Talent, Geography, and Offshore R&D*

Jingting Fan†

This draft: March 2021

Abstract

I model and quantify the impact of a new dimension of globalization: offshore R&D. In the model, firms work with researchers to develop new product blueprints and then engage in offshore production and exporting. Cross-country differences in the distributions of firm knowhow and worker ability generate a 'talent-acquisition' motive for offshore R&D, while frictions impeding trade and separation of production from R&D lead to a 'market-access' motive. I discipline the model using empirical facts documented from a new firm-level dataset. Counterfactual experiments show that the two motives can account for a significant part of the observed offshore R&D. Incorporating offshore R&D amplifies the gains from globalization by a factor of 1.3 and generates new implications on the impacts of traditional forms of global integration, namely trade and offshore production.

Keywords: Multinational firms, offshore R&D, global value chain, gains from openness
JEL Classification: F21 F23 F40 O32

*This paper is based on a chapter of my dissertation. I am deeply grateful to Nuno Limão for his encouragement and guidance and to my committee Şebnem Kalemli-Özcan and John Shea for their invaluable feedback. I thank Costas Arkolakis, Ariel Burstein, Kamran Bilir, Jonathan Eaton, Gordon Hanson, Wolfgang Keller, Eunhee Lee, Wenlan Luo, James Markusen, Luca Opromolla, Natalia Ramondo, Andres Rodriguez-Clare, Felipe Saffie, Lixin Tang, Jim Tybout, Daniel Wilson, Stephen Yeaple, and participants at seminars and conferences for helpful comments. All errors are mine.

†Department of Economics, Penn State University. Email address: jxf524@psu.edu

1 Introduction

Consider the creation and commercialization of a new product. Engineers develop prototypes. Firms, with their knowhow—insights into consumer behavior, experience with production processes, and brand recognition—oversee prototype development and carry out marketing and manufacturing production. Both the talent of engineers and the knowhow of firms are vital for this process, but globally, these two factors are distributed unevenly. While emerging countries such as China, India, and Eastern Europe have some of the biggest pools of engineers and researchers (National Science Board, 2018), the vast majority of the world's best-run firms and most-recognized brands are from industrialized countries.[1]

Firms can go global—by carrying out R&D overseas—to overcome this spatial mismatch and scale up the return to their knowhow. Given that firms engaging in multinational activities are often the largest ones accounting for a substantial share of global commerce,[2] it is perhaps unsurprising that R&D at the overseas affiliates of these firms is quantitatively significant. As Figure 1 shows, in many host countries, R&D carried out by foreign firms has increased over the past decades and now amounts to a substantial share of domestic R&D.

The sizable offshore R&D has implications beyond the bottom line of individual firms. By determining the location and efficiency of R&D, it directly affects product prices and availability. In a global economy where inventions in one place are being produced in, and exported to, multiple countries, it further influences the impacts of trade and offshore production policies. Yet despite its potential importance, the prior quantitative studies on globalization, reviewed below, have focused on trade and multinational production, and overlooked offshore R&D. The goal of this paper is to model offshore R&D and quantify its global impacts—both its direct effect on the welfare of countries and its interaction with trade and offshore production.

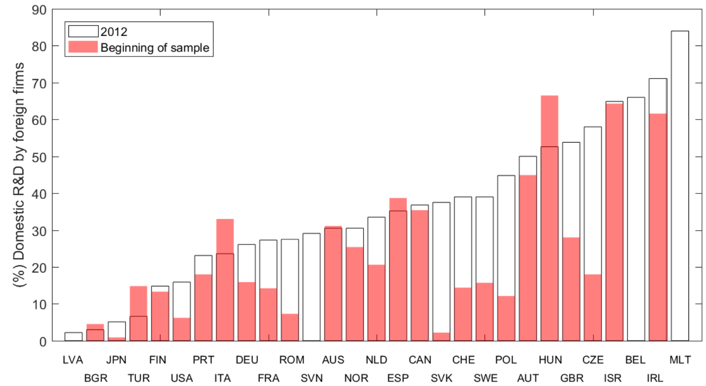
A hurdle to this goal is, being an integral part of a firm's global expansion strategy, offshore R&D is determined jointly with production. A model of offshore R&D is thus necessarily a model of multinational production. As production depends in part on the market access of countries, the model should embrace that geography is as important a determinant of offshore R&D as the endowment of talent and knowhow. To build and discipline such a model for global-scale analysis in turn calls for data on R&D and production of multinational corporations (MNCs) from a broad set of geographic regions. However, although new firm-level datasets have made it possible to measure production of MNCs from many countries (see, e.g., Alviarez, 2019; Cravino and Levchenko, 2017), systematic data on R&D are not readily available.

The first contribution of this paper is to assemble such a dataset and study empirically the joint

[1]Ninety out of the world's hundred most valuable brands are from the G7 countries (Swant, 2019), a group accounting for only a third of world GDP. Data on management practices show well-managed firms are also concentrated in developed countries (Bloom et al., 2012). One possible explanation for the abundance of knowhow in the early industrialized countries is that knowhow is acquired through learning by doing. Once acquired, however, it becomes valuable assets for firms. A quote from Lewis Platt, a former HP executive, underscores both the value of knowhow and its slow accumulation: 'If HP knew what HP knows, we would be three times as profitable.'

[2]Multinational firms account for about a third of global production and half of global export (OECD, 2018). Their role is even more prominent in R&D. In the U.S., for example, about 90% of business R&D are carried out by either the affiliate of foreign firms or headquarters of American MNCs (National Science Board, 2018).

Figure 1: The Level and Growth of Offshore R&D, 1985-2012



Notes: The measure for country i is $\frac{\text{Business enterprise R\&D expenditures in country } i \text{ by foreign firms}}{\text{Total business enterprise R\&D expenditures in country } i}$. Uncolored bars are for 2012; colored bars are for the beginning of the sample, which differs by country and dates back to as early as 1985. Data source: the OECD.

production and R&D decisions of MNCs. Production data and the ownership network connecting affiliates to their parents are from Orbis. I construct R&D data from PATSTAT Global, a database of administrative patent records from over 90 patent offices. The *address of inventors* provided in patent applications identifies where the invention underlying each patent is made. I match owners of patents to firms in Orbis and then measure the invention of an MNC in a host using the number of its patents invented there. Further refined using the information on firm ownership network and patent families, my measure is robust to where an MNC patents an invention (e.g., the USPTO, the EPO, or both) and which of its affiliates is listed as the patent owner. It is strongly correlated with bilateral offshore R&D measured using expenditures but has the advantage of being widely available as a firm-level panel for many countries.

I document three facts. First, the invention intensity of an affiliate, measured as the ratio between the number of patented inventions and sales, is higher in host countries with better human capital, and increases as host human capital improves over time. This is consistent with a 'talent-acquisition' motive arising from the relative abundance of a host in talented workers. Second, invention and production within a firm tend to colocate in both cross section and over time. Such colocation hints at frictions impeding the separation of R&D from production. Coupled with trade costs, these frictions incentivize firms to conduct R&D in countries where goods can be produced and shipped to major destinations cheaply—a 'market-access' motive. Third, invention and production of overseas affiliates both decrease in distance to the headquarters. This highlights the second role of geography: limiting firms' ability to mobilize their knowhow to overseas affiliates.

These facts should be respected by any quantitative models of multinational R&D and production. The second contribution of this paper is to develop and quantify the first such model to interpret the data and conduct experiments. I find that, first, offshore R&D generates significant but unevenly distributed gains. It benefits especially developed countries, for which it is a complement to trade and offshore production; for developing countries, it is a substitute. Second, because of the interaction between R&D, production, and trade—both within firms and through the general equilibrium—incorporating offshore R&D is important for assessing the impacts of trade and offshore production policies.

In my model, firms differ along two aspects of knowhow: innovation management efficiency—which governs their efficiency in working with researchers in R&D—and production management efficiency, which governs their manufacturing productivity. Firms can enter foreign hosts to perform R&D, a process that converts inputs of local researchers into new differentiated varieties. Workers differ in ability and choose to be a researcher or a manufacturing worker. I embed this offshore R&D decision into a multi-country general equilibrium model of global production and trade (Arkolakis, Ramondo, Rodríguez-Clare and Yeaple, 2018). After a product is developed by an R&D center, whether onshore or offshore, the firm chooses which countries to sell it to and for each destination, where production should take place. Thus, in an organization typical of modern MNCs, an American company can develop a new product in Germany, produce it in China, and then export to India. Motivated by the empirical patterns, I incorporate geographic frictions in separating production from R&D, and in separating both from the headquarters.

I calibrate the model to data from 37 major countries. Two crucial aspects of the model are the distributions of worker ability and firm knowhow, and geographic frictions. I parameterize each country's distribution of firm knowhow using the World Management Survey built by Bloom, Genakos, Sadun and Van Reenen (2012) and its talent distribution using a cognitive test score database developed by Hanushek and Woessmann (2012). This pins down the 'talent-acquisition' motive. To discipline the 'market-access' motive, I disentangle various types of geographic frictions using an indirect inference strategy. Specifically, I parameterize the cost of doing offshore R&D as a function of the distance between a host and the headquarters. I further capture parsimoniously the colocation and headquarter effects for production by specifying the cost of producing a variety in a host to be a weighted function of that host's proximity to the headquarters and its proximity to where the variety is invented. The weight and distance elasticities are then pinned down by matching the coefficients from the reduced-form regressions.

The model matches the data closely along dimensions not directly targeted. In further support of the model, I show through counterfactual experiments that the measured endowment distributions and geography are both first order determinants of offshore R&D. For example, improving the talent distribution of Brazil (worst among sample countries) to the world average increases the foreign share of R&D in Brazil by half; decreasing the management efficiency of the U.S. to the world average increases the foreign share of R&D in the U.S. by two thirds. Market access plays large but heterogeneous roles: for a typical country, being shut off the access to foreign consumers through exporting and to foreign producers through offshore production simultaneously results in a decline in inward offshore R&D by around half.

I use the model for two purposes. First, I use it as a device to measure how offshore R&D is connected to firms' global production and its contribution to national income. My calibration interprets the reduced-form patterns as that proximity to R&D centers is more important for production than proximity to the headquarters. As a result, the model infers that on average 70% of R&D in overseas affiliates is for local production, with the remaining conducted for production in surrounding countries and at the headquarters. This share is higher in hosts with a low production cost, consistent with the market-access motive. The model also highlights offshore R&D as a significant source of profit for countries with the best knowhow. For example, for the U.S., the profit accruing to products developed in its overseas R&D centers accounts for about a third of the total profit, or 7% of its national income. In contrast, for developing countries with poor knowhow endowments, both shares are close to zero.

Not measurable in typical datasets, these inferred patterns have broad policy implications. For example, governments around the world have granted generous R&D credits to attract foreign firms. The close connection between R&D and affiliate production means that in addition to R&D, these firms will likely also bring in their production, especially when the host has a low manufacturing cost; conversely, the growth in the relocation of manufacturing to developing countries might have an influence on foreign R&D in these countries. The significance of profit from overseas R&D is relevant for valuing the intangible wealth of nations.

The second use of the model is to examine the normative implications of offshore R&D. Analytically and quantitatively, I show that the option to conduct R&D abroad generates sizable gains. The average welfare gains from this channel, defined analogously to the gains from trade, are around 3.3%. Compared to a restricted model with only trade and offshore production, incorporating offshore R&D amplifies the gains from openness by a factor of 1.3. This amplification is substantially larger for developed countries, primarily because a higher fraction of their income is generated through offshore R&D. Existing quantitative work on globalization overlooking this channel thus not only underestimates the overall gains from openness, but also biases the comparison of the gains across countries.

The weights in the specification for offshore production cost of proximity to headquarters and proximity to R&D centers are critical for the inferred gains. My estimate suggests that proximity to R&D center weights more; if I had assumed that only proximity to the headquarters matters, I would have found much larger gains from R&D and openness, especially for emerging countries. This shows the value of disciplining the model using micro data.

I conclude the experiments by highlighting two important interactions between offshore R&D and two traditional forms of globalization, trade and offshore production. The first interaction is through a general equilibrium effect. In the model, trade and offshore production together enable countries to specialize in innovation or production according to their comparative advantage. By increasing effective R&D capacity everywhere, offshore R&D strengthens the comparative advantage of countries specializing in innovation and weakens the comparative advantage of those specializing in production. For this reason, it is a complement to trade and offshore production for developed countries and a substitute for developing countries.

The second interaction is through the within-firm linkage between R&D and production. Because for many firms, the varieties they develop abroad constitute an important source of profit, policies that change the profitability of these varieties can have a large impact on the firm. For this reason, liberalization in offshore production between two countries can ultimately benefits a third country that owns R&D centers there. Similarly, R&D subsidies in a host can have a global influence because foreign owners of R&D centers there also benefit. I show that such indirect effects are large for countries with a strong presence in overseas R&D, such as the U.S. These results are relevant for evaluating multilateral investment agreements, especially since many of them cover R&D-related aspects, such as requirements on protection of foreign intellectual property rights.

This paper contributes to a growing stand of literature on the impacts of MNCs (c.f., McGrattan and Prescott, 2009; Burstein and Monge-Naranjo, 2009; Garetto, 2013; Ramondo and Rodríguez-Clare, 2013; Arkolakis, Ramondo, Rodríguez-Clare and Yeaple, 2018; Tintelnot, 2016; Alviarez, 2019). Within this literature, the most closely related papers are Arkolakis, Ramondo, Rodríguez-Clare and Yeaple (2018), which studies how trade and offshore production affect country specialization and welfare, and Tintelnot (2016), which estimates a model of multinational production using firm-level data for general equilibrium counterfactuals. Different from these studies, the present paper focuses on offshore R&D, a prevalent but underexplored activity. I show that offshore R&D is important for both the gains from openness and the effects of trade and offshore production policies. Also related is Bilir and Morales (2020), which examines the impact of R&D within American MNCs using a production function estimation approach. The present paper instead models firms' R&D and production choices explicitly in a general equilibrium setting, thereby speaking to the aggregate implications of offshore R&D.

This paper also contributes to the literature explaining the pattern of FDI, dating at least as far back to as the theoretical work by Helpman (1984) and Markusen (1984). Recent studies in this area have examined the determinants of M&A FDI (Nocke and Yeaple, 2008), the role of firm heterogeneity (Helpman, Melitz and Yeaple, 2004), the dynamics of MNCs (Garetto, Oldenski and Ramondo, 2019; Fillat and Garetto, 2015), and have looked into the role of MNCs in shaping consumer preference and creating market access (Head and Mayer, 2019; Wang, 2019). The contribution of this paper is two folds. Empirically, I document new facts on the joint allocation of R&D and production across a broad set of countries, complementing existing studies, most of which focus on either production or R&D alone, often using data from a single home or host country.[3] Quantitatively, I develop a model that captures complex structures seen in modern MNCs. The model allows me to separate the roles of talent and various geographic frictions in the global operation of MNCs.

Finally, this paper is related to the research on the measurement of the global value chain (c.f., Johnson and Noguera, 2012; Koopman, Wang and Wei, 2014; Antràs and De Gortari, 2020). Most existing measurements are based on sectoral production. Increasingly, countries in the global value chain (GVC) also specialize in tasks (innovation, production, etc.). One obstacle to systematic measurements of country specialization by task is data availability. Compared to merchandise trade and sectoral production, data on cross-border flows of headquarter services and R&D are far less accessible. By combining a new firm-level dataset and a structural model, I characterize empirically a GVC with four stages: headquarter services embedded in the provision of knowhow, R&D, manufacturing, and marketing. This contributes to the measurement of R&D in the GVC.

[3] For example, Irarrazabal, Moxnes and Opromolla (2013) and Keller and Yeaple (2013) document gravity for affiliate production. On R&D, Hall (2011) documents important cross-border R&D by MNCs using aggregate patent data; Siedschlag, Smith, Turcu and Zhang (2013) estimates R&D location choice for MNCs from the EU. Because they define R&D based on the industry of an affiliate, their sample includes only 446 R&D location decisions.

2 Data and Facts

2.1 Data Sources and Empirical Sample

I assemble a dataset on the invention and production activities of firms from different countries, linked together by an ownership network. This subsection outlines the main sources of data and preparation procedures; Online Appendix A provides additional details on these procedures and supplementary data, and results from several validation and robustness exercises.

Financial and ownership data. The financial and ownership data are from the Historic Disk of the Orbis Database extracted in April 2017. I use the 2016 vintage of shareholder data to identify the parent of each firm, defined as the entity holding more than 50% control over the firm either directly or indirectly. These ownership data span across country borders, so I can link firms to their parents, domestic or overseas. Firms not linked to a parent are assumed to be independent.

My primary measure of production is sales; for robustness I will also use value added, which has more limited availability. The financial data come at the year-firm identifier level. I group all firms in a country owned by the same parent into one and treat it as an affiliate. This step gives me a total of around 185 million parent firm-host country-year observations between 1996 and 2016. The vast majority of parent firms have only one host—their home countries. Appendix Table A.1 reports the coverage of the financial data for the set of countries included in this study.

Patent data. I measure the invention activity of MNCs using the *addresses of patent inventors*, so, if a firm is granted many patents invented by researchers in Japan, for example, I infer that the firm invest heavily in R&D in Japan (but see the caveats discussed below).

My primary data source is PATSTAT Global, which covers patent applications across 90 national and regional patent offices. I match individual patents to firms and then identify where each patent is invented based on the location of its inventor(s). Around 25 million granted patents are matched to 681,241 unique parent firms in the Orbis Database. Only less than a quarter of these patents are from the USPTO (see Appendix Table A.2 for sample composition and the availability of inventor location information by patent office). Thus, by using a comprehensive patent database, I expand coverage among firms patenting outside the U.S., which will provide valuable variation when I relate patent invention to country characteristics. Relative to measures of MNC R&D based on R&D spending data, which are rarely available for more than one host or one home country at a time, the advantage of my measure is that detailed information on the universe of patents in most countries is publicly accessible.

A general caveat in interpreting patent invention as R&D is that being a decision of firms', patenting depends on the characteristics of both the firm and the country. In regression analysis below, I will purge out these factors by including a rich set of controls and fixed effects. In quantitative analysis in later sections, I will further need as an input the share of inventions in a host made by foreign firms. I show in Appendix Figure A.2 that aggregate and bilateral foreign invention shares calculated using patent counts are strongly correlated with those calculated using R&D expenditures, so using patent data does not introduce systematic biases for aggregate shares.

There are two additional caveats on measurement that are specific to the multinational setting of this paper.[4] First, MNCs are in principle free to assign the ownership of patents to any its affiliates, so the location of the assignee (the owner of a patent) needs not be where the invention was conducted. My use of the location of inventors circumvents this problem.[5] Second, because a patent only grants protection within the country in which it is issued, firms often seek patents in different countries (e.g., the USPTO and the JPO) for the same underlying invention. To avoid double counting, I keep only one patent among each family of patents covering the same invention.[6] These two choices also imply that my measure of invention is independent of where an MNC files for a patent, and which of its affiliates is assigned as the owner of the patent. This further alleviates the concern that my measure picks up firms' self-selection into filing patents in particular hosts.

I aggregate individual patents to obtain the total number of patents at the level of parent firm, year, and host country, with host countries identified by the location of inventors. In doing so, it is necessary to take a stand on how to classify patents with inventors from different countries. In the baseline analysis, I split patents based on the number of inventors. For example, if an invention is made jointly by two inventors in India and one in the U.S. headquarters, I attribute two-thirds of it to India and one-third to the U.S. For robustness, I will also use an alternative measure that counts each of the hosts co-inventing a patent as having invented the entire patent.

Sample country and time. I supplement firm-level datasets with country characteristics. To have a consistent sample as the quantitative section, I restrict to 37 host countries, selected based on the coverage of the World Input-Output Database and the Penn World Table, both of which are crucial for calibration. To reduce measurement errors from the discrete patent outcomes and to smooth out yearly fluctuations in country characteristics, I average the data by five-year intervals between 1996 and 2016.

