihs. R&D
Imp Dum	0.0181***		0.391***		
	(0.000473)		(0.0603)		
Log Imp		0.0125***		0.210***	
		(0.000288)		(0.0102)	
Ihs. Imp					0.0214***
					(0.000417)
N	3387000	536000	39000	33500	3387000
R-squared	0.389	0.472	0.568	0.594	0.414
Firm-Year FE	Yes	Yes	Yes	Yes	Yes
Country-Year-Ind-FE	Yes	Yes	Yes	Yes	Yes

Panel B: Imp on R&D Offshoring.

	(1) Imp Dum	(2) Imp Dum	(3) Log Imp	(4) Log Imp	(5) Ihs. Imp
R&D Dum	0.172***		1.637***		
	(0.00405)		(0.0329)		
Log R&D		0.00727***		0.294***	
		(0.00114)		(0.0141)	
Ihs. R&D					0.504***
					(0.00771)
N	3387000	39000	536000	33500	3387000
R-squared	0.448	0.632	0.467	0.663	0.501
Firm-Year FE	Yes	Yes	Yes	Yes	Yes
Country-Year-Ind-FE	Yes	Yes	Yes	Yes	Yes

Notes: Table presents coefficient estimates for regressions specified in Appendix Section C. In Panel A, the dependent variable y_{ilt} is R&D and the independent variable x_{ilt} is import. In Panel B, the opposite is true. For each variable, the extensive margin (captured by the dummy variable), the intensive margin (captured by the logged variable), and the combination of both margins (captured by the inverse hyperbolic transformation) are considered. Industries are identified by 3-digit NAICS codes. * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$. Standard errors are clustered at the firm level and reported in the parentheses.

Table A5: Cross-Country Interdependence of R&D and Production Offshoring

Panel A: R&D Offshoring on Imp.	(1)	(2)	(3)	(4)	(5)
	R&D Dum	R&D Dum	Log R&D	Log R&D	Ihs. R&D
Imp Dum	0.0180***		0.391***		
	(0.000844)		(0.0625)		
Region Imp Dum	0.00169***		-0.0482		
	(0.000261)		(0.0682)		
Log Imp		0.0134***		0.208***	
		(0.000539)		(0.0134)	
Log Region Imp		0.00100***		0.00734	
		(0.000376)		(0.0147)	
Ihs. Imp					0.0213***
					(0.000862)
Ihs. Region Imp					0.000927***
					(0.000200)
N	3387000	400000	39000	30000	3387000
R-squared	0.389	0.483	0.568	0.593	0.414
Firm-Year FE	Yes	Yes	Yes	Yes	Yes
Country-Year-Ind-FE	Yes	Yes	Yes	Yes	Yes

Panel B: Imp on R&D Offshoring.	(1)	(2)	(3)	(4)	(5)
	Imp Dum	Imp Dum	Log Imp	Log Imp	Ihs. Imp
R&D Dum	0.171***		1.639***		
	(0.00652)		(0.0371)		
Region R&D Dum	0.0442***		0.171***		
	(0.00367)		(0.0298)		
Log R&D		0.00838***		0.296***	
		(0.00172)		(0.0190)	
Log Region R&D		0.00228		0.0431*	
		(0.00166)		(0.0229)	
Ihs. R&D					0.502***
					(0.0121)
Ihs. Region R&D					0.0968***
					(0.00655)
N	3387000	25500	536000	22000	3387000
R-squared	0.449	0.637	0.467	0.689	0.501
Firm-Year FE	Yes	Yes	Yes	Yes	Yes
Country-Year-Ind-FE	Yes	Yes	Yes	Yes	Yes

Notes: Table presents coefficient estimates for regressions specified in Appendix Section C. In Panel A, the dependent variable y_{ilt} is R&D in country l, the independent variable x_{ilt} is import from country l, and the other independent variable x_{irt} is import from the same region as country l but excluding itself. In Panel B, the independent and dependent variables are switched. For each variable, the extensive margin (captured by the dummy variable), the intensive margin (captured by the logged variable), and the combination of both margins (captured by the inverse hyperbolic transformation) are considered. Industries are identified by 3-digit NAICS codes. * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$. Standard errors are clustered at the firm level and reported in the parentheses.