2.2 Descriptive Statistics

I combine the financial and patent datasets. The majority of firms never patented; some firms with patents have no available financial information. For regression analysis, I focus on firms that both have some financial information and have filed at least one patent—invented anywhere—during the sample period; the rest of this subsection provides descriptive statistics on this combined sample. For tabulating country-level statistics in quantification later, I will use the full set of firms

[4] There are also two advantages that are specific to the multinational setting in using patents to measure R&D. First, MNCs have an incentive to manipulate intangibles to shift profits across tax jurisdictions (Guvenen et al., 2017). Since the addresses of inventors appearing on a patent application does not affect the taxes firms own, they are less likely to be manipulated. Second, patents from different countries are conceptually closer to each other than R&D expenditures. R&D in a country trying to catch up with the frontier might be more about learning what others already know than about pushing the frontier. Patents in that country, on the other hand, requires pushing the frontier, even if only by a tiny step. In this sense, they are more comparable. To the extent that patents differ in quality, it is relatively straightforward to adjust for it using the number of forward citations.

[5] The location of an inventor is based on his/her mail address on patent application and is *not* his/her nationality. One potential concern is that if, after spending years on a project, an inventor moves across countries right before applying for a patent, then treating the inventor address as the location of R&D introduces measurement errors. Appendix A.2 performs a back-of-envelop calculation to show that errors introduced this way, if any, are likely small.

[6] I identify patents from different offices covering the same invention using their common priority, established when the first patent on an invention is filed anywhere.

Table 1: Sample Structure: Firms, Production Facilities, and R&D Centers

	(1)	(2)	(3)	(4)
			R&D center count	
Period	Firm count	Prod. facility count	Baseline	Liberal
1996-2000	26,635	43,395	30,506	36,818
2001-2005	48,473	76,641	52,243	61,539
2006-2010	75,336	113,401	78,924	91,703
2011-2016	86,335	131,493	88,569	102,886
Total	236,779	364,930	250,242	292,946

Note: Reported values are the numbers of distinct firms (Column 1), distinct production facilities (Column 2), and distinct R&D centers (Columns 3 and 4) in the matched R&D-financial sample. Each row is for a five-year interval between 1996 to 2016.

Table 2: Descriptive Statistics of the Firm-Level Sample

Panel A: Manufacturing and non-manufacturing shares					
	firm count	% of total revenue		% of total patents	
			% offshore		% offshore
Mfg.	43220	39.68	38.21	56.32	15.23
Non-Mfg.	43115	60.32	24.60	43.68	9.86
All	86335	100.00	30.00	100.00	12.89

Panel B: The distribution of firms by the number of affiliates						
	by the number of prod. sites			by the number of R&D Centers		
	All	Mfg.	Non-Mfg.	All	Mfg.	Non-Mfg.
1	77447	39,027	38,420	82,498	41,203	41,295
2	3361	1,602	1,759	2,373	1,158	1,215
3	1317	567	750	554	273	281
4	770	337	433	270	148	122
5	513	219	294	180	120	60
>=6	2927	1,468	1,459	460	318	142
Total	86335	43220	43115	86335	43220	43115

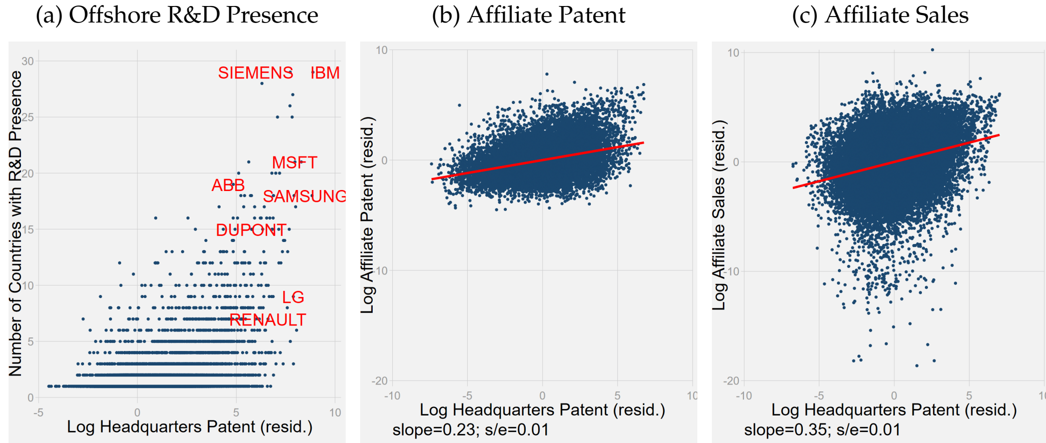
Note: This table reports summary statistics for the last period (2011-2016). Panel A reports the sample composition by manufacturing and non-manufacturing firms. Panel B reports the number of firms with different numbers of production affiliates and R&D centers.

to ensure good representation of the economy; Appendix A.3 provides a summary of that sample.

Sample Coverage and Entry Patterns. Table 1 gives an account of sample coverage. The first column is the number of parent firms in each period. It increases over the periods, reflecting the broadening coverage of the database. The second column is the number of production affiliates. I define a host as a production affiliate of a firm if it has positive sales over the period. On average, each firm operates in 1.5 countries. The third column is the R&D center count. As a baseline, I define a host country as an R&D center of a firm if it invents at least one full patent there. Firms have on average 1.05 R&D centers. Column 4 uses a more liberal definition for R&D centers, which only requires a host to have invented a partial patent. Under this definition, the average R&D center per firm is 1.4.

Table 2 provides a few descriptive statistics, focusing on the last period. Panel A reports the composition of the sample by whether a parent firm is in manufacturing or not. Manufacturing

Figure 2: The Role of Firm Heterogeneity in Invention and Offshoring



Notes: Figures use data from the last period only. Panel (a) plots the number of countries a firm enters for offshore R&D against the normalized (by taking out home and 2-digit industry fixed effects) log number of patents at the headquarters. A few manufacturing conglomerates and technology companies are annotated. Panels (b) and (c) are residual plots of log affiliate number of patents and sales, respectively, against the log number of patents at the headquarters, controlling for host country, home country, and 2-digit industry fixed effects. Standard errors reported in Panels (b) and (c) are clustered by firm.

accounts for about half of the firm count, 40% of total revenue, and 56% of total patenting. About 38% of revenue and 15% of patents of these firms are generated in their overseas affiliates. Non-manufacturing firms account for a higher share of revenue but a lower share of patents; their overseas affiliates also carry out non-trivial shares of both sales and invention.

Panel B count firms by the number of countries in which they have a production/R&D presence. About 10% of all firms carry out production in more than one country, around 3% in more than 6. Compared to offshore production, offshore R&D is both less common and less spread out: only 5% of firms conduct invention in more than 1 country; less than 1% do so in more than 6 countries. Interestingly, these patterns do not differ systematically between manufacturing firms and others.

The Role of Firm Heterogeneity. Figure 2 shows participation in offshore R&D depends on firms' innovation at home. Panel (a) plots the number of R&D centers of a firm against their numbers of inventions at the headquarters. To rule out that firms from certain countries or industries patent more intensively than others, shown in the horizontal axis are the residuals, netting out home country and 2-digit industry fixed effects. The figure shows that firms with more inventions at home enter more countries for offshore R&D. I will interpret this as selection into offshore R&D by firms' innovation efficiency.

Panels (b) and (c) of Figure 2 are residual plots at firm-host country level. Panel (b) shows a positive correlation in the number of inventions between affiliates and headquarters, controlling for home country, host country, and industry fixed effects. The right panel shows that affiliate sales is positively correlated with headquarter invention, with an elasticity of 0.35, significantly above the 0.23 elasticity for affiliate invention depicted in Panel (b). The positive correlation in both panels is consistent with correlated innovation efficiency between parents and affiliates. A larger elasticity for affiliate sales than for affiliate invention hints at an additional source of hetero-

Table 3: Human Capital and Affiliate Invention Intensity

	(1)	(2)	(3)	(4)
Dependent variable:	ln (patent/sales)			R&D Indicator
human capital index	0.979***	3.013**	3.431**	0.181**
	(0.260)	(1.365)	(1.334)	(0.076)
ln(GDP per capita)	-0.308	-0.622	-0.710*	0.081***
	(0.250)	(0.439)	(0.350)	(0.023)
IPR protection		0.563***	0.404**	0.020
		(0.205)	(0.176)	(0.015)
R&D subsidies		0.508	0.572	0.011
		(0.384)	(0.403)	(0.029)
ln (researchers)			0.421**	0.067***
			(0.172)	(0.016)
ln(GDP)	0.057			
	(0.091)			
log (sales)				0.004***
				(0.001)
Observations	21031	11803	11464	80253
R^2	0.252	0.677	0.675	0.637
Within R^2	0.029	0.010	0.015	0.005
Distance measures	Y	-	-	-
Firm-period FE	Y	Y	Y	Y
Affiliate FE	-	Y	Y	Y

Note: The outcome variable is log of the ratio between patent counts and affiliate sales. The explanatory variables are country characteristics. Column 1 is a cross-sectional regression that controls for firm-period fixed effects and bilateral distance measures, including geographic distance and a set of dummies (see the text). Columns 2 through 4 control for time-invariant host characteristics through affiliate fixed effects. Standard errors (in parenthesis) are clustered by host country and by firm. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

geneity that determines affiliate sales beyond how many inventions an affiliate makes. The model developed in the next section will incorporate both ingredients.

2.3 Three Facts on Location Choices for Invention and Production

I document three facts on the spatial distribution of invention and production within a MNC.

Fact 1: The invention intensity of an affiliate increases in host human capital quality. The first fact investigates how affiliate invention intensity varies with host characteristics. I pay particular attention to host human capital, as it is rated as a primary factor by managers choosing locations for R&D centers (Thursby and Thursby, 2006). My specification is as follows:

$$y_{foht,t} = FE + \overrightarrow{\gamma_x} \cdot \overrightarrow{X}_{oh,t} + \epsilon_{ohf,t}.$$

The outcome variable is the log ratio of patent invention over sales, for an affiliate in host h of firm f from country o in period t. The explanatory variables $\overrightarrow{X}_{oh,t}$ include the measure of host talent, the human capital index from the Penn World Table, along with other controls. FE is fixed effects.

A concern in interpreting $y_{foht,t}$ as the invention intensity of an affiliate is that it might pick up that firms apply for more patents in more attractive markets. Note that in constructing y_{foh}, it is the number of patents *invented* in country h, which could be granted by any authorities, rather than the number of patents granted by the authority of country h, that is being counted. Selective patenting in more attractive hosts thus will not in itself bias the measure (see also Appendix Figure A.1 for additional evidence). I will include country fixed effects and time-varying country characteristics to soak up remaining variations in the propensity of patenting across hosts. Relatedly, some firms patent their inventions more frequently than others. This source of heterogeneity will be absorbed by firm-level fixed effects.

Table 3 reports the results. The first column controls for firm-period fixed effects and four measures of bilateral distance between the home and the host: geographic distance, and indicators for whether countries o and h share an official language, are contiguous to each other, or have a colonial tie. This specification exploits within-firm, cross-host, variation, and finds significant positive correlation between the invention intensity of an affiliate and the human capital index. The size and income of the host, on the other hand, do not seems to be important.

Columns 2 adds affiliate fixed effects, so all invariant country characteristics are absorbed. I further control for the protection of intellectual property rights (IPR, Park, 2008) and R&D subsidies (OECD) of the host, two policy measures likely correlated with R&D and patenting.[7] The specification shows that an improvement in host human capital over time is correlated with an increase in the patent sales ratio. In terms of magnitude, a one standard deviation increase in the human capital index (≈ 0.38) more than doubles the outcome variable. Column 3 further includes a narrower measure of talent, the number of researchers in a country. An increase in this measure is positively correlated with the invention intensity of affiliates, but the coefficient for the human capital index barely changes, suggesting that a broader interpretation of talent is warranted. Finally, Column 4 examine whether an improvement in human capital is associated with the entry of foreign R&D centers through the extensive margin, controlling for the sales of the affiliate. I find that both measures of talent are associated with entry.

Most existing quantitative work on MNCs either abstracts from R&D entirely or assumes it takes place only at the headquarters (Ramondo and Rodríguez-Clare, 2013; Tintelnot, 2016; Arkolakis et al., 2018). The first fact suggests that such simplification misses systematic variation in the R&D intensity of MNC affiliates around the world.

Fact 2: Invention and production tend to be located together. The second fact is on the colocation of invention and production. The specification is:

$$y_{fh,t} = FE + \gamma_{R\&D}\, x_{fh,t} + \overrightarrow{\gamma}_{dist} \cdot \overrightarrow{dist}_{fh,t} + \epsilon_{fh,t},$$

where f, h, t, indicates firm, host, and period, respectively. Variables $y_{fh,t}$ and $x_{fh,t}$ measure production and invention of firm f in host h. $\overrightarrow{dist}_{fh,t}$ is a vector consisting of four measures of the average distance between host h and all other countries in which firm f has an R&D center. $\overrightarrow{dist}_{fh,t}$ is firm specific because firms differ in their geographic presence. Coefficients $\gamma_{R\&D}$ and $\overrightarrow{\gamma}_{dist}$ capture the colocation patterns.[8]

[7] Because changes in GDP and income are highly correlated, I include only one of them.

[8] Since the quantitative analysis will be based on a static model, I focus on the contemporaneous effect in this specification. Including leads and lags of offshore R&D measures as additional explanatory variables result in positive and

Table 4: Colocation of Invention and Production

Dependent var.	(1) prod. indicator	(2) ln (sales)	(3)	(4)	(5)	(6)	(7)
R&D Indicator$_{fh,t}$	0.281***	1.164***		1.042***			
	(0.003)	(0.024)		(0.026)			
ln(patent)$_{fh,t}$			0.331***		0.329***	0.205***	0.181***
			(0.012)		(0.012)	(0.044)	(0.042)
ln (distance)$_{fh,t}$				-0.024	-0.328**		
				(0.025)	(0.144)		
common language$_{fh,t}$				0.220***	0.408		
				(0.051)	(0.267)		
contiguity$_{fh,t}$				0.143***	0.224		
				(0.049)	(0.235)		
colonial tie$_{fh,t}$				0.090**	-0.563*		
				(0.046)	(0.306)		
Observations	7494979	119659	19519	119503	19519	14090	8839
R^2	0.704	0.495	0.572	0.496	0.572	0.969	0.963
Within R^2	0.042	0.045	0.093	0.047	0.094	0.022	0.020
Firm-period FE	Y	Y	Y	Y	Y	Y	Y
Host-period FE	Y	Y	Y	Y	Y	Y	-
Home-host FE	Y	Y	Y	Y	Y	-	-
Host-industry FE	Y	Y	Y	Y	Y	-	-
Affiliate FE	-	-	-	-	-	Y	Y
Host-industry-period FE	-	-	-	-	-	-	Y

Note: Column 1 estimates the relationship between having an R&D center in a host and the probability of having a production facility in the same host. Columns 2 through 7 estimate the relationship between having an R&D center (and the size of the R&D center) and the size of the production facility in the same host. Columns 4 and 5 also control for the average distance of a production facility to other countries in which the firm has an R&D center. Industry effects are at two-digit level. Standard errors (in parenthesis) are clustered by firm. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

The OLS estimate of $\gamma_{R\&D}$ and $\overrightarrow{\gamma}_{dist}$ might suffer from an omitted variable bias. For example, firms might conduct both invention and production in hosts with a comparative advantage in their industry; both activities might also grow in response to an expansion in the economy or market access of a host. To rule out these factors, I include host-period, home-host, and host-industry fixed effects. I further control for firm-period fixed effects to absorb across-the-board growth of a firm. As discussed previously, these controls also help purge out systematic variations in patenting propensities across firms and hosts, so $x_{fh,t}$ can be interpreted as R&D.

Table 4 reports the results. The first column focuses on the extensive margin and shows that having an R&D center in a host increases the probability of having a production affiliate in the same country by 0.28, or about ten times the mean value of the outcome variable (0.027). Column 2 focuses on the intensive margin of production and finds that MNCs on average produce 116% more in hosts where they have an R&D center. Column 3 uses log patent count as the explanatory variable. The sample size is substantially smaller as only affiliates with inventions are included, but the result is similar qualitatively: sales are correlated with the number of inventions at the affiliate level. Based on Columns 2 and 3, respectively, Columns 4 and 5 further include the average value of the four distance measures between h and all countries in which firm f has an R&D center. The distance coefficients are generally consistent that the proximity to sibling R&D centers in other countries also have an effect.

statistically significant coefficients for contemporary and lagged offshore R&D, and small and insignificant coefficient for future offshore R&D. Such dynamic results, available upon request, are also consistent with a colocation effect.

These results are indicative of frictions impeding the separation of invention and production, but they can also be due to idiosyncratic reasons—beyond what can be controlled for by host and firm fixed effects—that make a host suitable for specific firms. To address this concern, Column 6 controls for affiliate fixed effects to ensure that the correlation is not due to high idiosyncratic match quality between a firm and a host. Column 7 further includes host-industry-period fixed effects, so it rules out the effect of *changes* in host economy that affects the entire industry. Both specifications find that as affiliates increase invention, their production also increase.[9]

A remaining possibility is that the correlation is driven by over-time *changes* in the idiosyncratic match quality between a host and particular affiliates. Given that the estimate changes little with the inclusion of host-industry-period fixed effects, it appears that the scope for such shocks to affect the estimate is limited. Nevertheless, I adopt an alternative IV strategy to further address this concern, under the following identifying assumption: controlling for host fixed effects and other time-varying host, industry, and firm characteristics, changes in the R&D environment of a host—e.g., R&D subsidies and the number of researchers—affect affiliate production through affiliate R&D. I discuss the IV results in Appendix A.5, which also suggest that affiliate invention is correlated with production. This strategy rules out that the colocation of invention and production is due to over-time changes in idiosyncratic match quality, thus complementing the specifications in Table 4. While both identification strategies are imperfect, the fact that they give a similar result despite exploiting orthogonal sources of variation, boosts the confidence that the finding is not due to an omitted variable bias.

Fact 3: Affiliate production and invention both decrease with distance to the headquarters, but at different rates. The third fact concerns the headquarter effects on affiliate invention and production. The specification is the following:

$$y_{foh,t} = FE + \overrightarrow{\gamma}_{dist} \cdot \overrightarrow{dist}_{oh} + \epsilon_{ohf,t},$$

in which the outcome variable $y_{foh,t}$ is measures of invention or production in host h of firm f from country o in period t. $\overrightarrow{dist}_{oh}$, the distance measures between headquarters o and host h, is the focus of this specification. I exclude headquarters from the sample, so the comparison is among affiliates of the same firm in different countries.

Columns 1 and 3 of Table 5 report the results of a linear probability model in which the outcome variable is an indicator for having a production affiliate or an R&D center in h. They show that geographic frictions play important but heterogeneous roles. Sharing a language is important for both invention and production, whereas distance and colonial ties matter more for production. These estimates are economically sizable compared to the mean value of the dependent variables (0.018 and 0.027, respectively). Columns 2 and 4 estimate the intensive margin effect of distance

[9]Two remarks on the within-affiliate specification: First, when exploiting over-time variations, firms growing from a small number of patents can have an extraordinary percentage growth rate. To avoid these firms having an outsized impact on the estimate, I weight firms by the square root of their patent counts. Not weighting would result in an estimate of 0.06 (s/e= 0.014) for Column 6 and an estimate of 0.09 (s/e= 0.02) for Column 7. Second, I do not include the average distance metrics as there are not enough over-time changes to estimate them precisely. Including these variables will not impact the coefficient for log(patent) materially.

Table 5: The Headquarter Effect on Invention and Production

	(1)	(2)	(3)	(4)	(5)	(6)
	Affiliate Invention		Affiliate Production			
Dependent var.	indicator	ln(patent)	indicator	ln (sales)	indicator	ln (sales)
ln(distance)$_{oh}$	-0.002**	-0.129***	-0.005***	-0.282***	-0.005***	-0.253***
	(0.001)	(0.034)	(0.002)	(0.028)	(0.001)	(0.027)
common language$_{oh}$	0.020***	0.258***	0.022***	0.162**	0.015**	0.094
	(0.004)	(0.072)	(0.009)	(0.064)	(0.007)	(0.061)
contiguity$_{oh}$	0.002	0.106	0.004	0.185***	0.003	0.174***
	(0.002)	(0.072)	(0.004)	(0.061)	(0.004)	(0.058)
colonial tie$_{oh}$	0.002	0.029	0.024***	0.153**	0.023***	0.129*
	(0.004)	(0.067)	(0.008)	(0.075)	(0.007)	(0.068)
R&D indicator$_{fh,t}$					0.375***	1.198***
					(0.019)	(0.031)
Observations	7295102	45364	7295102	103131	7295102	103131
R^2	0.124	0.336	0.302	0.420	0.339	0.446
Within R^2	0.004	0.012	0.006	0.012	0.058	0.056
Firm-period FE	Y	Y	Y	Y	Y	Y
Host-industry FE	Y	Y	Y	Y	Y	Y
Host-period FE	Y	Y	Y	Y	Y	Y

Note: Columns 1 and 2 estimate the relationship between affiliate invention and the proximity of the host to the headquarters. Columns 3 to 6 estimate the relationship between the proximity to the headquarters and affiliate production, among which Columns 5 and 6 also include the R&D center indicator. Headquarters are excluded from all regressions. Standard errors (in parenthesis) are clustered by country pair. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

to the headquarters on affiliate activities. Sharing a common language is more important for invention, but other types of geographic frictions are in general more important for production.

These results do not necessarily mean that proximity to the headquarters matters for both invention and production—given Fact 2, as long as one activity is hindered by distance to the headquarters, regressions on the other activity will show expected signs for geographic frictions. To show qualitatively that the headquarter effects for both are at play, Columns 5 and 6 add the R&D center indicator in the regressions for affiliate production. The estimated geographic coefficients decrease slightly but remain statistically significant, consistent with a headquarter effect for production. Without a formal model, it is challenging to quantitatively separate the strength of colocation and headquarter effects. I will disentangle these forces by combining these estimates with a structural model in later sections.

Robustness and additional evidence. Together, these facts highlight two motives for firms conducting R&D in a host: to take advantage of the host talent and to be close to production. The variation in R&D intensity across countries and the differential headquarter effects also show that while R&D and production are interconnected decisions, firms do have the flexibility in choosing the appropriate site for each. These facts call for a model in which R&D and production are explicitly introduced as separate decisions of MNCs. In the rest of this paper, I develop such a model to interpret the data and to quantify the aggregate implications of offshore R&D.

In Appendix A.4, I show that the three facts are robust to different measures of affiliate R&D (e.g. citation) and production (value added), and that estimates are similar when only the manufacturing industry is included. In Appendix A.6, I provide additional evidence for a key assumption of the model. This result will be discussed as I introduce the model below.

3 The Model

3.1 Environment

Worker. There are N countries, indexed by $i = 1, 2, ...N$. Country i is endowed with L_i measure of workers, whose ability α is from a distribution with a cumulative distribution function (cdf) $A_i(\alpha)$. Workers with ability α choose whether to work in manufacturing production, where everyone has the same effective productivity and earns a common wage of W_i^l (l for low-skill), or to work in high-skill jobs—R&D and marketing—and earn a wage of $W_i^h \times \alpha$.[10] Letting $\hat{\alpha}_i$ denotes the ability level above which a worker chooses high-skill jobs, we have

$$W_i^h \hat{\alpha}_i = W_i^l.$$

The supply of high and low skill efficiency units, denoted by L_i^h and L_i^l respectively, are given by:

$$L_i^h = L_i \cdot \int_{\alpha > \hat{\alpha}_i} \alpha \, dA_i(\alpha)$$
$$L_i^l = L_i \cdot A_i(\hat{\alpha}_i).$$

Firm. Country i is endowed with E_i measure of firms differing in their innovation management efficiency, $\tilde{z}^R \in \tilde{\mathbb{Z}}^R$, drawn from a cdf denoted by $G_i^E(\tilde{z}^R)$. This cdf captures the state of knowhow in country i, which I interpret as adopted management practices conducive for innovation but can also more broadly capture the accumulated technological expertise in country i. Firms build R&D centers in different countries, which recruit local researchers to develop new varieties, and then engage in offshore production and trade.

Consumption. A representative consumer in country i decides the allocation of consumption expenditures according the preference represented by the following utility function:

$$U_i = (\int_{\Omega_i} q_i(\omega)^{\frac{\sigma-1}{\sigma}} d\omega)^{\frac{\sigma}{\sigma-1}},$$

where Ω_i denotes the set of product varieties available in country i, $q_i(\omega)$ is the consumption of variety ω, and $\sigma > 1$ is the elasticity of substitution. Let the aggregate consumption expenditure in country i be X_i. The demand for variety ω is:

$$q_i(\omega) = p_i(\omega)^{-\sigma} \frac{X_i}{P_i^{1-\sigma}},$$

[10] By allowing for an endogenous occupation choice, I am able to use external data on ability distributions to calibrate $A_i(\alpha)$. An alternative approach is to assume each country has an exogenous endowment of researchers and other types of workers. Parameterizing that alternative model boils down to assuming that the endowment of different types of workers in a country matches the observed occupation shares, which risks attributing variation in other country characteristics that drive occupation choice, such as firm knowhow, to talent endowment.

where $p_i(\omega)$ is the price of variety ω in country i and $P_i = [\int_{\Omega_i} p_i(\omega)^{1-\sigma} d\omega]^{\frac{1}{1-\sigma}}$ is the price index.

3.2 Firm Decisions: Overview and Empirical Support

This subsection gives an overview of firm decisions and discusses the empirical support for the assumption on the independence among R&D centers. I will use the following notations: o denotes a firm's headquarters, the country where a firm originates and obtains its draw of $\tilde{z}^R$; i denotes the country where a product is invented—the location of the R&D center; m denotes the country where manufacturing production takes place; and d denotes the destination country where a product is consumed.

Consider a firm from country o. Knowing its innovation efficiency in the home country, $\tilde{z}^R$, the firm decides how many R&D centers to open and in which countries. To open an R&D center in country i, it pays a fixed cost of c^R_{oi} in country i high-skill efficiency unit, a cost that varies by pairs of countries. Motivated by Figure 2b and Fact 3, I assume that firms can transfer part of their knowhow to affiliates. Letting $\phi^R_{oi} \leq 1$ be the fraction of innovation efficiency transferred, the innovation efficiency for an R&D center in country i owned by a country o firm is $z^R = \tilde{z}^R \phi^R_{oi}$. This efficiency governs how many varieties can be developed for a given number of researchers.

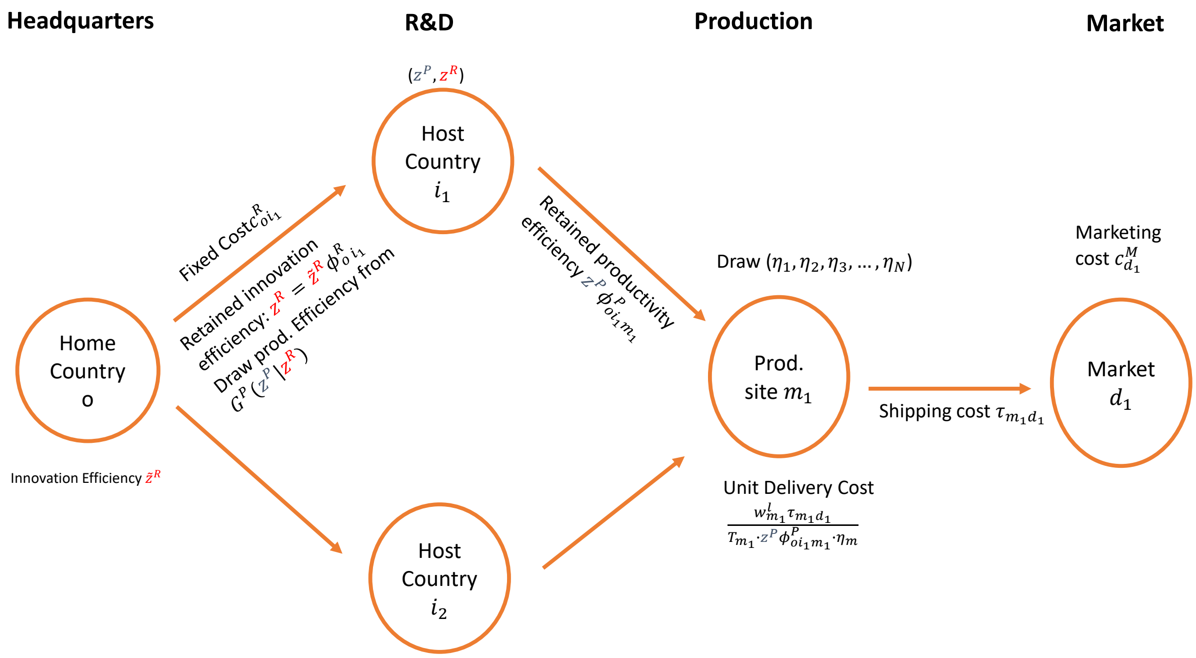
Figures 2b and 2c together imply that the affiliate *sales per invention* increase in the parent innovation efficiency. I interpret this as more innovative firms on average having higher manufacturing efficiency/product quality. In keeping with this interpretation, each R&D center upon entry also obtains a random draw of production management efficiency, denoted $z^P \in \mathbb{Z}^P$, which is common to all products developed by the R&D center. The distribution from which z^P is drawn increases in z^R in the sense of first-order stochastic dominance. I use $G^P(z^P|z^R)$ to denote the cdf of production efficiency draws and use $g^P(z^P|z^R)$ to denote its probability density function (pdf).

Firms may open R&D centers in multiple countries, but at most one in each. The firm decides how many products to develop in each center. To sell products to destination country d, a *per-variety* fixed marketing cost of c^M_d in terms of country d high-skill labor must be paid.

Manufacturing production can take place in countries where a firm does not have an R&D center. By separating production from R&D and the headquarters, firms can take advantage of cheaper production labor and save on shipping fees. However, geographic separation makes it difficult for researchers and the management to communicate with the production line, hurting production efficiency. I use $\phi^P_{oim} \leq 1$ to denote the fraction of productivity that a firm can transfer to its production center. Motivated by Facts 2 and 3, I will parameterize ϕ^P_{oim} to be a function of both the distance between o and m and the distance between i and m, thus allowing proximity to both the headquarters and R&D centers to play a role.

Production uses only labor. For an R&D center with production efficiency z^P, the *deterministic* component of its labor productivity in country m is $T_m \cdot z^P \phi^P_{oim}$, with T_m being the country-wide manufacturing efficiency at m, capturing the influence of infrastructure, institution, etc. I further assume that there is a stochastic element, η_m, idiosyncratic to individual varieties, which enters productivity multiplicatively, so the *variety-level* productivity in m is $T_m \cdot z^P \phi^P_{oim} \eta_m$. The cost of producing one unit in m and delivering it to d is $\frac{W^l_m \tau_{md}}{T_m z^P \phi^P_{oim} \eta_m}$. This takes into account the wage for

Figure 3: Overview of the Model



low-skill worker, W^l_m, and iceberg shipping fee, τ_{md}.

Figure 3 illustrates firm decisions. Importantly, I assume that varieties developed by a firm, in the same or different R&D centers, are differentiated from each other and from varieties developed by all other firms. This assumption is consistent with how R&D is organized in many conglomerates. General Electric, for example, organizes its ten research labs by scientific disciplines in five countries (the U.S., Germany, India, China, and Brazil). With this assumption, firms make offshore R&D decisions for each country independently, which greatly simplifies the model.[11,12]

I test this assumption empirically. In Appendix A.6, I show that affiliate invention responds to R&D subsidies, the availability of researchers, and IPR protection in the host, but *does not* respond to these factors in either the headquarters or other countries in which the firm has a presence. This suggests interdependence among R&D centers is not a first-order feature of my data.

Given the independence, in the remainder of this section, I first explain the production and trade decisions of a firm after a variety has been developed. I then describe the innovation decision

[11]More than 70% of FDI flows in the data are in the form of mergers and acquisitions (Nocke and Yeaple, 2008). The differentiated-variety assumption adopted in the present paper is consistent with this form of FDI—multinationals transfer their managerial knowhow to newly acquired foreign firms, which have their own brands and products, and help them carry out independent product development and manufacturing production.

[12]The model treats R&D at headquarters and R&D in offshore centers symmetrically. Bilir and Morales (2020) find that for the U.S. multinationals, R&D at headquarters have stronger spillover effects to their affiliates overseas than R&D at one affiliate to other affiliates. An extension of this model that allows firms to first invest in R&D to build up 'core management capacity' before performing product innovation at home and abroad would be consistent with this pattern. Whether we treat R&D at the headquarters differently will not affect the qualitative conclusions from the empirical section as firm fixed effects are always included; my quantitative experiments based on the present model can be viewed as a short-term version that does not allow for endogenous accumulation of the core innovation efficiency.

of R&D centers and entry into offshore R&D. Finally, I define the equilibrium and characterize the welfare gains from openness.

3.3 Production and Trade

Consider a variety developed by an R&D center (z^P, z^R) in country i from o. The R&D center obtains N idiosyncratic productivity draws, one for each potential production site, denoted by $\boldsymbol{\eta} = (\eta_1, \eta_2, .., \eta_N) \in \boldsymbol{H}$. The cost of serving country d through country m is $c_{oimd} = \frac{W_m^l \tau_{md}}{T_m z^P \phi_{oim}^P \eta_m}$.

Knowing the realization of $\boldsymbol{\eta}$, the firm decides whether to sell this variety to market d. Entry to d for each variety requires hiring c_d^M high-skill labor units from d for marketing, each costing W_d^h. After paying this, the firm chooses the lowest cost production location for each of its varieties. Because there are no fixed costs in offshore production, all countries are potential production sites. With monopolistic competitive markets, the price of this variety in country d is a constant markup over the lowest unit delivery cost among all choices:

$$p_{oid}(z^P, \boldsymbol{\eta}) = \min_m \{ \frac{\sigma}{\sigma - 1} \frac{W_m^l \tau_{md}}{T_m z^P \phi_{oim}^P \eta_m} \}.$$

The operation profit from manufacturing and selling this variety to country d is

$$\pi_{oid}(z^P, \boldsymbol{\eta}) = \mathbb{1}(p_{oid}(z^P, \boldsymbol{\eta}) < \hat{p}_d) \cdot [\frac{1}{\sigma} \frac{X_d}{P_d^{1-\sigma}} p_{oid}(z^P, \boldsymbol{\eta})^{1-\sigma} - c_d^M W_d^h],$$

where $\hat{p}_d \equiv \left(\frac{\sigma W_d^h c_d^M}{X_d} \right)^{\frac{1}{1-\sigma}} P_d$ is the cutoff price below which a firm can sell enough to recover the marketing cost. The total revenue in d from this variety is

$$r_{oid}(z^P, \boldsymbol{\eta}) = \mathbb{1}(p_{oid}(z^P, \boldsymbol{\eta}) < \hat{p}_d) \cdot \frac{X_d}{P_d^{1-\sigma}} p_{oid}(z^P, \boldsymbol{\eta})^{1-\sigma}.$$

For tractable aggregation, I assume the following for the rest of this paper:

Assumption 1.

$$a) \quad F(\boldsymbol{x}) \equiv Prob(\eta_1 \leq x_1, ..., \eta_N \leq x_N) = \begin{cases} 1 - (\sum_{m=1}^{N} \frac{1}{N} x_m^{-\theta}), \forall m \in \{1, ...N\}, \ x_m \geq 1 \\ 0, \ \exists m \in \{1, ..., N\}, \ x_m < 1 \end{cases}$$

$$b) \quad \hat{p}_d < \frac{\sigma}{\sigma - 1} \frac{1}{z^P} \frac{W_m^l \tau_{md}}{T_m \phi_{oim}^P}, \ \forall \ o, \ i, \ m, \ d, \ z^P \in \mathbb{Z}^p$$

$F(\boldsymbol{x})$ is a special case of the multivariate distribution proposed in Arkolakis et al. (2018). It has two attractive implications. First, $p_{oid}(z^P, \boldsymbol{\eta})^{1-\sigma}$ follows a Pareto distribution.[13] Second, together

[13]More precisely, $p_{oid}(z^P, \boldsymbol{\eta})^{1-\sigma}$ follows a Pareto *away* from its minimum support. Assumption 1.b implies that among firms actively selling to d, that minimum support never binds. The economic content of this assumption is that the marketing cost is high enough so that for all firms there are bad enough realizations of $\boldsymbol{\eta}_m$ such that they will not enter d. As long as this assumption is maintained, c_d^M can vary by firm and its exact values would not matter.

with Assumption 1.b, it implies that the expected fraction of sales in country d that is fulfilled through production in country m, denoted by ψ_{oimd}, is

$$\psi_{oimd} = \frac{\frac{1}{N}(\frac{T_m \phi^P_{oim}}{W^l_m \tau_{md}})^\theta}{\tilde{\zeta}^\theta_{oid}}, \text{ where } \tilde{\zeta}_{oid} \equiv [\sum_m \frac{1}{N}(\frac{T_m \phi^p_{oim}}{W^l_m \tau_{md}})^\theta]^{\frac{1}{\theta}}.[14] \tag{1}$$

z^P does not enter the expression because it is common to all production sites.

Letting $\bar{r}_{oid}(z^P)$, $\overline{\pi}_{oid}(z^P)$, and $\bar{c}^M_{oid}(z^P)$ denote the expected values of per-variety sales revenue, operation profit, and marketing cost, associated with market d before $\boldsymbol{\eta}$ realizes, I show in the appendix that these objects are given by:

$$\begin{aligned} \bar{r}_{oid}(z^P) &\equiv \mathbb{E}(r_{oid}(z^P, \boldsymbol{\eta})|z^P) = \frac{\theta(\sigma-1)^\theta \sigma^{1-\frac{\theta\sigma}{\sigma-1}}}{\theta-(\sigma-1)} X_d^{\frac{\theta}{\sigma-1}} P_d^\theta (w^h_d c^M_d)^{\frac{\theta+1-\sigma}{1-\sigma}} (\tilde{\zeta}_{oid} z^P)^\theta \\ \bar{c}^M_{oid}(z^P) &\equiv \mathbb{E}(c^M_{oid}(z^P, \boldsymbol{\eta})|z^P) = \frac{\theta-(\sigma-1)}{\theta\sigma} \bar{r}_{oid}(z^P) \\ \overline{\pi}_{oid}(z^P) &\equiv \mathbb{E}(\pi_{oid}(z^P, \boldsymbol{\eta})|z^P) = \frac{1}{\sigma} \bar{r}_{oid}(z^P) - \bar{c}^M_{oid}(z^P). \end{aligned} \tag{2}$$

The total operation profit from a variety by an R&D center in country i is therefore

$$\overline{\pi}_{oi}(z^P) = \sum_d \overline{\pi}_{oid}(z^P). \tag{3}$$

3.4 R&D Decisions

R&D centers recruit high-skill workers to develop new varieties. Letting h be the measure of high-skill efficiency units recruited, the total number of new varieties developed, v, is:[15]

$$v = z^R \cdot h^\gamma, \ 0 < \gamma < 1.$$

Two comments on this setup are in order. First, I model R&D as the invention of differentiated products. In quantification, I will map varieties to patents, which has been documented to be a strong predictor of new product introduction (Argente et al., 2018). An alternative is to model R&D as process innovation, which improves manufacturing efficiency of both the affiliate in the same host and sibling affiliates through spillovers. I wish to note that such spillovers are internalized within the firm and are also present in my setup, where inventions in an R&D center influence the sales of affiliates in other countries because they too can produce these varieties. Ultimately, both models imply that R&D in a host increases the revenues of the firm both locally and globally, and are observationally equivalent when only affiliate level sales/value added data are available.

[14]Note $\sum_m \psi_{oimd} = 1$. $\tilde{\zeta}_{oid}$ differ slightly from other work using similar parametric assumption (with the multiplier $\frac{1}{N}$) because I specify $F(\mathbf{x})$ to be a function of $\frac{1}{N}$, thereby normalizing the support of η_m, $m \in 1, 2, ..., N$ to be $[1, \infty)$.

[15]In an earlier version of this paper (Fan, 2017), I assume that the quantity and quality of researchers are imperfect substitutes and that R&D output is log supermodular in z^R and researcher ability. That feature leads to positive assortative matching between firms and researchers and has additional implications on firms' location choice. Since most of the qualitative predictions of the model do not depends on that feature, I abstract from it here for simplicity.