Table A6: Effect of Trump Tariffs - Neighboring Countries

	(1) Log Imp	(2) Ihs. R&D	(3) R&D Dum	(4) Log R&D
Treat × Post	-0.0976***	-0.0852**	-0.0113**	-0.127*
	(0.0254)	(0.0428)	(0.00557)	(0.0652)
Treat Neighbor × Post	-0.0293	-0.0340	-0.00657	-0.0833
	(0.0263)	(0.0408)	(0.00536)	(0.0613)
N	185000	185000	185000	16500
R-squared	0.890	0.877	0.838	0.894
Firm-Country-FE	Yes	Yes	Yes	Yes
Country-Year-FE	Yes	Yes	Yes	Yes

Notes: Table presents coefficient estimates of the DID regressions as specified in Equation (4). See Section III.B for details on the quasi-experiment, i.e. Trump Tariffs. The treatment dummy equals one for firm-country pairs whose tariff rate was affected during Trump Tariffs. The "Treat Neighbor" dummy equals one if the firm's tariff rate in neighboring countries (excluding the focal country) was affected during Trump Tariffs. The Post dummy equals one for year 2019 and zero for years 2014-2017. The four columns correspond to four outcome variables: log import, inverse hyperbolic transformation of R&D expenditure, indicator for positive R&D expenditure, and log R&d expenditure. * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$. Standard errors are clustered at the firm level and reported in the parentheses.

Table A7: Transition Probabilities of Production and Innovation Offshoring Decisions

Panel A: Transition of Import Status.

	$100 * \Pr(\text{Imp}_{t+1} = 0)$	$100 * \Pr(\text{Imp}_{t+1} = 1)$
$\text{Imp}_t = 0$	93.98	6.02
$\text{Imp}_t = 1$	17.69	82.31

Panel B: Transition of R&D Status.

	$100 * \Pr(\text{R\&D}_{t+1} = 0)$	$100 * \Pr(\text{R\&D}_{t+1} = 1)$
$\text{R\&D}_t = 0$	99.55	0.45
$\text{R\&D}_t = 1$	12.92	87.08

Notes: Table presents the transition probabilities of the discrete R&D and import states. Calculation is based on a firm-country and year panel. Rows represent states in the current year and columns represent states in the next year. Data on firms' import status comes from the LFTTD, and data on their R&D status is from the BRDIS.

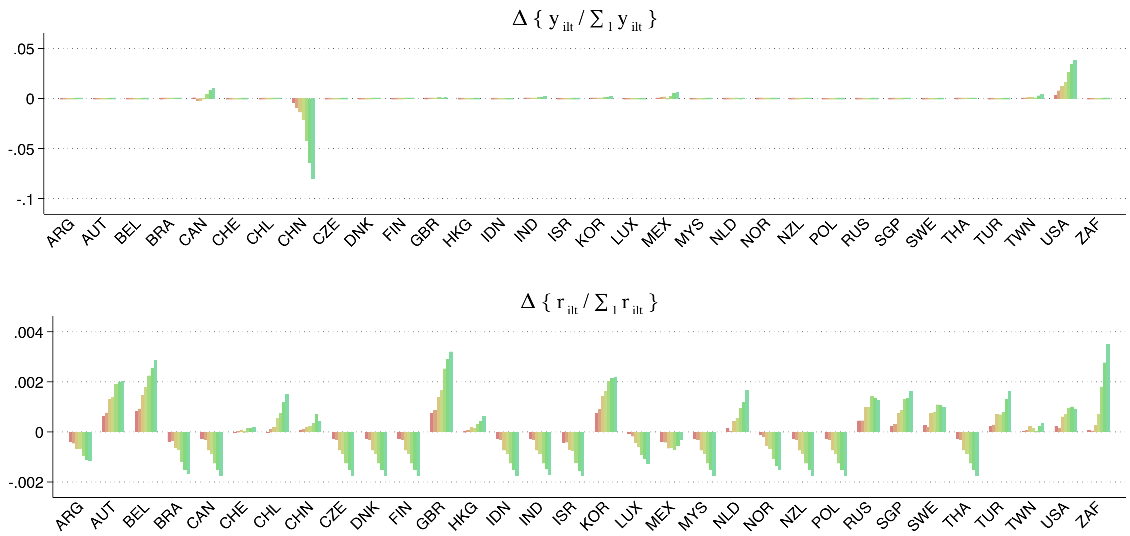
Table A8: Validation of Production Offshoring Potential Estimates

	(1) Log # Importers	(2) Log # Importers	(3) Log # Importers
$\ln \hat{\theta}_{lt}$	0.776***	0.764***	0.362***
	(0.0233)	(0.0216)	(0.0751)
N	450	450	450
FE	No FEs	Year FEs	Country FEs