Second, $\gamma < 1$ implies decreasing return to the size of research teams. This reflects the importance in R&D of affiliate-level fixed factors, embodied in z^R. An example of such fixed factors is the capacity of top managers in supervising product development, which does not scale easily.[16]

The optimization problem for an R&D center is

$$\pi^R_{oi}(z^P, z^R) = \max_h \overline{\pi}_{oi}(z^P) z^R h^\gamma - W^h_i h,$$

where $\overline{\pi}_{oi}(z^P)$ is the expected per-variety profit. The first order condition implies

$$h_{oi}(z^P, z^R) = \Big(\frac{\gamma \overline{\pi}_{oi}(z^P) z^R}{W^h_i}\Big)^{\frac{1}{1-\gamma}},$$

which gives the number of varieties invented as:

$$v_{oi}(z^P, z^R) = z^{R\,\frac{1}{1-\gamma}} \Big(\frac{\gamma \overline{\pi}_{oi}(z^P)}{W^h_i}\Big)^{\frac{\gamma}{1-\gamma}}. \tag{4}$$

The total operation profit from these varieties is $\overline{\pi}_{oi}(z^P) \cdot v_{oi}(z^P, z^R)$. Of this, a γ fraction goes to researchers. The rest is the return to firm knowhow, given by:

$$\pi^R_{oi}(z^P, z^R) = (\gamma^{\frac{\gamma}{1-\gamma}} - \gamma^{\frac{1}{1-\gamma}}) \Big(\frac{1}{W^h_i}\Big)^{\frac{\gamma}{1-\gamma}} (\overline{\pi}_{oi}(z^P) z^R)^{\frac{1}{1-\gamma}}. \tag{5}$$

This return also represents the total variable profit from doing R&D in country i.

3.5 Offshore R&D and the Distribution of R&D Centers

Let $\pi^R_{oi}(z^R)$ be the expected profit (across all possible z^P draws) for an R&D center in country i, given its innovation efficiency z^R. We have:

$$\pi^R_{oi}(z^R) = \int_{\mathbf{Z}^P} \pi^R_{oi}(z^P, z^R) g^P(z^P | z^R) dz^P$$

To characterize firms' decision to open offshore R&D centers, I assume the following about $g^P(z^P|z^R)$.

Assumption 2. *The distribution from which an R&D center draws its production efficiency z^P increases in the innovation efficiency of the R&D center in the sense of first-order stochastic dominance.*

Firms compare the expected profit from building an offshore R&D center to the fixed cost of setting up the center, $c^R_{oi} W^h_i$. By Equations (2) and (5) , $\pi^R_{oi}(z^P, z^R)$ increases in z^P and z^R. Assumption 2 then implies that $\pi^R_i(z^R)$ increases in z^R, so there exists a cutoff $\hat{z}^R_{oi}$, such that firms from country o will perform offshore R&D in country i if and only if its innovation efficiency at the

[16]See Antras et al. (2006) for an analysis of an offshoring model in which managers can only supervise a fixed number of workers. Alternatively, decreasing returns might stem from increasing coordination costs, free-riding, and disagreement among researchers, as a team expands.

headquarters $\tilde{z}^R$ is above $\hat{z}^R_{oi}$. This cutoff is given by:

$$\pi^R_{oi}(\hat{z}^R_{oi}\phi^R_{oi}) = c^R_{oi}W^h_i.$$

The measure of R&D centers from o to i, denoted by R_{oi} is thus:

$$R_{oi} = E_o \cdot \Big(1 - G^E_o(\hat{z}^R_{oi})\Big),$$

in which E_o is the measure of firms from country o and $G^E_o(\tilde{z}^R)$ is the cdf of the innovation efficiency at the headquarters o. The cdf of the innovation efficiency of the affiliated R&D centers in country i, denoted G^R_{oi}, is characterized by:

$$R_{oi}G^R_{oi}(z^R) = \mathbb{1}(z^R > \hat{z}^R_{oi}\phi^R_{oi}) \cdot E_o \cdot G^E_o(\frac{z^R}{\phi^R_{oi}}).$$

The associated pdf is:

$$g^R_{oi}(z^R) = \frac{1}{R_{oi}}\mathbb{1}(z^R > \hat{z}^R_{oi}\phi^R_{oi}) \cdot E_o \cdot g^E_o(\frac{z^R}{\phi^R_{oi}}) \cdot \frac{1}{\phi^R_{oi}}.$$

Finally, the density distribution of the R&D centers from country o over (z^P, z^R) is:

$$g_{oi}(z^P, z^R) = g^R_{oi}(z^R) \cdot g^P(z^P|z^R).$$

3.6 Aggregation

I now derive the expressions for aggregate outcomes. Letting $V_{oi}(z^P)$ be the marginal density of varieties invented in country i by R&D centers from country o with production efficiency z^P, i.e.,

$$V_{oi}(z^P) = R_{oi}\int_{\mathbb{Z}^R} v_{oi}(z^P, z^R) \cdot g_{oi}(z^P, z^R) dz^R, \tag{6}$$

where $v_{oi}(z^P, z^R)$ is given by Equation (4). The total measure of varieties across all z^P is:

$$V_{oi} = \int_{\mathbf{Z}^P} V_{oi}(z^P) dz^P.$$

In the appendix, I show that the price index in country d satisfies

$$P_d^{1-\sigma} = \theta(\frac{\sigma}{\sigma-1})^{-\theta}\frac{1}{\theta-(\sigma-1)}\Big(\frac{\sigma W^h_d c^M_d}{X_d}\Big)^{\frac{\theta-(\sigma-1)}{1-\sigma}} P_d{}^{\theta-(\sigma-1)} \sum_o \sum_i \tilde{\zeta}^\theta_{oid} \int_{\mathbf{Z}^P} (z^P)^\theta V_{oi}(z^P) dz^P. \tag{7}$$

The sales to country d of the varieties developed in i by firms from o, denoted by X_{oid}, is

$$X_{oid} = \theta(\frac{\sigma}{\sigma-1})^{-\theta}\frac{1}{\theta-(\sigma-1)}(\sigma W^h_d c^M_d)^{\frac{\theta-(\sigma-1)}{1-\sigma}}\Big(\frac{X_d}{P_d^{1-\sigma}}\Big)^{\frac{\theta}{\sigma-1}} \tilde{\zeta}^\theta_{oid} \int_{\mathbf{Z}^P} (z^P)^\theta V_{oi}(z^P) dz^P. \tag{8}$$

Among theses sales, let X_{oimd} be the value fulfilled through production in country m. Because Equation (1) holds for each variety making up X_{oid}, it applies to the aggregate flows, i.e.,

$$X_{oimd} = \psi_{oimd} X_{oid}.$$

The sales revenue is shared among participants at different stages of the process. First, the bulk of sales ($\frac{\sigma-1}{\sigma}$) is the manufacturing value added taking place in country m. Letting Y_{om} be the manufacturing income generated through production for firms from o in country m, we have

$$Y_{om} = \frac{\sigma-1}{\sigma} \sum_{i,d} X_{oimd}$$

The markup is split among three participants. From Equation (2), a $\frac{\theta-(\sigma-1)}{\theta}$ fraction is used for marketing in destination d. The total marketing costs incurred in country d by firms from o, denoted by C^M_{od}, is:

$$C^M_{od} = \frac{1}{\sigma} \cdot \Big(\frac{\theta-(\sigma-1)}{\theta}\Big) \sum_{i,m} X_{oimd}.$$

The reminder of the markup is value added by researchers in country i, in the forms of both product development and the overhead of R&D centers, and the net profit to firm owners at o. Denoting the variable profit of firms from o operating R&D centers in country i by Π_{oi}, we have:

$$\Pi_{oi} = \frac{1-\gamma}{\sigma} \cdot \Big(\frac{\sigma-1}{\theta}\Big) \sum_{m,d} X_{oimd}.$$

The payment to researchers working directly on product developed in country, denoted by I_{oi}, is:

$$I_{oi} = \frac{\gamma}{1-\gamma} \Pi_{oi}.$$

The payment to researchers in the form of overhead R&D cost is

$$C^R_{oi} = E_o \cdot [1 - G^R_o(\hat{z}^R_{od})] \cdot c^R_{oi} W^h_i.$$

The labor market clearing condition for low-skill workers is

$$W^l_d L^l_d = \sum_o Y_{od}.$$

The labor market clearing condition for high-skill workers is

$$W^h_d L^h_d = \underbrace{\sum_o I_{od}}_{\text{product development}} + \underbrace{\sum_o C^M_{od}}_{\text{marketing entry}} + \underbrace{\sum_o C^R_{od}}_{\text{R\&D overhead}}.$$

Total income equals total expenditures

$$X_d = \underbrace{W_d^h L_d^h + W_d^l L_d^l}_{\text{Labor Income}} + \sum_i \underbrace{(\Pi_{di} - C_{di}^R)}_{\text{Net Profits}}. \tag{9}$$

Definition 1. *The competitive equilibrium is defined as a set of allocations and prices, such that workers and firms optimize, all goods and labor markets clear, and all firm-level decisions are consistent with aggregate allocations and prices (see appendix for details).*

3.7 The Gains from Openness

MNCs integrate countries into a GVC, starting from headquarter services embedded in knowhow provision, to R&D, production, and eventually marketing and consumption. When firms carry out R&D abroad, participants at different stages benefit. First, high-skill workers in the host and firm owners share the rent from the newly created varieties. Because production tends to colocate with R&D, low-skill workers in the host and nearby countries benefit through a higher demand for labor. Finally, high-skill workers in destination countries benefit from an increase in labor demand for marketing; everyone benefits from having more varieties.

Given the interaction among these forces, characterizing the welfare implications of a general policy on offshore R&D is challenging. To illustrate the intuition, I focus on the effect on the aggregate welfare, defined as the total real income of a country, of a particular counterfactual: moving the economy from the observed level of international linkages to complete isolation.[17] For *only this section*, I make the following assumption:

Assumption 3. *a) The distribution from which firms draw production efficiency z^P is independent of z^R but can vary arbitrarily across pairs of origin country and R&D location, o and i.*
b) Firm innovation efficiency, $\tilde{z}^R$, follows a Pareto distribution: $G_o^E(x) = 1 - (\frac{x}{\underline{Z}_o^R})^{-\kappa_R}$, with $\underline{Z}_o^R$ being the country specific lower bound and $\kappa_R > (1-\gamma)$ being the tail coefficient.
c) L_i^h and L_i^l are exogenous, i.e., there is no occupation choice.

Part (a) of the assumption departs from Assumption 2 as it rules out positive correlation between z^P and z^R at the firm level. On the other hand, z^P can be from any distributions and differ across pairs of origin and host countries, so heterogeneity at country-pair level or the headquarter effects on z^P are allowed. Parts (b) and (c) are technical assumptions that simply expressions. Under Assumption 3, we have the following:

Proposition 1. *Under Assumption 3, the gains from openness for country d, defined as the percentage change in $\frac{X_d}{P_d}$ as d moves from complete isolation to the observed equilibrium, is*

$$GO_d = \left(\frac{X_{dddd}}{\sum_m X_{ddmd}}\right)^{-\frac{1}{\theta}} \times \left(\frac{\sum_{o,m} X_{odmd}}{X_d}\right)^{-\frac{1}{\theta}} \times \left(\frac{\sum_m X_{ddmd}}{\sum_{o,m} X_{odmd}}\right)^{-\frac{1}{\theta}} \times \left(\frac{I_{dd}}{\sum_o I_{od}}\right)^{\frac{\gamma}{\theta}} \times f\left(\frac{\sum_o I_{od}}{X_d}, \frac{I_{dd}}{\sum_o I_{od}}\right) \times \frac{X_d}{Y_d} - 1, \tag{10}$$

[17] I use the total real income of a country, $\frac{X_d}{P_d}$, as a measure for the aggregate welfare, with X_d defined in Equation (9). To the extent that different workers are differentially affected by openness, this aggregate welfare function implicitly assumes that national governments can use lump-sum transfers for redistribution, so only the total income matters.

where $f(\frac{\sum_o I_{od}}{X_d}, \frac{I_{dd}}{\sum_o I_{od}})$ is a function of model parameters and two ratios: the share of variable R&D expenses in income $\frac{\sum_o I_{od}}{X_d}$, and the fraction of domestic R&D expenses by local firms $\frac{I_{dd}}{\sum_o I_{od}}$.

Proof. See appendix. □

This expression illustrates forces through which economic integration affects welfare. The first term, $\frac{X_{dddd}}{\sum_m X_{ddmd}}$, measures the importance of trade. In the absence of colocation and headquarter effects, i.e., when ϕ^P_{oim} is independent of both i and o, this term equals $\frac{\sum_{o,i} X_{oidd}}{\sum_{o,i,m} X_{oimd}}$, which is the fraction of sales in country d that is manufactured locally, the same as in the familiar gains from trade formula. The second term, $\frac{\sum_{o,m} X_{odmd}}{X_d}$, is the fraction of country d expenditure on goods invented in d. It captures the benefit from having access to goods invented elsewhere, by either domestic or foreign firms. The first two terms capture the direct effect of trade and the possibility to spatially separate production from invention.

The third term, $\frac{\sum_m X_{ddmd}}{\sum_{o,m} X_{odmd}}$, reflects the importance of foreign firms in domestic R&D. Intuitively, the smaller this term is, the more country d depends on foreign affiliates and the larger the losses are when foreign R&D centers are shut down. This channel is counteracted by the ability of domestic firms to make up for the lost varieties, as the exit of foreign firms free up researchers to local firms. Holding the total number of researchers constant, the strength of this offsetting channel is captured by the share of local firms in R&D, $(\frac{I_{dd}}{\sum_o I_{od}})^{\frac{\gamma}{\theta}}$, the fourth term.

Together, the first four terms capture the direct impact of openness on the *real wage* of production workers. The last two terms account for the indirect influence of openness on the aggregate real income through changes in income compositions. More specifically, $\frac{X_d}{Y_d}$ captures that openness allows some countries to specialize in innovation and earn a higher share of income through innovation rent. For such countries, $\frac{X_d}{Y_d}$ is higher and the welfare gains are above and beyond the gains in the real wage. $f(\frac{\sum_o I_{od}}{X_d}, \frac{I_{dd}}{\sum_o I_{od}})$ adjusts for the endogenous response to openness of the fraction of the high skill working in R&D and of the income share from innovation. These forces are captured by two sufficient statistics, $\frac{\sum_o I_{od}}{X_d}$ and $\frac{I_{dd}}{\sum_o I_{od}}$.

In the appendix, I compare this formula to that from Arkolakis et al. (2018). Slightly different versions of the first two and the last terms in Equation (10) also appear in their formula. The third and fourth terms, $(\frac{X_{ddd}}{\sum_o X_{odd}})^{-\frac{1}{\theta}} \times (\frac{I_{dd}}{\sum_o I_{od}})^{\frac{\gamma}{\theta}}$, which summarize the direct effect of offshore R&D, is novel. To gauge the importance of this effect, consider a special case with ϕ^P_{oim} independent of i, in which case $\frac{\sum_m X_{ddmd}}{\sum_{o,m} X_{odmd}} = \frac{\sum_{i,m} X_{dimd}}{\sum_{o,i,m} X_{oimd}} = \frac{I_{dd}}{\sum_o I_{od}}$, so the product of the third and fourth terms simplifies to $(\frac{I_{dd}}{\sum_o I_{od}})^{-\frac{1-\gamma}{\theta}}$. Consider the median country in the quantitative section, with about 30% of its R&D done by foreign affiliates. The value of $(\frac{I_{dd}}{\sum_o I_{od}})^{-\frac{1-\gamma}{\theta}}$ is around 1.05, when $\gamma = 0.25$ and $\theta = 4.5$—all else equal, offshore R&D brings about 5% direct welfare gains.

In addition to bringing about direct welfare gains, offshore R&D also affects the measured gains from trade and offshore production. With R&D and production colocation, my model infers different values for the first two terms of the formula. In the rest of this paper, I discipline these channels using the firm-level data and show that doing so is important for quantifying both the gains from offshore R&D and the overall gains from openness.

4 Parameterization

The quantitative analysis focuses on the same 37 countries as in Section 2. I parameterize the model to match the following data: the geographic coefficients estimated in Section 2, statistics on the firm size distribution, and endowment, income, and openness by country. This section describes the procedures, starting with functional form assumptions.

4.1 Additional Assumptions

Talent and innovation efficiency distributions. I parameterize the ability distribution for workers in country i, $A_i(\alpha)$, to be log normal, i.e., $\log(a_i) \sim N(\mu_\alpha^i, {\sigma_\alpha^i}^2)$.

I assume firm innovation efficiency is drawn from a truncated Pareto distribution as below:

$$G_o^E(\tilde{z}^R) = \frac{(\underline{Z}_o^R)^{-\kappa_o^R} - (\tilde{z}^R)^{-\kappa_o^R}}{\underline{Z}_o^{R^{-\kappa_o^R}} - \overline{Z}_o^{R^{-\kappa_o^R}}}, \tag{11}$$

in which the $\underline{Z}_o^R$ and $\overline{Z}_o^R$ are the lower and upper bounds of the support and κ_o^R is the tail parameter of the distribution, all of which can vary by country.[18]

Relationship between z^P and z^R. I assume that $G(z^P|z^R)$, the cdf of z^P draws for an R&D center with efficiency z^R, takes the following form:

$$G(z^P|z^R) = \text{Prob}(z^P \in H|z^R) \cdot G_H^P(z^P) + [1 - \text{Prob}(z^P \in H|z^R)] \cdot G_L^H(z^P), \tag{12}$$

in which $G_H^P(z^P)$ and $G_L^P(z^P)$ are two distributions (high and low) from which firms draw z^P. The probability of drawing from the high distribution increases in z^R:

$$\text{Prob}(z^P \in H|z^R) = \frac{\exp(\delta_0 + \delta_1 \times z^R)}{1 + \exp(\delta_0 + \delta_1 \times z^R)}. \tag{13}$$

δ_0 and δ_1 are to be estimated. A positive δ_1 means that innovative firms tend to be more productive on average. $G_H^P(z^P)$ and $G_L^P(z^P)$ are both Pareto distributions with different supports but the same tail index κ^P:

$$G_H^P(z^P) = 1 - (\frac{\underline{z}_H^P}{z^P})^{\kappa_P},\ G_L^P(z^P) = 1 - (\frac{\underline{z}_L^P}{z^P})^{\kappa_P}, \quad \underline{z}_L^P < \underline{z}_H^P.$$

Specifying $G(z^P|z^R)$ as in Equation (12) makes the model tractable as it circumvents the need for numerically integrating over the two dimensional (z^P, z^R) space (see, e.g. Equation (6) and (7)).

Geographic frictions. I parameterize frictions impeding offshore production and R&D as log

[18] The distribution converges to the Pareto distribution as $\overline{Z}_o^R \to \infty$.

linear functions of various distance measures and a host-specific fixed effect as follows:

$$\begin{cases} \log(\phi^P_{oim}) = s \cdot \log(\phi^P_{im}) + (1-s) \cdot \log(\phi^P_{om}),\ s \in [0,1],\ \text{where} \\ \log(\phi^P_{om}) = \mathbb{1}(o \neq m) \cdot [\phi^P_m + \overrightarrow{\beta^{P,om}} \cdot \overrightarrow{dist_{om}}] \\ \log(\phi^P_{im}) = \mathbb{1}(i \neq m) \cdot [\phi^P_m + \overrightarrow{\beta^{P,im}} \cdot \overrightarrow{dist_{im}}] \end{cases}$$
$$\begin{cases} \log(\phi^R_{oi}) = \mathbb{1}(o \neq i) \cdot [\phi^R_i + \overrightarrow{\beta^R} \cdot \overrightarrow{dist_{oi}}] \\ c^R_{oi} = \mathbb{1}(o \neq i) \cdot \exp\left(\phi^{cR}_i + \overrightarrow{\beta^{cR}} \cdot \overrightarrow{dist_{oi}}\right) \end{cases} \tag{14}$$

The first block of Equation (14) defines ϕ^P_{oim}, the retained production efficiency in country m, as the geometric average of ϕ^P_{im} and ϕ^P_{om}, with s capturing relative importance of the two terms. ϕ^P_{om} and ϕ^P_{im} are themselves functions of ϕ^P_m, which captures host-specific effects on inward offshore production, and bilateral distance measures. As in the reduced-form analysis, four distance measures are included: geographic distance and indicators for whether the two countries share an official language, are adjacent, or have a colonial tie. The second block in Equation (14) defines the retained efficiency and the fixed cost in offshore R&D, ϕ^R_{oi} and c^R_{oi}. Through the inclusion of ϕ^R_i and ϕ^{cR}_i, the parameterization allows different degrees of openness to foreign R&D by host. Finally, the indicator functions in these definitions normalize ϕ^P_{om}, ϕ^R_{oi}, ϕ^P_{im} to 1 and c^R_{oi} to zero for domestic activities.

4.2 Parameters Assigned Directly

I directly assign values to some parameters of the model and this subsection explains how.

Labor endowment. Because patenting and (especially) trade are by and large a manufacturing activity, I interpret the model as for the manufacturing industry. I set the number of workers in a country, L_i, to the manufacturing employment of a country, calculated using its total employment from the Penn World Table and manufacturing share from the World Bank.

Elasticities. The elasticity of substitution between varieties, σ, determines markups. I set this parameter to be 4. This value implies markups of 30%, in line with recent findings using firm-level data from the U.S. and elsewhere (De Loecker and Eeckhout, 2018). θ governs the dispersion of the idiosyncratic component of affiliate productivity. As Equation (2) suggests, the share of sales devoted to marketing is $\frac{\theta-(\sigma-1)}{\theta\sigma}$. According to a recent survey of chief marketing officers in the U.S.,[19] firms spend around 8-10% of revenues on marketing. I set $\theta = 4.5$, so marketing accounts for 8.3% of sales. This choice of θ is also close to the estimate off trade flows by Arkolakis et al. (2018). With θ and σ given, γ determines the share of sales spent on developing new varieties ($\frac{\gamma(\sigma-1)}{\theta\sigma}$), which I map to the revenue share of R&D expenses in the data. Compustat U.S. manufacturing firms on average spend 4% of sales on R&D. I set $\gamma = 0.25$ so $\frac{\gamma(\sigma-1)}{\theta\sigma} \approx 0.04$. After deducting production cost, marketing, and product development, the remaining component of sales are the variable profit. Accounting for about 12% of sales, this component covers both the

[19] See https://cmosurvey.org/about/

overhead in offshore R&D and the net return to firm knowhow.

Trade costs. I normalize $\tau_{mm} = 1$ and assume symmetric trade costs. As shown in Appendix C.2, the approach of Head and Ries (2001) applies to this model, so we can back out trade costs as:

$$\tau_{md} = \tau_{dm} = \left(\frac{\sum_{o,i} X_{oimd}}{\sum_{o,i} X_{oimm}} \cdot \frac{\sum_{o,i} X_{oidm}}{\sum_{o,i} X_{oidd}}\right)^{-\frac{1}{2\theta}}.$$

Although X_{oimd} is not observable, the four sums in the expression are observable from conventional input-output tables. I aggregate across manufacturing industries in the World Input Output Database and pin down bilateral trade costs using this equation.

Knowhow and talent distributions. Calibrating knowhow and talent distributions requires comparable data across countries. I use the World Management Survey by Bloom et al. (2012) for the former and the cognitive test score data by Hanushek and Woessmann (2012) for the latter.

The World Management Survey provides firm-level management scores for a number of countries. In the survey, interviewers rate each firm based on its talent management policy and production efficiency along various dimensions. The overall management score for a firm is a summary of these sub scores and has been shown to be strongly correlated with objective measures of firm performance (Bloom et al., 2012). The talent management score intends to capture whether firms follow good management practices for retaining and incentivizing its talent, so it closely maps to the capacity of a firm in R&D, where talent plays a crucial role. I use it to calibrate innovation management distributions. I obtain three distribution statistics of $\tilde{z}^R$ for each country: mean, standard deviation, and skewness, and pick the parameters governing the distribution in Equation (11) so that the model statistics match their empirical counterparts.[20]

I define the production management score to be the average of targeting, operation, and monitoring scores, which captures the efficiency of a firm in carrying out production. With this, I estimate δ_0 and δ_1 in Equation (13). Specifically, I classify a firm as being from the high productivity distribution if its production management score falls into the top 5% of all firms in the sample (corresponding to about top 12% in the U.S.). I then estimate the relationship between a firm's innovation management score and the probability that it is from a high productivity distribution using the Logit specification implied by Equation (13). This procedure, reported in Appendix C.2, determines $\delta_0 = -5$, $\delta_1 = 0.21$.[21]

For talent distributions, from the test score database I obtain the average cognitive score and

[20]The talent management score in the data follows a symmetric bell-shaped distribution. Since it is well known that the firm size distribution has a fat tail, I calculate the three statistics based on the exponent of the original scores, effectively treating the reported scores as logarithm of actual innovation efficiency. Some countries in quantification are not covered by the World Management Survey. I impute their statistics based on income and the geographic regions of countries. Appendix C.1 reports the distribution statistics for management scores.

[21]The choice of the top 5% cutoff is motivated by the importance of the right-tail firms in international business. A high cutoff allows me to better capture the activities of these firms. An implicit assumption in using the Logit regression to estimate δ_0 and δ_1 is that firms drawing their production efficiency from $G_L^P(z^P)$ constitute the bottom 95% in the production efficiency distribution, whereas firms drawing from $G_H^P(z^P)$ constitute the top 5%. This assumption does not hold strictly because under the Pareto assumption, the supports of $G_L^P(z^P)$ and $G_H^P(z^P)$ overlap. Given the choice of the cutoff (5%), however, the calibrated $\underline{Z}_H^P$ will be large enough so the overlap is negligible. Figure C.1 in the appendix plots the probability of a firm drawing from $G_H^P(z^P)$ against z^R for the U.S.

Table 6: Parameters Calibrated Externally

Symbol	Description	Value	Source
σ	elasticity of substitution between varieties	4	markup (30%)
θ	dispersion of offshore production draws	4.5	marketing expenditure/sales (8.3%)
γ	researcher share of variable profit	0.25	R&D expense/sales (4%)
δ_0	probability of high production efficiency	-5	estimated (Table C.3)
δ_1	dependence of z^P on z^R	0.21	estimated (Table C.3)
$\{\tau_{md}\|m,d=1,...,N\}$	trade costs	-	World Input Output Database
$\{G_o^E(\tilde{z}^R)\|o=1,...,N\}$	innovation efficiency dist.	-	Bloom et al. (2012)
$\{A_i(\alpha)\|i=1,...,N\}$	talent dist.	-	Hanushek and Woessmann (2012)
$\{L_i\|i=1,..,N\}$	labor endowment	-	World Bank & PWT

the shares of students reaching 'basic' and 'top' performance level, which are defined according to a common standard across countries. I set the distribution of the U.S. to the standard log normal (i.e., $\mu^{\alpha}_{US} = 1$, $\sigma^{\alpha}_{US} = 1$), and then determine talent distributions of other countries by matching their distribution statistics relative to those of the U.S.[22]

Table 6 summarizes the parameters determined directly.

4.3 Parameters Determined in Equilibrium

The remaining parameters, determined jointly, include production efficiency distribution parameters, $\underline{z}^P_L$, $\underline{z}^P_H$, and κ_P, country-specific productivity, $\{T_m|m = 1,...,N\}$, measures of firms in a country $\{E_o|o = 1,..,N\}$, frictions to offshore production and R&D, which include bilateral coefficients $\{\overrightarrow{\beta^{P,om}}, \overrightarrow{\beta^{P,im}}, \overrightarrow{\beta^{R}}, \overrightarrow{\beta^{cR}}\}$, host-specific components $\{\phi^P_m, \phi^R_i, \phi^{cR}_i|i, m = 1,..,N\}$, and the weight parameter s. I describe the intuition on how each parameter is determined and the numerical algorithm that recovers them jointly.

Firm production efficiency parameters. Having pinned down the distributions of innovation efficiency and the way it maps into productivity, governed by Equation (13), parameters in $G^P_H(z^P)$ and $G^P_L(z^P)$ are the remaining degrees of freedom for the firm size distribution. I normalize $\underline{z}^P_L$ to 1, and pick $\underline{z}^P_H$ and κ_P jointly. κ_P governs the shape of the firm size distribution at the very top, while $\underline{z}^P_H$ has an influence on the scale of the top 5% relative to the rest of firms. The targets are the Pareto tail coefficient for the firm size distribution, and the fraction of firms with fewer than 20 and 100 employment. All these targets are based on the U.S. data and will be compared to the corresponding statistics of the U.S. in the model economy.

Measure of firms. The measure of firms from country o, E_o, along with their innovation efficiency, determines the fraction of the world patents invented by firms from o, $\frac{\sum_d V_{od}}{\sum_{o,d} V_{od}}$. I normalize E_{US} and chose $\{E_o : o \neq \text{US}\}$ so $\frac{\sum_d V_{od}}{\sum_{o,d} V_{od}}$ in the model is aligned with its empirical counterpart.

Country-level efficiency. Given the distributions of firm knowhow and talent, T_m captures the residual variation in the overall efficiency of country m. I choose T_m so that $\frac{X_m}{P_m}$ matches the real

[22]Specifically, in the U.S., 91% and 7.3% students reach 'basic' and 'top' performance, respectively, which implies that the cutoffs for basic performance is 0.25 and for top performance is 4.27; the U.S. distribution has a mean of 4.9. I calibrate μ^{α}_i and σ^{α}_i for other countries so that their shares of population above the two cutoffs and their mean scores relative to that of the U.S. match the data on the L^2 norm.

GDP of country m.

Geographic frictions to offshore R&D and production. I determine the three sets of host-specific parameters, $\{\phi_m^P, \phi_i^R, \phi_i^{cR} | i, m = 1, .., N\}$, by matching the openness of a country to inward offshore R&D (both extensive and intensive margins) and inward offshore production. The moments that I target are the foreign share of patents invented in a host $\frac{\sum_{o,\, o\neq i} V_{oi}}{\sum_o V_{oi}}$, the foreign share of R&D center counts $\frac{\sum_{o,\, o\neq i} R_{oi}}{\sum_o R_{oi}}$, and the foreign share of manufacturing production $\frac{\sum_{o,\, o\neq m} Y_{om}}{\sum_o Y_{om}}$. The empirical counterparts of these moments are aggregated from firm-level data used in Section 2.2, adjusted as described in Appendix C.1 to address concerns on the data from specific countries.[23]

Given $\{\phi_m^P, \phi_i^R, \phi_i^{cR} | i, m = 1, .., N\}$, I infer s and $\{\overrightarrow{\beta^{P,om}}, \overrightarrow{\beta^{P,im}}, \overrightarrow{\beta^{R}}, \overrightarrow{\beta^{cR}}\}$ using an indirect-inference procedure that targets reduced-form estimates from five specifications in Section 2.3 that are most informative about these coefficients. The first three specifications are Columns 1, 2, and 4 of Table 5, which estimate the headquarter effects for the intensive margin of offshore production, and for the extensive and intensive margins of offshore R&D. These three regressions help identify $\overrightarrow{\beta^{P,om}}$, $\overrightarrow{\beta^{R}}$, and $\overrightarrow{\beta^{cR}}$. The remaining parameters, $\overrightarrow{\beta^{P,im}}$ and s, jointly determine the relative importance in production of access to R&D centers versus headquarters. I choose two regressions identifying precisely this as the target. First, Column 6 of Table 5. Controlling for distance to headquarters, the coefficient for the R&D center indicator conveys information on the importance of access to R&D. Second, Column 4 of Table 4. This specification controls for the home-host fixed effects and identifies the effect of being close to sibling R&D centers.[24]

To implement the indirect inference procedure, I simulate a sample of 50,000 firms from the model, with the sample weight of each country equal to its size. I then use this sample to estimate the five regressions above, including the exact same set of controls as in the corresponding specifications. I choose $\{\overrightarrow{\beta^{P,om}}, \overrightarrow{\beta^{P,im}}, \overrightarrow{\beta^{R}}, \overrightarrow{\beta^{cR}}\}$ and s to minimize the L^2 norm between regression coefficients based on the simulated data and those based on the actual data. With 22 target coefficients and 17 parameters, the model is over-identified. In constructing the objective function, I weight regression coefficients using the inverse of their empirical standard errors, acknowledging that these coefficients are not estimated with the same precision.

Numerical implementation. I determine the parameters in this subsection using a nested fixed point algorithm.[25] In the outermost layer, I choose $\underline{z}_H^P$ and κ^P to match the moments on

[23]In keeping with the interpretation of the model as for manufacturing, I keep only firms whose parents are in manufacturing (based on their core industry) in calculating countries' shares of production/R&D carried out by foreign firms. Because manufacturing firms can have non-manufacturing affiliates, my calculation of foreign shares (e.g. $\sum_{o,o\neq m} Y_{om}/\sum_o Y_{om}$) will include these affiliates. The rationale for this choice is that non-manufacturing affiliates might support other parties within the firm, in which case it is more appropriate to view them as part of the core business. I note that this choice does not automatically inflate $\sum_{o,o\neq m} Y_{om}/\sum_o Y_{om}$ because non-manufacturing affiliates of *domestic* manufacturing firms are also counted, so the numerator and denominator are consistently defined. An alternative choice is to calculate $\sum_{o,o\neq m} Y_{om}/\sum_o Y_{om}$ using all manufacturing affiliates, without regard to the industry of the parent. This alternative gives very similar foreign shares so it will not affect the quantitative findings substantively.

[24]Note that s and $\overrightarrow{\beta^{P,im}}$ can be separated because I have imposed that the weights of ϕ_{im}^P and ϕ_{om}^P sum to one. I do not target the regressions with patent counts as the explanatory variable because these regressions use much smaller samples and are more susceptible to the attenuation bias.

[25]An often used alternative is to cast the entire procedure as a constraint optimization problem, with all equilibrium conditions and the set of just-identified moments being the constraint. This alternative is not suitable here because each evaluation of the constraint optimization problem requires estimating firm-level regressions, which is costly given

Table 7: Parameters Calibrated in Equilibrium

Parameter and value				Description	Moment	Model	Data
A. Firm size dist. parameters							
$\underline{z}_L^P = 1$ (normalized)					% of firms with emp.<100	0.99	0.99
$\underline{z}_H^P = 2.01$				Firm z^P draws	% of firms with emp.<20	0.94	0.95
$\kappa^P = 6.15$					Power law coefficient of firm size dist.	1.05	1.05
B. Country-specific parameters and fixed effects							
$\{T_m \mid m = 1, ..., N\}$				country-specific man. TFP	$\{\frac{X_m}{P_m}\}$	-	Columns 1-5 of Table C.1
$\{E_o \mid o = 1, .., N\}$				measure of domestic firms	$\{\frac{\sum_i V_{oi}}{\sum_{o,i} V_{oi}}\}$	-	
$\{\phi_m^P \mid m = 1, .., N\}$				host effect in production	$\{\frac{\sum_{o,\, o\neq m} Y_{om}}{\sum_o Y_{om}}\}$	-	
$\{\phi_i^R \mid i = 1, .., N\}$				host effect in R&D	$\{\frac{\sum_{o,\, o\neq i} V_{oi}}{\sum_o V_{oi}}\}$	-	
$\{\phi_i^{cR} \mid i = 1, .., N\}$				host effect in R&D overhead	$\{\frac{\sum_{o,\, o\neq i} R_{oi}}{\sum_o R_{oi}}\}$	-	
C. Bilateral Geographic Coefficients							
s=0.82				$\phi_{oim}^P = (\phi_{im}^P)^s (\phi_{om}^P)^{1-s}$			
distance	**lang**	**border**	**colony**				
-0.063	0.014	0.045	0.022	$\overrightarrow{\beta^{P,im}}$	reduced-form estimates on colocation and headquarter effects	Table C.4 Panel B	Table C.4 Panel A
-0.032	0.038	0.022	0.054	$\overrightarrow{\beta^{P,om}}$			
-0.061	0.17	0.043	0.026	$\overrightarrow{\beta^{R}}$			
0.12	-0.027	-0.10	-0.005	$\overrightarrow{\beta^{cR}}$			

firm size. In the middle loop, I search over the space of $\{\overrightarrow{\beta^{P,om}}, \overrightarrow{\beta^{P,im}}, \overrightarrow{\beta^{R}}, \overrightarrow{\beta^{cR}}\}$ and s to minimize difference in regression coefficients. In the innermost loop, I solve for the competitive equilibrium, while choosing $\{T_m \mid m = 1, ..., N\}$, $\{E_o \mid o = 1, .., N\}$, and$\{\phi_m^P, \phi_i^R, \phi_i^{cR} \mid i, m = 1, .., N\}$ to match their respective targets described above. Once the innermost loop is completed, I simulate the model to obtain regression coefficients to feed into the middle loop. Appendix C.4 provides details on the implementation of the algorithm.

Parameter values and model fit. Table 7 summarizes the results from this procedure. Panel A is the parameters determined in the outermost loop: $\underline{z}_H^P = 2$ and $\kappa^P = 6.15$. The model matches closely the three moments of the U.S. firm size distribution. Panel B is the country specific parameters that are matched perfectly by design in the innermost layer. The target values of the moments are reported in Appendix Table C.1.

Panel C of Table 7 reports the coefficients governing the costs of offshore R&D and production. I find $s = 0.82$, meaning that proximity to the R&D team is more important than proximity to the headquarters. This reflects the colocation pattern seen in the data. All geographic coefficients have expected signs. In line with the heterogeneous headquarter effects documented before, the distance coefficients are different between R&D ($\overrightarrow{\beta^{R}}$) and production ($\overrightarrow{\beta^{P,om}}$). For example, sharing an official language with the headquarters increases official R&D efficiency by 17% but increases offshore production efficiency by only 3.8%. Appendix Table C.4 reports the model and data moments for parameters in Panel C side by side. It shows that the model fits data well: all but three of the 22 coefficients are within the 95% confidence intervals of empirical estimates.

that a large number of firms are simulated, and many fixed effects are included. The nested approach circumvents this problem because it only runs these regressions for when the innermost loop is satisfied.

Table 8: Fit of Non-targeted Moments

Additional moments on firm size in U.S.	**Model**	**Data**
Fraction of firms with emp. < 10	0.91	0.90
Share of emp. in firms with emp. > 500	0.611	0.47
Share of R&D by parents of MNCs	0.84	0.79
The efficiency advantage of foreign affiliates		
Foreign affiliate advantage	0.21	0.15
coefficient of variation across countries	1.272	1.158
correlation with host log GDP per capita	-0.11	-0.25
Entry into Offshore R&D		
% of firms with R&D centers in 1 country	93.5	95.3
2 countries	1.9	2.7
3 countries	0.7	0.6
4 countries	0.6	0.3
5 countries	0.3	0.3
>= 6 countries	3.0	0.7

4.4 Validation using Non-targeted Moments

I evaluate the fit of the model on non-targeted moments in Table 8.

Employment and R&D concentration. The upper panel reports additional statistics on firm size. In the data, 90% of firms have fewer than 10 employees and 47% of employment are in firms with more than 500 employees. The model matches the former well and over predicts the latter. In 2014, about 79% of the business enterprise R&D in the U.S. is conducted by parent firms of U.S. MNCs. This share is slightly higher in the model.

The multinational managerial advantage. In the model, firm knowhow, disciplined using management scores, is the main source of firm heterogeneity. Self-selection by knowhow into offshore R&D implies that foreign affiliates tend to have higher management scores than indigenous firms and that this advantage is larger in low-income host countries populated with poorly managed firms.

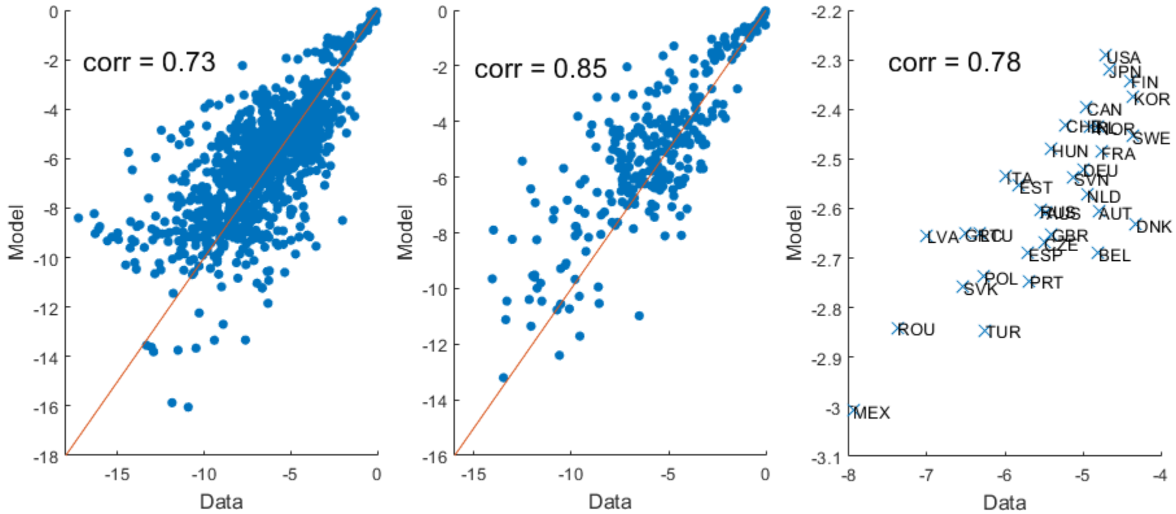
To validate these implications quantitatively, I calculate the foreign affiliate managerial advantage for each country, defined as the percentage difference between the average innovation efficiency of foreign affiliates and that of indigenous firms, and compare this measure to their empirical counterparts, constructed using the World Management Survey. As the middle panel of Table 8 shows, the predicted average foreign affiliate management advantage and its cross-country variations are both similar to those in the data. In addition, both the model and the data show a negative correlation between this measure and host income.