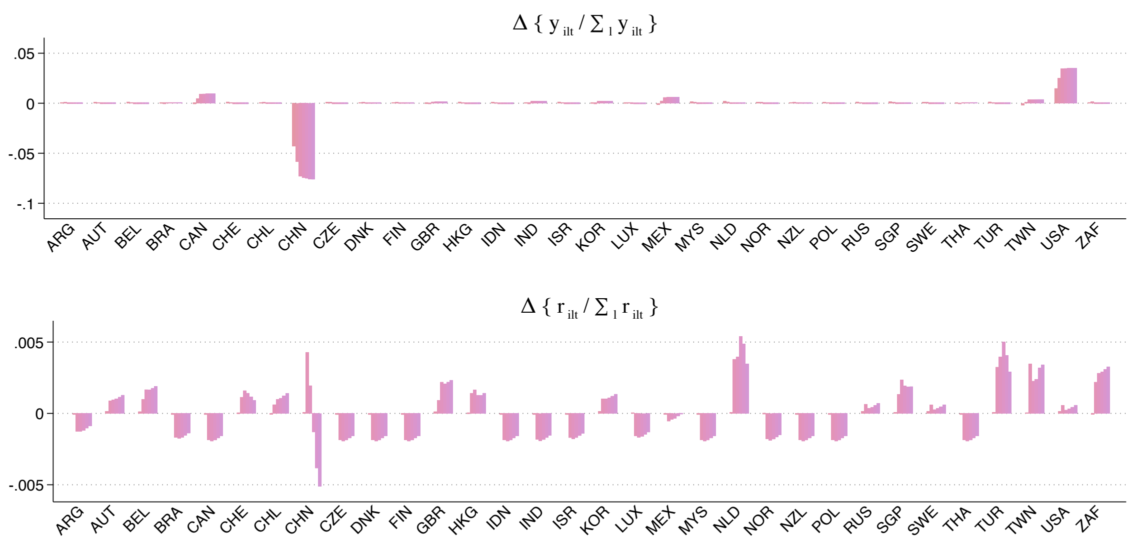
Notes: Table presents coefficient estimates from regressing the log number of importing firms on the country's log estimated production offshoring potential, $\ln \hat{\theta}_{lt}$. Regression is based on a country-year panel. The independent variable $\ln \hat{\theta}_{lt}$ (see Section V for more details on its definition) is obtained from running an OLS regression of Equation (8). * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$. Standard errors are reported in the parentheses.