Offshore R&D entry. In the lower panel of Table 8 are the shares of firms with R&D centers in different numbers of countries. The model fits the data reasonably well except for the share of firms entering more than 6 countries: with all firms from a country facing the same offshoring entry cost, the most efficient firms in the model tend to enter more countries than in the data.

Bilateral offshore production and R&D. The calibration matches the overall inward offshore

Figure 4: The Fit of the Model on Aggregate Moments

(a) Offshore Production (b) Offshore R&D (c) Share of High-Skill Workers



Notes: The left and middle panels show the fit of the model in log bilateral production shares ($\log(\frac{Y_{om}}{\sum_o Y_{om}})$) and log bilateral offshore R&D shares ($\log(\frac{V_{oi}}{\sum_o V_{oi}})$), respectively. The right panel shows the fit of the model in log high-skill occupation shares. The vertical axis is $\log(1 - A_i(\hat{a}_i))$; the horizontal axis is log of the share of researchers in industrial employment obtained from the OECD.

R&D and production of each host by design, but when it comes to *bilateral* offshore production and R&D, the only information being used is the coefficients identified from within-firm variations. I assess whether the model matches the data on bilateral offshore production and R&D shares. Figure 4a and Figure 4b plot model-predicted log bilateral offshore production and R&D shares against their data counterparts. Both figures show a close fit, except for some country pairs on the left end, which have very small values in the data.

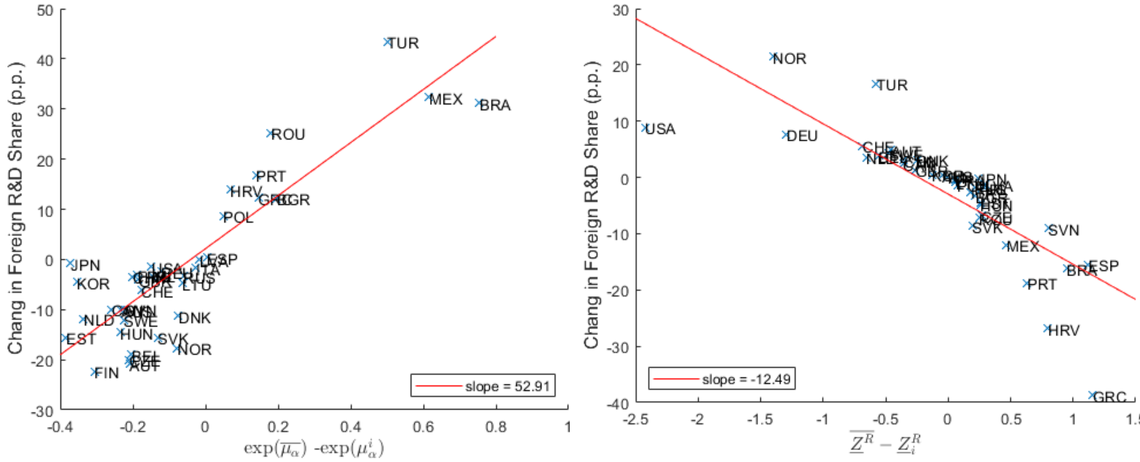
Occupation shares. By matching R&D by headquarter country of firms and the real income by country, my calibration indirectly disciplines the shares of national income from high-skill occupations and other sources. Whether the implied sources of income translate into a reasonable occupation distribution of workers boils down to if the log normal distribution, parameterized to match cognitive test scores, is a good approximation to the reality. As a validation of these assumptions, Figure 4c plots the share of workers sorting into high-skill occupations against the share of researchers in industrial employment from the OECD. The range of variation in the data is larger than in the model, likely because the definition of skill in the data is narrower. Despite this difference, the correlation between the two variables is quite high.

4.5 Evaluating the Two Motives of Offshore R&D

I evaluate the quantitative relevance of the two driving forces of offshore R&D, arising from the relative abundance in talent and market access of the host, respectively. To this end, in a set of experiments, I change either the measured distributions of talent and knowhow, or the market access, of a country. To isolate the impacts of changes in other countries, I change parameters for one single country at a time, keeping parameters of all other countries at the benchmark values.

Figure 5: The Role of Endowment Distributions

(a) Talent distribution (b) Firm Distribution



Notes: The vertical axis is the change (in p.p.) in the share of R&D by foreign firms. The horizontal axis in the left panel is $\exp(\overline{\mu}^i_\alpha) - \exp(\mu^i_\alpha)$, in which μ^i_α is the talent parameter of country i in the baseline economy (with $\exp(\mu^i_\alpha)$ being the median of the talent distribution), and $\overline{\mu_\alpha} \equiv \frac{\sum_i \mu^i_\alpha}{N}$ is the average talent parameter in the world. The horizontal axis in the right panel is $\overline{\underline{Z}^R} - \underline{Z}^R_i$, in which $\underline{Z}^R_i$ is the lower support of the firm knowhow distribution in country i and $\overline{\underline{Z}^R} \equiv \frac{\sum_i \underline{Z}^R_i}{N}$ is the average lower support in the world.

The role of endowment distributions. I first quantify the role of relative talent endowment. For each host, I separately set the location parameters for its talent and knowhow distributions, μ^i_α and $\underline{Z}^R_i$, to their respective mean values among the sample countries, $\overline{\mu_\alpha}$ and $\overline{\underline{Z}^R}$.

Figure 5a plots the change in inward offshore R&D for each country after its talent quality parameter μ^i_α is changed to the world average. Countries with $\mu^i_\alpha < \overline{\mu_\alpha}$, e.g. Brazil, see a surge in inward offshore R&D in response. Intuitively, as the domestic talent distribution improves, both domestic and foreign firms increase R&D in the host. The increase in the R&D of foreign firms is larger because more of them can now make enough profit to recoup the fixed entry cost. The change in inward offshore R&D correlates strongly with the improvement in talent in the experiment. According to the fitted line, a 0.2 increase in the horizontal axis, or 15% of the sample range, increases inward offshore R&D by around 10 p.p.

Figure 5b shows the change in offshore R&D shares as we set $\underline{Z}^R_i$ to the world average, $\overline{\underline{Z}^R}$. Countries with a high $\underline{Z}^R_i$, e.g., the U.S., experience an increase in inward offshore R&D. As domestic firms in a country become less competitive, its wage decreases and price increases, making the country more attractive to foreign entrants. This channel is quantitatively significant as well. According to the fitted line, increasing the horizontal axis by 1, roughly a quarter of the sample range, would decrease inward offshore R&D by 12.5 p.p.

The role of host country market access. In the model, foreign access consists of two components: access to foreign consumers through exporting and access to foreign manufacturers through offshore production. I consider their separate and joint impacts.

In the first experiment, I increase the export cost of host m, $\tau_{md}, m \neq d$, to infinity. This shuts down the direct access to foreign consumers of a host. The shares of R&D by foreign firms in this scenario are reported under Column 2 of Table 9, which shows an across-the-board increase

Table 9: The Determinants of Offshore R&D

Country	Benchmark	Eliminate Access to Consumer	Producer	Both	Country	Benchmark	Eliminate Access to Consumer	Producer	Both
	(1)	(2)	(3)	(4)		(1)	(2)	(3)	(4)
AUS	39.77	56.10	25.43	25.37	IRL	73.60	76.24	20.43	15.92
AUT	45.51	68.22	2.50	2.46	ITA	43.88	59.84	29.33	29.26
BEL	58.91	78.93	1.88	1.88	JPN	2.29	2.38	1.39	1.39
BGR	19.01	30.40	8.60	8.60	KOR	6.57	8.42	2.56	2.52
BRA	57.57	68.02	47.72	47.72	LTU	20.30	47.57	6.51	6.51
CAN	52.21	57.17	34.08	33.48	LVA	6.04	8.41	2.58	2.58
CHE	41.60	47.11	8.46	7.74	MEX	54.99	66.42	48.31	48.29
CHN	41.48	50.15	34.41	34.41	NLD	30.27	41.38	0.00	0.00
CZE	27.26	70.16	7.03	7.03	NOR	31.65	43.33	0.00	0.00
DEU	28.20	32.40	18.51	18.32	POL	14.63	50.67	2.24	2.24
DNK	38.45	63.29	0.56	0.46	PRT	26.22	52.65	2.79	2.79
ESP	22.10	42.92	3.57	3.60	ROU	53.45	72.33	40.15	40.15
EST	28.51	85.51	14.43	14.43	RUS	11.59	26.06	0.96	0.96
FIN	24.92	36.99	0.65	0.65	SVK	28.39	79.59	10.42	10.42
FRA	26.60	30.38	16.79	16.62	SVN	17.29	69.91	6.86	6.86
GBR	62.96	72.13	32.72	31.56	SWE	43.25	55.57	3.15	3.11
GRC	50.08	64.55	31.43	31.43	TUR	20.11	33.97	3.93	3.93
HRV	53.37	76.34	32.06	32.06	USA	15.77	15.91	10.82	10.80
HUN	55.81	79.43	44.66	44.66					

Notes: The numbers reported in this table are the share of domestic R&D expenditures incurred by affiliates of foreign companies in each country. All numbers are in percent. 'Benchmark' is for the baseline equilibrium; 'Consumer' is for when the access of a host to foreign consumers through exporting is shut down; 'Producer' is for when the access of a host to foreign producers through offshore production is shut down; 'Both' combines changes in 'Consumer' and 'Producer' for each country.

compared to the baseline (Column 1).

This result might be surprising at the first glance, given the following partial equilibrium intuition: eliminating access to foreign consumers reduces the return to doing R&D in a host, which should reduce the entry of MNCs. This direct channel is dampened because, unable to export, firms can still serve foreign consumers by offshoring production to other countries. Moreover, a general equilibrium effect kicks in: the reduction in export lowers production wages, pushing marginal workers into R&D, which makes the country more attractive for R&D. An increase in export costs of a host thus has a similar effect to a decrease in its production efficiency, which strengthens its comparative advantage in innovation.

The second experiment increases the cost of offshore production from an R&D host country to infinity (by setting $\phi^{p}_{oim} = 0, m \neq i$), so inventions in i, by both domestic and foreign firms, can be produced only locally. Column 3 of Table 9 shows that, most countries experience a decrease in offshore R&D from the baseline equilibrium. Unlike in the previous experiment, here the general equilibrium effect acts in the same direction as the partial equilibrium effect: when the option of offshore production is eliminated, R&D centers in the host countries have to produce locally to serve both foreign and domestic customers, which increases wages, making the country less attractive for R&D. An increase in offshore production costs can therefore be viewed similarly to a reduction in R&D innovation efficiency of a country, which strengthens its comparative advantage in production.

I shut down both types of market access simultaneously (Column 4). The resulting foreign R&D shares are much smaller than when only export is shut down (Column 2), because when offshore production is infeasible, countries can no longer specialize in innovation. On the other hand, without this specialization channel, the partial equilibrium effect of export on offshore R&D matters little: barring a few exceptions, Columns 4 is essentially the same as Column 3.

These experiments show that when calibrated to match the data, the two main forces incorporated in the model are quantitatively important. The difference in partial and general equilibrium intuition also underscores the value of the quantitative multi-country model.

5 Measuring the Role of Offshore R&D in Production and Income

I use the model for two purposes. The first is as a measurement device to infer the role of offshore R&D in firms' global organization of production, and the contribution of offshore R&D to national income. Both aspects have broad policy relevance, but neither are directly observable.

Distribution of manufacturing by R&D modes. Columns 1 through 4 of Table 10 decompose R&D by its source (whether done by local firms) and use (whether for local production). Columns 1 and 3 are the shares of R&D by domestic and foreign firms, calibrated to match the data.

Columns 2 and 4 are inferred through the lens of the model. The second column is the share of R&D done by domestic firms at home for local production, measured by the revenue of the varieties developed, i.e., $\frac{\sum_d X_{oood}}{\sum_{m,d} X_{oomd}}$. These shares average 82.9%, reflecting that it is costly to separate production from the headquarters and the R&D center at the same time. The fourth column reports the local production share of varieties developed by foreign affiliates. On average, 70% of foreign R&D leads to offshore production in the same host.[26] The shares vary across host. Figure 6a plots this share for each host against its average manufacturing cost, approximated by $\log(\frac{W_m^l}{T_m})$. It shows that offshore R&D in low-income countries are more likely to lead to local production than that in high-income countries.

Policy makers around the world extend generous R&D credits to foreign firms. The relatively high concentration of production around R&D sites in most hosts implies first, intended or not, such policies have an indirect effect in attracting manufacturing of foreign firms. Conversely, policies or changes in economic fundamentals that lead to the relocation of production to emerging countries can result in a reshuffle of R&D across countries as well.[27]

Offshore R&D and the sources of national income. Offshore R&D enables firms to apply their knowhow globally and shape the sources of income for all. Columns 5 to 9 decompose the income of a country into manufacturing production, profit (total and that from overseas inventions), R&D,

[26] That the majority of affiliate R&D is conducted for local production is consistent with Bilir and Morales (2020). Using a different data set (American MNCs) and a different approach (production function estimation), they find that R&D in an affiliate mostly applies to the affiliate itself and has limited spillovers on sibling affiliates.

[27] Making a host more attractive in offshore production can have direct and indirect effects on offshore R&D. The direct effect works to attract offshore R&D as firms can now produce alongside the R&D centers more easily; however, this also pushes the country to specialize in production, which could crowd out inward offshore R&D through a general equilibrium effect. Appendix C.3 investigates this possibility and shows that the net effect is ambiguous and depends on country characteristics.

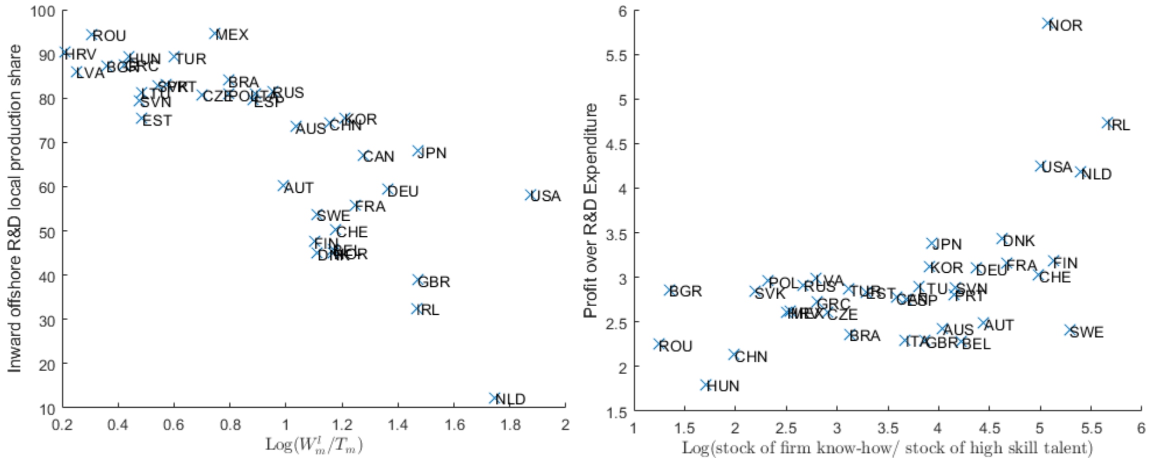
Table 10: Decomposition of R&D and Sources of Firm Profit

	Source and use of R&D				Source of income (% of total income)				
	% by domestic firms		% by foreign firms		mfg.	profit		R&D	mkt.
Country		% of local prod.		% of local prod.		total	inventions abroad		
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
AUS	60.2	87.3	39.8	73.5	82.6	6.4	0.24	2.6	8.3
AUT	54.5	78.8	45.5	60.1	82.1	6.8	0.44	2.7	8.3
BEL	41.1	63.6	58.9	45.9	85.0	4.6	0.72	2.0	8.3
BGR	81.0	97.2	19.0	87.4	79.2	9.2	0.01	3.2	8.3
BRA	42.4	97.4	57.6	84.1	76.6	10.6	0.01	4.5	8.3
CAN	47.8	84.9	52.2	67.1	75.5	11.9	2.37	4.3	8.3
CHE	58.4	71.3	41.6	50.3	74.1	13.2	2.16	4.4	8.3
CHN	58.5	95.0	41.5	74.3	76.7	10.2	0.00	4.8	8.3
CZE	72.7	92.7	27.3	80.7	83.8	5.7	0.01	2.2	8.3
DEU	71.8	70.5	28.2	59.4	76.9	11.2	1.69	3.6	8.3
DNK	61.6	66.5	38.4	44.9	79.3	9.6	2.40	2.8	8.3
ESP	77.9	90.8	22.1	79.5	81.4	7.6	0.09	2.7	8.3
EST	71.5	93.2	28.5	75.4	82.3	6.9	0.05	2.4	8.3
FIN	75.1	73.0	24.9	47.6	73.0	14.2	2.23	4.5	8.3
FRA	73.4	71.0	26.6	55.8	76.4	11.6	1.39	3.7	8.3
GBR	37.0	46.8	63.0	39.0	83.9	5.4	1.99	2.4	8.3
GRC	49.9	97.3	50.1	87.5	77.9	10.1	0.02	3.7	8.3
HRV	46.6	98.3	53.4	90.4	78.5	9.5	0.00	3.6	8.3
HUN	44.2	97.7	55.8	89.2	80.4	7.2	0.00	4.0	8.3
IRL	26.4	47.3	73.6	32.5	70.5	17.5	12.24	3.7	8.3
ITA	56.1	92.9	43.9	80.9	78.2	9.4	0.14	4.1	8.3
JPN	97.7	83.9	2.3	68.0	72.1	15.1	1.75	4.5	8.3
KOR	93.4	90.0	6.6	75.6	74.4	13.1	0.65	4.2	8.3
LTU	79.7	94.8	20.3	81.3	81.3	7.7	0.11	2.7	8.3
LVA	94.0	97.2	6.0	86.0	80.3	8.5	0.01	2.9	8.3
MEX	45.0	98.9	55.0	94.7	76.9	10.7	0.01	4.1	8.3
NLD	69.7	14.9	30.3	12.3	80.5	9.0	3.60	2.2	8.3
NOR	68.3	68.9	31.7	45.1	67.7	20.4	11.80	3.5	8.3
POL	85.4	92.1	14.6	80.9	81.6	7.5	0.11	2.5	8.3
PRT	73.8	94.9	26.2	83.0	81.3	7.7	0.01	2.7	8.3
ROU	46.5	98.5	53.5	94.4	84.3	5.1	0.00	2.3	8.3
RUS	88.4	91.7	11.6	81.6	79.1	9.4	0.06	3.2	8.3
SVK	71.6	94.5	28.4	82.8	85.1	4.8	0.02	1.7	8.3
SVN	82.7	94.7	17.3	79.5	79.7	8.9	0.02	3.1	8.3
SWE	56.7	74.9	43.3	53.8	77.9	9.7	0.78	4.0	8.3
TUR	79.9	97.9	20.1	89.3	75.4	12.0	0.00	4.2	8.3
USA	84.2	67.6	15.8	58.3	66.0	20.8	7.65	4.9	8.3
Mean	65.6	82.9	34.4	69.5	78.3	10.0	1.5	3.4	8.3

Notes: All numbers are in percent. Columns 1 and 3 report the source of R&D, i.e., whether it is by domestic firms (Column 1) or foreign firms (Column 3). Columns 2 and 4 report the fraction of R&D devoted to local production. Columns 5 to 9 decompose the fractions of income of a country from different activities: manufacturing production, profit (total and that accrued from products developed offshore), R&D, and marketing.

Figure 6: The Use of Offshore R&D and Its Influence on Income Compositions

(a) Local Production Share for Inward R&D (b) Ratio Between Profit and R&D Income



Notes: The left panel plots the share of inward offshore R&D devoted to local production against log average production cost. The right panel plots the ratio between profit and researcher income against the relative abundance of firm knowhow, measured as log of the ratio between the stock of firm knowhow and the stock of worker ability. The stock variables are constructed as the product of the measures of firms (workers) and their mean knowhow (ability).

and marketing. In autarky, the sources of income are the same across countries. In the open economy, while marketing expenditures still account for a fixed share of income, the importance of other sources are altered by firms' global operations. In particular, advanced countries populated with the most efficient firms tend to earn a higher share of income from profit. For instance, profit accounts for 21% of income in the U.S., more than a third of which is from overseas R&D centers—an important part of the U.S. knowhow only realizes its value through offshore R&D. In contrast, only 5% of income in Slovakia is from profit and almost all of it is generated from varieties invented domestically. This finding has implications on valuing the intangible assets of nations'.

The relative abundance of a country in knowhow versus talent governs the incentive for offshore R&D, thus having a direct impact on the compositions of income, especially the relative share of profit and R&D income. Without offshore R&D, the ratio between profit and R&D income is $\frac{\gamma}{1-\gamma}(=3)$; in the baseline economy, this ratio instead ranges from 1.7 to 6. Figure 6b plots this ratio against the relative abundance in knowhow. Countries relatively more abundant in firm knowhow derive a higher share of income from profit than from R&D.

6 Welfare Implications

The descriptive analyses show that offshore R&D is an integral part of firms' global production decisions and has a first-order impact on the income distribution for all. In this section, I conduct counterfactual experiments to uncover its normative implications.

6.1 Offshore R&D, Bridge R&D, and the Gains from Openness

I examine the gains from various forms of economic integration by eliminating each channel from the model separately. I define the gains from offshore R&D as the increase in the aggregate real income of a country ($\frac{X_d}{P_d}$) as the economy moves from the equilibrium where offshore R&D is prohibitively inefficient ($\phi^R_{oi} = 0, i \neq o$) to the baseline equilibrium; I define the gains from trade and offshore production analogously. I then compare the gains from these individual channels to the gains from openness.

The impact of offshore R&D. The first column in Table 11 is the gains from offshore R&D. The simple average across countries is 3.3%, roughly the same as the gains from trade. This average, however, masks important heterogeneity. While all countries are better off with offshore R&D, advanced countries enjoy substantially higher benefits than emerging economies.

The heterogeneity arises because with offshore R&D, advanced countries are getting a bigger slice of the innovation rent. As discussed in Section 5, offshore R&D increases the profit of the headquarters and the income of researchers in the host. Advanced countries, being main owners of firm knowhow and also generally more open to inward offshore R&D, benefit more from both sources. I measure the combined importance of these two sources using the share of total innovation rent—the sum of profit and researcher income in a country—that is generated through offshore R&D, i.e., the income of domestic researchers working at foreign affiliates plus the profit of domestic firms from overseas R&D. Figure 7a shows this measure explains most variations in the gains from offshore R&D.

I also calculate an equilibrium in which firms can carry out R&D abroad, but only to develop products to be manufactured in either the same host or the home country. This rules out overseas R&D dedicated to production in other foreign countries. Comparison between this equilibrium and the baseline economy gives the gains from 'bridge R&D,' reported in Column 2. For hosts where foreign R&D is mostly for local production, e.g., Mexico and Brazil, the gains from 'bridge R&D' are a small fraction of the gains from offshore R&D. For other countries, 'bridge R&D' is relatively more important, either because as hosts of foreign R&D centers, these countries have many inventions being manufactured in third countries (the case of the Netherlands, the U.K.), or because they are important manufacturing bases for bridge R&D conducted in surrounding countries (the case of Latvia, Poland).

Interaction among the three forms of globalization. Columns 3 and 4 report the gains from trade and offshore production, respectively. As expected, small countries close to major markets (e.g., Belgium) gain more from both channels, whereas large and remote countries (e.g., Brazil) gain less. Column 5 is the gains from openness, which average 15.6%.

We can understand the interaction among the three forms of integration by comparing the sum of the gains from these individual channels to the gains from openness. If the former is larger, it means the benefit of further openness is greater once a country is already open in other dimensions, so the three channels are complements for welfare; conversely, the three channels are substitutes.[28] I use as a measure of complementarity the ratio between the sum of gains from

[28]Note that the gains from individual channels are calculated from the calibrated equilibrium with the other two

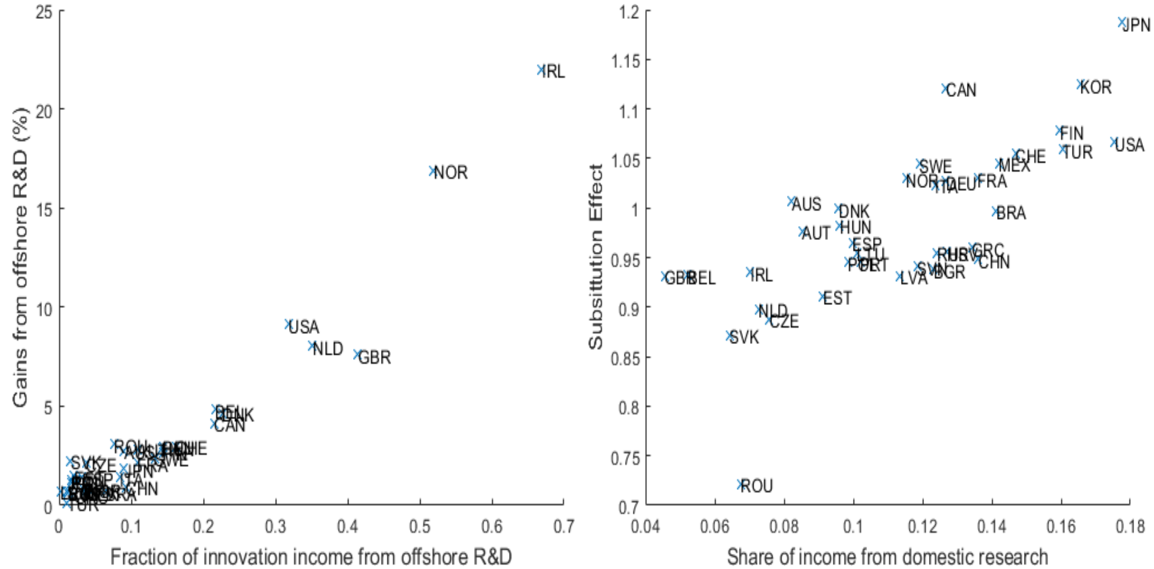
Table 11: Offshore R&D and the Gains from Openness

	Baseline model					Re-calibrated Alternative models		
	Offshore R&D					No Off. R&D	$s = 0$ (or $\phi^P_{oil} = \phi^P_{ol}$)	
Country	All (1)	Bridge (2)	Trade (3)	Off. prod. (4)	Openness (5)	Openness (6)	Off. R&D (7)	Openness (8)
AUS	2.76	1.95	0.62	7.98	11.29	9.08	1.73	14.15
AUT	2.81	2.70	4.02	10.14	17.38	15.27	4.24	23.97
BEL	4.84	4.51	5.60	21.09	33.83	27.60	3.99	39.40
BGR	0.64	0.54	2.79	2.26	6.07	5.90	13.14	22.26
BRA	0.68	0.18	0.71	0.79	2.19	1.58	7.04	10.43
CAN	4.14	1.82	3.74	5.55	11.98	8.08	10.97	26.11
CHE	2.98	1.86	4.59	9.06	15.77	13.66	6.54	25.81
CHN	0.93	0.12	0.50	1.14	2.70	1.89	1.89	5.31
CZE	2.14	2.03	3.68	6.65	14.05	12.88	6.64	23.76
DEU	2.96	1.64	4.40	10.10	17.00	15.19	4.12	23.19
DNK	4.64	3.44	4.35	13.88	22.88	19.62	6.02	32.12
ESP	1.41	1.26	1.43	5.55	8.70	8.23	5.23	15.94
EST	1.47	1.39	3.79	5.43	11.74	11.45	8.60	24.33
FIN	2.69	1.71	4.73	8.53	14.79	13.13	4.66	22.72
FRA	2.14	1.48	3.48	9.61	14.79	14.14	3.65	21.01
GBR	7.61	5.09	1.66	29.72	41.85	26.80	3.43	43.36
GRC	0.53	0.36	1.41	1.56	3.65	3.39	14.42	20.54
HRV	0.79	0.54	2.13	1.41	4.54	3.95	18.59	27.01
HUN	2.89	1.07	6.04	2.20	11.33	8.21	13.62	31.55
IRL	22.01	8.36	9.02	22.05	56.77	24.89	13.57	72.34
ITA	1.40	0.72	2.23	2.71	6.20	4.96	8.74	17.49
JPN	1.85	0.81	2.12	4.82	7.41	6.24	3.46	13.89
KOR	0.81	0.43	2.53	2.93	5.58	5.26	4.79	13.03
LTU	1.26	1.14	3.22	4.58	9.49	9.16	12.03	26.43
LVA	0.71	0.72	2.95	3.06	7.21	7.27	14.13	25.56
MEX	0.66	0.20	2.08	0.77	3.35	2.79	13.67	18.58
NLD	8.02	6.42	6.07	60.05	82.57	75.21	1.80	74.61
NOR	16.90	5.18	7.13	11.08	34.09	15.17	21.74	56.21
POL	1.24	1.27	2.37	5.48	9.63	9.63	8.05	20.88
PRT	1.13	1.10	1.39	4.34	7.26	7.01	7.49	17.06
ROU	3.08	1.23	0.84	3.80	10.70	7.70	11.71	31.54
RUS	0.68	0.69	1.23	3.60	5.78	5.93	5.08	12.73
SVK	2.26	2.29	4.29	8.31	17.03	16.59	8.85	32.58
SVN	0.71	0.66	3.78	3.13	8.10	8.10	9.88	20.88
SWE	2.38	2.10	4.34	8.79	14.84	13.09	1.70	18.80
TUR	0.12	0.09	1.78	0.69	2.45	2.44	14.11	17.72
USA	9.11	2.81	5.18	9.64	22.44	12.70	4.81	26.58
mean	3.34	1.89	3.30	8.45	15.61	12.28	8.22	26.21
std	4.42	1.84	1.92	10.60	15.97	12.32	4.93	14.64

Notes: All numbers are in percent. Columns 1 to 5 reports the gains from offshore R&D, offshore bridge R&D, trade, offshore production, and openness, respectively. Column 6 reports the gains from openness in a restricted model without offshore R&D, calibrated to match the same patterns of trade and offshore production as in the baseline equilibrium. Columns 7 and 8 report the gains from offshore R&D and openness for a re-calibrated model that assumes that $s = 0$, in which case there is no colocation

Figure 7: The Gains from Offshore R&D and Interaction with Trade and Offshore Production

(a) Gains from Off. R&D (b) Interaction with Trade and Off. Prod.



Notes: The left panel shows that the gains from offshore R&D (vertical axis) are higher in countries where offshore R&D account for a higher share of the innovation rent. The innovation rent is the sum of researcher income and profit in a country. Among it, the fraction accounted by offshore R&D is the profit of domestic firms doing R&D abroad plus the income of domestic researchers working in foreign R&D centers. On the right panel, the vertical axis is the ratio between the sum of individual gains and the gains from openness. The horizontal axis is the share of purely domestic innovation rent (profit and researcher income generated without foreign engagement) in total income.

individual channels and the overall gains from openness. This ratio, plotted in the vertical axis of Figure 7b, shows that the three channels are complements for some and substitutes for others.

This difference is due to the specialization of countries in the world economy. Trade and offshore production allows some countries to specialize in R&D and others in production. With efficient firms mobilizing innovation knowhow globally, offshore R&D increases R&D capacity in all countries. This enhances the comparative advantage of countries specializing in innovation but—as foreign R&D centers bid up domestic wages—weakens the comparative advantage of those specializing in production. As a result, the three channels are complements for the former group of countries and substitutes for the latter group of countries.

I use as a proxy for the comparative advantage of a country in innovation the share of its income from the rent created by purely domestic innovation, i.e., the profit and researcher income from varieties developed by domestic researchers with domestic firms only. This proxy captures the comparative advantage of a country in innovation, stemming from both firm knowhow and talent distributions. Figure 7b shows that this proxy is strongly correlated with the degree of complementarity/substitution of the three channels for countries.

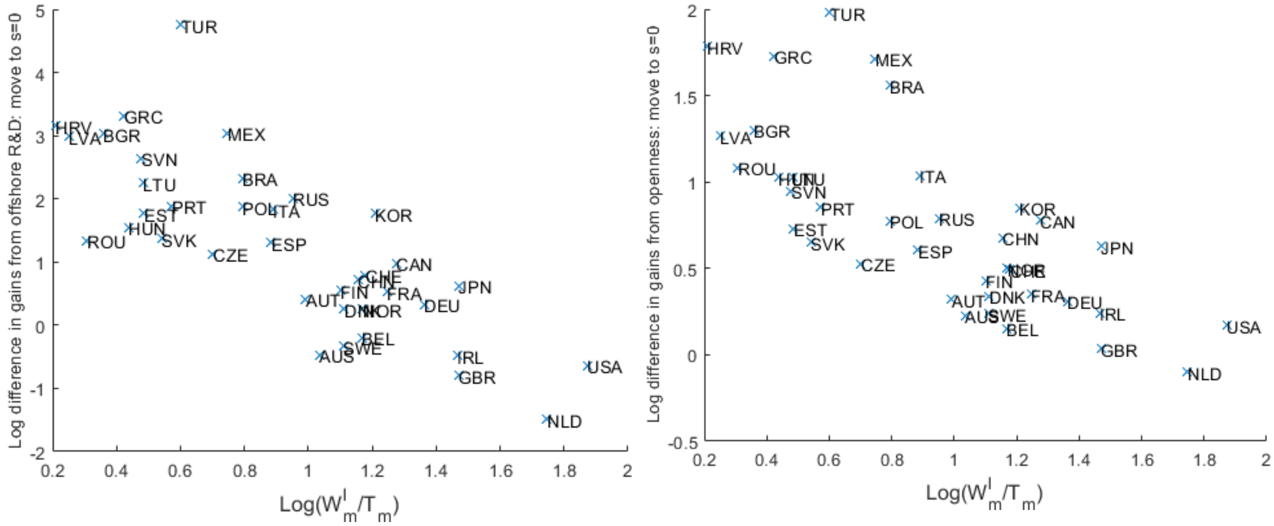
Impact on the gains from openness. To assess the importance of incorporating offshore R&D for inferring the gains from openness, I calculate the gains from openness in a restricted model without offshore R&D, calibrated to match the same trade and offshore production data as in the

channels both present. If the sum of individual channels is larger than the gains from openness, it means for at least one of the channels, the effect of integration is larger when the other two channels are already present.

Figure 8: Geographic Frictions and the Gains from R&D and Openness

(a) Setting $s = 0$ redistributes the gains from off. R&D

(b) Setting $s = 0$ amplifies the gains from openness



Notes: The horizontal axis in both panels is log of effective production cost in a country. The vertical axis is log of the ratio between the gains in the alternative model with $s = 0$ and the gains in the baseline model. The left panel focuses on the gains from offshore R&D; the right panel focuses on the gains from openness.

baseline equilibrium. Column 6 of Table 11 shows that this experiment generates 12.3% average gains from openness.

Comparison between Columns 5 and 6 suggests that offshore R&D amplifies the gains from openness by a factor of 1.3 (15.6/12.3) on average. This amplification differs significantly across countries. Advanced countries making substantial profit from offshore R&D see the biggest increase in inferred gains. For example, the inferred gains from openness of the U.S. almost double when offshore R&D is incorporated. For emerging countries, the amplification is generally smaller and could be slightly negative (e.g., Russia). The uneven changes between the baseline and restricted models further underscore the importance of incorporating offshore R&D—its omission not only underestimates the gains from openness, but also biases the comparison of the gains across countries.

The role of geographic frictions. In the model, the cost of offshore production is specified as a combination of distance to headquarters and distance to R&D centers. My estimates suggest $s = 0.82$ and implies on average 70% of R&D in offshore centers is devoted to localized production. Now I show that the value of s, which I discipline carefully using firm-level data, is crucial for correctly inferring the welfare gains.

To this end, I calculate the gains from R&D and openness in a re-calibrated model with $s = 0$, under which a higher fraction of offshore R&D is conducted for production at the headquarters. The last two columns of Table 11 report the results. The average gains from offshore R&D increase to 8.2% from the benchmark value of 3.3%, but the increases are concentrated in emerging countries. In fact, the gains from offshore R&D for the U.S. shrink by almost half.

The difference arises because the baseline model infers R&D as mostly for production in host countries, which competes with local firms for manufacturing workers and the product market, whereas the alternative model infers it as mostly for headquarter production, which competes with other firms from the home country. Because developing countries with lower production costs have a higher local production share of inward offshore R&D, moving to the alternative model with $s = 0$ leads to a more significant reduction in the competition faced by their domestic firms. This in turn results in a larger increase in the inferred gains from offshore R&D. The opposite is true for developed countries. Figure 8a plots the log change in gains from offshore R&D from the baseline to the alternative model with $s = 0$ and shows that indeed it is the countries with higher effective production costs that see the inferred gains from offshore R&D decreased.

Figure 8b shows that the negative slope continues to hold for the changes in the gains from openness from the baseline to the alternative model. Different from the gains from offshore R&D, however, the change are positive for almost all countries. This across-the-board increase in the inferred gains occurs because the alternative model implies a higher degree of integration through offshore production. Recall that in calibration I match by country the inward offshore production share, $\frac{\sum_{o,o\neq m} Y_{om}}{\sum_o Y_{om}}$, an important part of which is accounted for by varieties developed locally by foreign R&D centers. The alternative model, in which more offshore R&D is for production at the headquarters, matches the same $\frac{\sum_{o,o\neq m} Y_{om}}{\sum_o Y_{om}}$ by allowing for more offshore production, increasing the openness of the economy and hence generating higher gains for all.

Taking stock, these counterfactual experiments demonstrate that offshore R&D represents a new and quantitatively important channel through which countries benefit from globalization. It is a substitute for trade and offshore production for poor countries but a complement for advanced countries. Furthermore, the nature of offshore R&D—whether the inventions are devoted to production in the host or elsewhere—matters for the gains from offshore R&D and openness. This shows the importance of disciplining the model using firm-level data.

6.2 Implications for the Global Incidence of FDI and R&D Policies

Most existing quantitative research on MNCs does not differentiate policies on offshore R&D and offshore production. Yet policy makers usually have at their disposal instruments that specifically target each of these two activities. I examine whether not differentiating R&D and production is an important restriction for practical policy evaluations. As an example, I focus on two forms of FDI liberalization among emerging countries, which has gained significance in the past decade.

Integration among emerging countries. I first consider a reform eliminating the overhead cost for offshore R&D between a set of emerging countries, including Brazil, China, Hungary, Mexico, Poland, Russia, Romania, and Turkey. In practice, this reduction in cost can take the forms of speedier approval of entry, subsidized land, or tax credits for the upfront investment in R&D. The first column of Table 12 reports the results. Not surprisingly, this policy benefits emerging countries. Yet their benefits are at the expense of developed countries, whose overseas R&D centers in emerging countries have to face tougher competition after the policy.

The second experiment is liberalization in bilateral offshore production, which increases ϕ^{P}_{im} by 20% between the set of emerging countries. I focus on the different distributions of the welfare gains across countries, rather than the level of welfare gains, because these two types of liberaliza-

Table 12: Implications for FDI and R&D Policies

Country	FDI Integration Among Emerging Economies			Higher U.K. R&D Efficiency	
	off. R&D	off. Prod.	off. Prod. w/o off. R&D	Baseline	w/o off. R&D
	(1)	(2)	(3)	(4)	(5)
BRA	0.9	1.1	0.8	0.0	0.0
CHN	1.1	1.2	1.3	0.0	0.0
HUN	8.1	9.1	4.2	0.0	0.0
MEX	0.7	1.9	0.8	0.0	0.0
POL	1.6	7.9	3.9	0.1	0.0
ROU	10.0	19.8	10.3	0.1	0.1
RUS	2.5	4.7	2.7	0.0	0.0
TUR	1.0	1.5	0.9	0.0	0.0
DEU	-0.1	0.3	-0.1	0.0	0.0
FRA	-0.1	0.3	-0.1	0.0	0.0
GBR	0.0	0.3	-0.1	5.3	5.7
NLD	0.0	0.8	-0.2	0.4	0.2
JPN	-0.1	0.5	-0.2	-0.1	0.0
USA	-0.6	1.4	-0.3	0.1	-0.1
mean (all)	0.5	1.4	0.6	0.2	0.2

Notes: The first two columns show that offshore R&D and production policies between the same set of countries have qualitatively different incidences on the rest of countries. Comparison between Columns 2 and 3 shows that the same offshore production policy has different effects when offshore R&D is overlooked. The last two columns show that the spillover effects on the rest of the world of an increase in R&D efficiency in U.K. are different if offshore R&D is overlooked.

tion do not necessarily have the same administrative burden or fiscal costs. As shown in Column 2 of Table 12, emerging countries still gain significantly, but differently from the first experiment, major developed countries are also better off—thanks to their presence in the emerging countries through offshore R&D, countries like the U.S. benefit from an increase in the profit of the varieties they develop there. These two experiments demonstrate that openness to R&D and production could have qualitatively different third country effects. This point is important for multilateral investment treaties, which often cover investment in intellectual properties.