Figure A1: Non-Linear Effects of U.S. Trade Policies Against China

Panel A: Decreasing production offshoring potentials

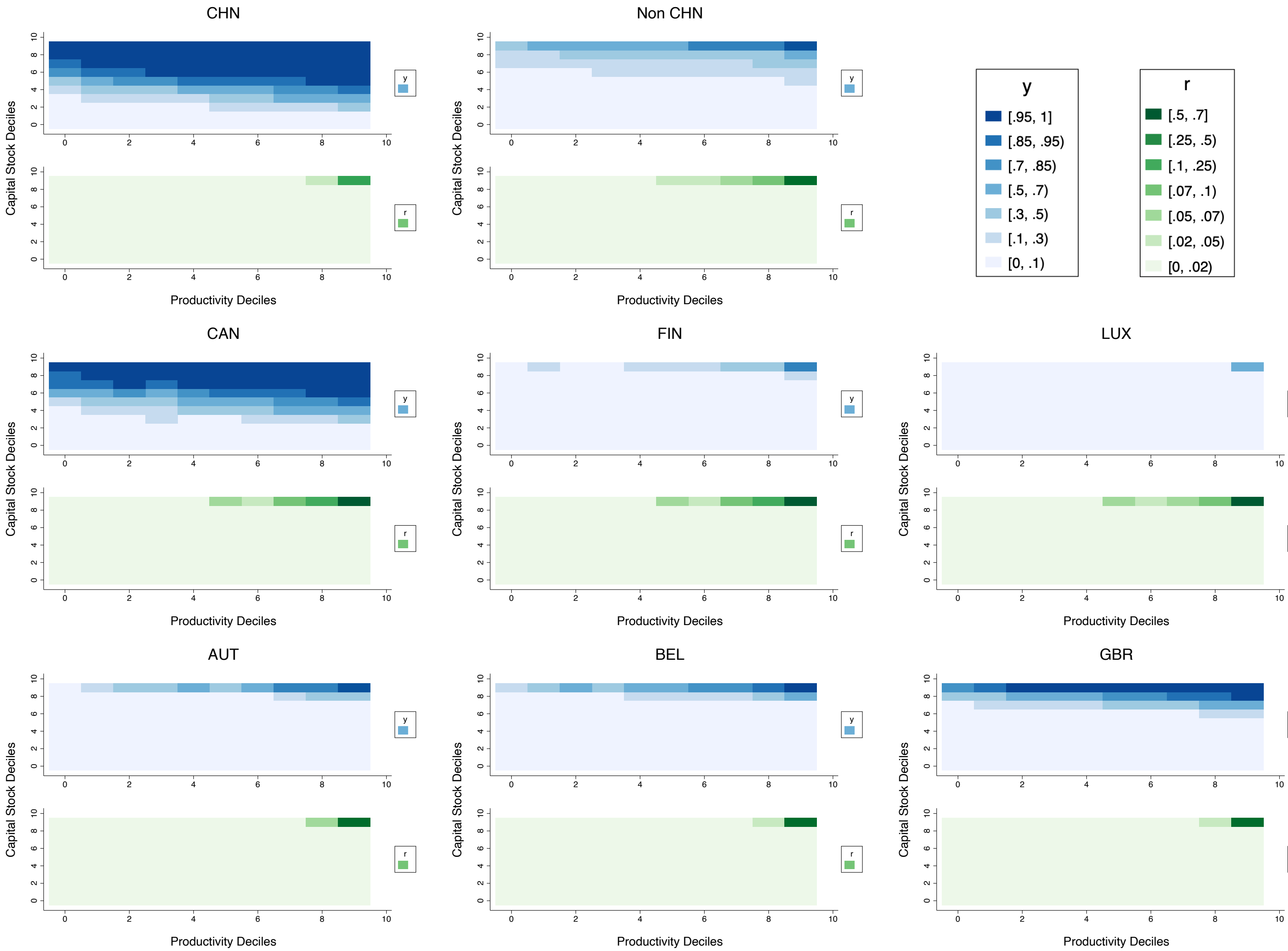


Panel B: Increasing Cost of Production



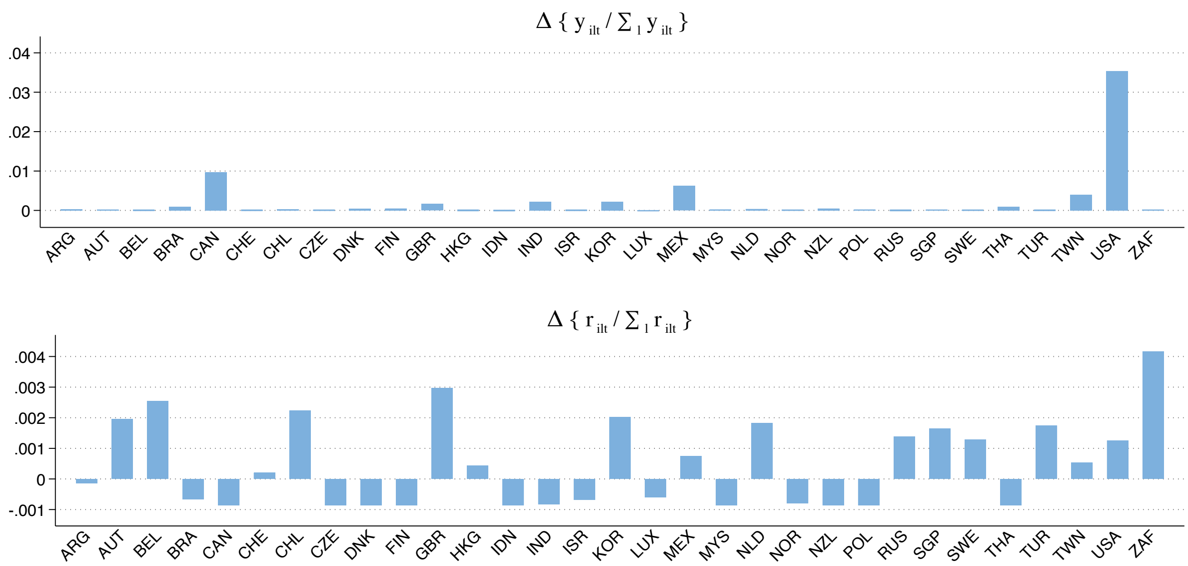
Notes: Figure presents the counterfactual results depicting the changes in countries' production and innovation shares in response U.S. trade policy shocks that adversely affect production offshoring to China. In Panel A, each country is represented by nine bars, corresponding to different reductions in China's production offshoring potential (20%, 25%, 40%, 50%, 75%, 90%, and 100%). In Panel B, each country is represented by seven bars, corresponding to different increases in the sunk and fixed cost of production in China ($2K, $5K , $30K , $50K , $65K , $85K , and $100K). See Section VII.B.2 for further details on these two policy counterfactual exercises.

Figure A2: Firm Heterogeneity in Production and Innovation Offshoring



Notes: Figure depicts the fraction of firms that engage in production and innovation offshoring to each country in 2017, simulated using the baseline model and categorized based on firms' productivity (on the x-axis) and capital stock (on the y-axis) deciles. The blue panels represent production offshoring, while the green panels represent innovation offshoring. The figure includes panels for the following countries: China, all countries except China, Canada, Finland, Luxembourg, Austria, Belgium, and the UK.

Figure A3: Simulated Effect of U.S-China Decoupling



Notes: Figure presents the counterfactual results depicting changes in countries' production and innovation shares in response to the decoupling of the U.S. and China. This counterfactual exercise is implemented by setting the costs of producing and innovating in China to infinity.