Because of the within-firm linkages between trade and production, and offshore R&D, incorporating the latter also affects policies on trade and offshore production. To make this point, I consider the same liberalization as in the second experiment, but in a restricted version of the model without offshore R&D. The welfare impacts of this experiment are reported in the third column of Table 12. Compared to the baseline economy, emerging countries generally benefit less—without foreign entrants, the overall R&D in these host does not expand as much to take advantage of the increasing access to overseas producers. Developed countries experience net losses: the production of their affiliates in emerging countries face an increase in competition as before, but now they cannot make up for the losses with the profit of the varieties they develop in these countries. The comparison between these two experiments shows that even if one's goal is solely to understand the effect of liberalizing offshore production, it is important to incorporate offshore R&D.

The global incidence of R&D policies. The presence of offshore R&D also implies that R&D policies can have a global impact simply because such policies would typically also apply to local R&D centers owned by foreign firms. This channel is independent of and in addition to the spillover effects of R&D across affiliates studied in the literature (e.g., Bilir and Morales, 2020). As an example, I consider a 20% increase in the efficiency of R&D taking place in the U.K.[29] This change increases the real income of the U.K. by 5.3%. Countries with extensive ties with the U.K. via offshore R&D are also better off. In total, 14% the total gains accrue to other countries. On the other hand, if offshore R&D is shut down, the same change in the R&D efficiency of the U.K. will benefit itself by 5.7%, which is 133% of the total gains—other countries, most notably the U.S., bear welfare losses.

7 Conclusion

Talented researchers and efficient firms are both necessary inputs to the development of new products, but they are distributed unevenly across countries. In a world separated by geographic frictions, MNCs organize their R&D and production to overcome this mismatch, in doing so integrating participants from different parts of the economy.

This paper develops and quantifies a model of firms' global R&D and production decisions, featuring 'talent-acquisition' and 'market-access' motives for offshore R&D. Quantitatively, offshore R&D brings about 3.3% welfare gains and amplifies the gains from openness by a factor of 1.3. These effects are especially large for developed countries, which derive a significant fraction of the value of their knowhow through offshore R&D. It thus has implications for measuring a country's intangible assets. Moreover, because of its interconnection with integration via trade and offshore production through both within-firm linkages and general equilibrium effects, understanding offshore R&D matters for these more familiar forms of globalization as well.

This paper abstracts from some aspects of the reality that might prove useful for measuring and theorizing about offshore R&D. For example, the model incorporates different tasks in bringing a product to consumers but has overlooked the role of sectors. Incorporating sectors and input-output linkages can shed light on the role of sectoral comparative advantage and its interaction with relative talent abundance. Second, constrained by the data, I have focused on offshore production within the boundary of firms. Enriching the model to accommodate outsourcing through arms' length transactions will paint a more complete picture of how offshore R&D affects country specialization and income.

References

Alviarez, Vanessa, "Multinational Production and Comparative Advantage," *Journal of International Economics*, 2019, *119*, 1–54.

Antràs, Pol and Alonso De Gortari, "On the geography of global value chains," *Econometrica*, 2020, *88* (4), 1553–1598.

[29] For foreign firms, this change resembles the 'patent box,' a policy implemented in the U.K. to attract innovative firms that reduces the corporate tax rate on revenues generated through R&D.

Antras, Pol, Luis Garicano, and Esteban Rossi-Hansberg, "Offshoring in a Knowledge Economy," *The Quarterly Journal of Economics*, 2006, *121* (1), 31–77.

Argente, David, Douglas Hanley, Salome Baslandze, Sara Moreira et al., "Patents to Products: Innovation and Firm Performance," *Unpublished Manuscript*, 2018.

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

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

Bloom, Nicholas, Christos Genakos, Raffaella Sadun, and John Van Reenen, "Management Practices Across Firms and Countries," *The Academy of Management Perspectives*, 2012, *26* (1), 12–33.

Burstein, Ariel T and Alexander Monge-Naranjo, "Foreign Know-how, Firm Control, and the Income of Developing Countries," *The Quarterly Journal of Economics*, 2009, p. 149.

Cravino, Javier and Andrei A Levchenko, "Multinational Firms and International Business Cycle Transmission," *The Quarterly Journal of Economics*, 2017, *132* (2), 921–962.

Fan, Jingting, "Essays on the Welfare Implications of International Economic Integration." PhD dissertation, University of Maryland 2017.

Fillat, José L and Stefania Garetto, "Risk, returns, and multinational production," *The Quarterly Journal of Economics*, 2015, *130* (4), 2027–2073.

Garetto, Stefania, "Input Sourcing and Multinational Production," *American Economic Journal: Macroeconomics*, 2013, *5* (2), 118–51.

_, Lindsay Oldenski, and Natalia Ramondo, "Multinational Expansion in Time and Space," Working Paper 25804, National Bureau of Economic Research May 2019.

Guvenen, Fatih, Jr. Mataloni Raymond J, Dylan G Rassier, and Kim J Ruhl, "Offshore Profit Shifting and Domestic Productivity Measurement," *NBER Working Paper 23324*, 2017.

Hall, Bronwyn H, "The internationalization of R&D," *Available at SSRN 2179941*, 2011.

Hanushek, Eric A and Ludger Woessmann, "Do Better Schools Lead to More Growth? Cognitive Skills, Economic Outcomes, and Causation," *Journal of Economic Growth*, 2012, *17* (4), 267–321.

Head, Keith and John Ries, "Increasing Returns Versus National Product Differentiation as an Explanation for the Pattern of US-Canada Trade," *American Economic Review*, 2001, *91* (4), 858–876.

_ and Thierry Mayer, "Brands in Motion: How Frictions Shape Multinational Production," *American Economic Review*, 2019, *109* (9), 3073–3124.

Helpman, Elhanan, "A Simple Theory of Ttrade with Multinational Corporations," *Journal of Political Economy*, 1984, *92* (3).

_, Marc J Melitz, and Stephen R Yeaple, "Export Versus FDI with Heterogeneous Firms," *American Economic Review*, 2004, *94* (1), 300–316.

Irarrazabal, Alfonso, Andreas Moxnes, and Luca David Opromolla, "The Margins of Multinational Production and the Role of Intrafirm Trade," *Journal of Political Economy*, 2013, *121* (1), 74–126.

Johnson, Robert C and Guillermo Noguera, "Accounting for Intermediates: Production Sharing and Trade in Value Added," *Journal of international Economics*, 2012, *86* (2), 224–236.

Keller, Wolfgang and Stephen Ross Yeaple, "The Gravity of Knowledge," *American Economic Review*, 2013, *103* (4), 1414–1444.

Koopman, Robert, Zhi Wang, and Shang-Jin Wei, "Tracing Value-added and Double Counting in Gross Exports," *American Economic Review*, 2014, *104* (2), 459–94.

Loecker, Jan De and Jan Eeckhout, "Global Market Power," Working Paper 24768, National Bureau of Economic Research June 2018.

Markusen, James R, "Multinationals, Multi-Plant Economies, and the Gains from Trade," *Journal of International Economics*, 1984, *16* (3), 205–226.

McGrattan, Ellen R and Edward C Prescott, "Openness, Technology Capital, and Development," *Journal of Economic Theory*, 2009, *144* (6), 2454–2476.

National Science Board, *Science and Engineering Indicators 2018*, Alexandria, VA: National Science Foundation, Available at https://www.nsf.gov/statistics/indicators/, 2018.

Nocke, Volker and Stephen Yeaple, "An Assignment Theory of Foreign Direct Investment," *The Review of Economic Studies*, 2008, *75* (2), 529–557.

OECD, *Multinational enterprises in the global economy: Heavily debated but hardly measured*, Retrieved at: https://www.oecd.org/industry/ind/MNEs-in-the-global-economy-policy-note.pdf, 2018.

Park, Walter G, "International Patent Protection: 1960–2005," *Research Policy*, 2008, *37* (4), 761–766.

Ramondo, Natalia and Andrés Rodríguez-Clare, "Trade, Multinational Production, and the Gains from Openness," *Journal of Political Economy*, 2013, *121* (2), 273–322.

Siedschlag, Iulia, Donal Smith, Camelia Turcu, and Xiaoheng Zhang, "What determines the location choice of R&D activities by multinational firms?," *Research Policy*, 2013, *42* (8), 1420–1430.

Swant, Marty, *The World's Most Valuable Brands*, Forbes Magazine, 2019.

Thursby, Marie and Jerry Thursby, *Here or There?: A Survey of Factors in Multinational R&D Location–Report to the Government-University-Industry Research Roundtable*, National Academies Press, 2006.

Tintelnot, Felix, "Global Production with Export Platforms," *The Quarterly Journal of Economics*, 2016.

Wang, Zi, "Headquarters Gravity: How Multinationals Shape International Trade," *Unpublished Manuscript*, 2019.

Appendix For Online Publication
Talent, Geography, and Offshore R&D

Jingting Fan
Pennsylvania State University

Contents

A Data and Facts

In this section I first explain in detail data sources, coverage, and cleaning procedures. I then validate patents as a measure of offshore R&D by showing that aggregate and bilateral offshore R&D statistics constructed from patents are in line with those constructed from R&D expenditures. I also show that the facts reported in the text are robust to alternative measures of production and R&D, the restriction to only the manufacturing industry, and an alternative IV strategy (for Fact 2). Finally, I present evidence in support of the assumption that R&D centers in different countries operate independently.

A.1 Data Extraction, Cleaning, and Coverage

The main source of the data for financial variables, ownership, and other firm characteristics is the Orbis Historic Disk Product, which links multiple vintages of Orbis products through firm identifiers and avoids some issues frequently arising in combining data vintages, as explained in Kalemli-Ozcan et al. (2019). The following describes my procedures for preparing the data, which largely follow those in Cravino and Levchenko (2017).

Financial data. I extract the financial data over 1996-2016 from the Historic Disk. The data are at firm identifier-year level. Each firm identifier represents a unique legal identity, possibly owned by another firm, an individual, or a family. I use sales (turnover) as the main measure of production and use value added for robustness. The initial extraction contains all firms with non-missing sales information for at least one year over the entire period. In a given year, firms might have multiple values of reported sales from different sources (local registry, annual report, or others), for consolidated or unconsolidated accounts. When different sources coexist, I take local registry over other sources as it is likely more accurate. Consolidated account might include sales of other firms belonging to the same conglomerate. For all analysis involving intensive margin measures of activities (i.e., sales and value added), I use only values reported in unconsolidated accounts and drop firms whose reporting is done solely in consolidated accounts; for analysis focusing on the extensive margin (i.e., whether an MNC has productive activities in a host), I keep the latter group of firms. The extracted data include 58 million unique firm identifiers and 196 million firm identifier-year observations, among which 168 millions have non-missing sales data from unconsolidated accounts.

Due to reporting lag and expanding coverage, the representation of the sample varies over time. Table A.1 reports aggregate statistics of the raw data for 2013.[1] Column 1 reports the ratio between the total sales of firms and the GDP of the country, which is well above 0.8 for most countries.[2] For empirical analysis, relatively low coverage in some countries does not pose a threat because systematic variations in sample representation will be absorbed by fixed effects; in quantification, however, I will need to calculate the overall foreign shares in R&D and production, in which case low coverage could lead to biases. I explain how foreign shares are calculated for these countries in Section C.1 of this appendix.

For a subset of countries, Eurostat provides total sales for 'total business economy; repairs of computers, personal and household goods; except for financial and insurance activities.' I aggregate total sales of firms in these industries. Column 2 shows that the sample representation is reasonably good in these countries. Column 3 shows similar levels of representation for manufacturing. In both columns, the ratio is above 1 for some countries. One reason for this is that only the total sales of firms are reported, so I treat all sales of a firm as from its reported core industry. To the extent that some manufacturing firms also generate revenue from finance, I will not be able to exclude such revenue in this calculation.

Columns 4 and 5 reproduce Columns 2 and 3, restricting to firms after the match with the patent data. The matching process will be explained below, but in short, a firm is in the patenting data if one of its affiliates, its parent, or its sibling affiliates within the same MNCs, has filed a patent in any country covered by the PATSTAT database. Given that most firms do not own patents, the post-match sample is much smaller, but as Columns 4 and 5 show, it still accounts for a substantial share of the economy.

Ownership data. I extract a snapshot of shareholder information from the Historic Disk.[3] For each firm ID, I identify its global ultimate owner (GUO), the entity holding a controlling share over the firm ID. This definition requires the owner to either directly hold more than 50% of the shares of the affiliate or—if the control is indirectly through other firms—hold more than 50% shares for every intermediate step along the ownership chain. This criterion ensures that the GUO has majority control over the affiliate. For firms that are not otherwise linked to a GUO, I assume that their GUO are themselves, which practically means they are all domestic firms. To the extent that firms are more likely to have unreported links to foreign firms than to domestic firms, my treatment underestimates the importance of MNCs.

[1] Except for Canada, which have a large number of missing value in 2013. I calculate statistics for Canada based on the 2014 data. Since in empirical analysis I will first average over each five-year interval, missing values in one year will not affect the measurement drastically.

[2] Aggregate sales could be higher than GDP because they count value added multiple times. The U.S. has a low coverage because most American firms report in consolidated accounts and are thus excluded from this calculation.

[3] I measure the ownership information in 2016 and assume that it does not change throughout the sample period. For regressions exploiting over-time variations, this measurement error, if any, likely attenuates results. Earlier vintages of the ownership data are in principle available, but with much more limited coverage.

Table A.1: Coverage of the Firm-Level Data

	Full sample			Sample with patents	
ISO	(1) $\frac{\text{total sales}}{\text{GDP}}$	(2) $\frac{\text{total sales exc. finance}}{\text{Eurostat total}}$	(3) $\frac{\text{mfg. sales}}{\text{Eurostat mfg.}}$	(4) $\frac{\text{total sales exc. finance}}{\text{Eurostat total}}$	(5) $\frac{\text{mfg. sales}}{\text{Eurostat mfg.}}$
AUS	2.97	-	-	-	-
AUT	4.67	0.80	0.73	0.45	0.55
BEL	2.50	0.72	0.68	0.38	0.48
BGR	1.51	1.07	0.81	0.23	0.23
BRA	0.23	-	-	-	-
CAN	1.49	-	-	-	-
CHE	2.93	0.39	0.44	0.03	0.08
CHN	0.99	-	-	-	-
CZE	2.19	0.96	0.90	0.42	0.59
DEU	1.95	0.71	0.63	0.36	0.45
DNK	2.00	0.61	0.59	0.31	0.49
ESP	1.53	0.92	1.03	0.43	0.64
EST	2.27	0.96	0.88	0.20	0.33
FIN	2.85	-	-	-	-
FRA	1.88	0.81	0.80	0.42	0.56
GBR	0.83	0.26	0.31	0.11	0.21
GRC	0.96	0.64	0.78	0.08	0.08
HRV	1.13	0.93	1.13	0.14	0.20
HUN	2.20	1.10	0.99	0.40	0.51
IRL	4.33	0.83	0.56	0.61	0.34
ITA	2.02	0.87	0.97	0.31	0.46
JPN	1.91	-	-	-	-
KOR	2.10	-	-	-	-
LTU	1.34	0.68	0.48	0.08	0.14
LVA	1.82	0.99	1.04	0.10	0.18
MEX	0.31	-	-	-	-
NLD	0.92	0.26	0.15	0.11	0.12
NOR	2.68	0.99	0.82	0.40	0.47
POL	1.04	0.74	0.93	0.30	0.55
PRT	1.60	0.92	0.79	0.30	0.27
ROU	0.90	0.98	0.95	0.32	0.44
RUS	1.62	-	-	-	-
SVK	2.59	1.17	0.95	0.46	0.63
SVN	2.00	0.82	0.77	0.25	0.36
SWE	3.05	0.89	0.76	0.42	0.59
TUR	0.32	-	0.41	-	0.13
USA	0.02	-	-	-	-
Average	1.83	0.81	0.75	0.29	0.37

Notes: This table reports aggregate statistics constructed from the firm-level data, divided by the corresponding official statistics. Columns 1 through 3 are for the full sample; Columns 4 and 5 are for firms in the patenting sample. 'Eurostat Total' refers to sales reported by Eurostat in 'total business economy; repairs of computers, personal and household goods; except financial and insurance activities.' This definition includes NACE sectors 05-63, 68-82, and 95. I calculate the sample counterpart of this statistics in Columns 2 and 4 by aggregating over firms whose core industry is in these sectors. 'Eurostat mfg.' refers to total manufacturing sales from Eurostat. I calculate the sample counterpart of this statistics in Columns 3 and 5 using only manufacturing firms. In columns 2 and 3, a couple of countries have ratios above 1. This is likely due to my treatment of the sales of multi-sector firms: only the total sales of a firm are reported, so I assume all sales is from the core industry.

Time-invariant firm characteristics. I define the home country and industry of an MNC to be the country and industry of the GUO, respectively. When the GUO is an individual or a family, in which case industry classification and country information are unavailable, I use instead the industry and location of the largest affiliate (by sales) within the MNC. Note that because Table A.1 reports the statistics at affiliate level, it is not impacted by this choice.

Patent data. I use patent-level data from PATSTAT Global to construct a measure of R&D. The database contains bibliographical data related to more than 100 million patent documents from 90 patent issuing authorities, including all major national, regional (e.g., the EPO), and global (e.g., the Patent Cooperative Treaty) patent offices. I link individual patents from this database to their assignees (their owners) using a crosswalk from the Orbis Intellectual Property Database. This crosswalk links patent applications to firms using a string matching algorithm based on the name, address, and industry classification of the owner. This process matches a total of around 25 million eventually granted patents world wide to $681,241$ unique firm identifiers from the Orbis registry.

I use the following procedures to further prepare the data.

1. **De-duplication.** Firms can and often do apply for multiple patents from different patent authorities for protection of the same underlying invention. Fortunately, all such patents need to establish a common priority, i.e., the first applied patent on the invention, and can therefore be identified as belonging to the same patent family in the PATSTAT database.[4] I keep a family as long as one of its many patents is linked to a firm ID and, within each family, keep only one patent—the one with the most complete inventor location information. This de-duplication process reduces the number of unique patents to around 17 million, about two-thirds of the original number.

2. **Excluding design patents.** I exclude design patents and patents with unidentified types. Together, the excluded patents account for about 2% of the sample. The resulting sample contains patents from 90 patent offices, with the top 10 biggest patent offices accounting for 90% of the observations. The USPTO patents account for about 20% of this sample.

3. **Excluding patents without inventor location information.** PATSTAT does not receive inventor location information from the Japanese Patent office (JPO). Patents from other offices sometimes also have missing inventor locations. I exclude patents from the JPO or otherwise have missing inventor location information.[5]

   Table A.2 summarizes the contribution of each patent office to the sample and the fraction of these patents with non-missing inventor location information. Columns 1 and 2 are for all patents that can be matched to a firm in the Orbis data (after steps 1 and 2 described above), dating back to the early 20th century. U.S. and China are two biggest patent offices in this period, followed by Germany, Korea, and the EPO. Column 2 reports the fraction of observations from each patent office with non-missing inventor location. For six out of ten top patent authorities, inventor location is available for more than 70% of patents.[6]

   Columns 3 and 4 reproduce Columns 1 and 2 for the period of my empirical analysis, 1996-2016. With increasing patenting in China, the top 10 offices now account for 94% of the sample. Aside from China, Australia, and Canada, all other major patent offices have close to universal availability of inventor location information. The increase in the availability of location information for

[4] A patent family is a collection of patents from different countries protecting the same invention.

[5] Note that such exclusion does not necessarily mean that the R&D underlying the patent is excluded from my database. As long as one patent within a family has inventor location information, it will be preserved. This is an advantage of using the full PATSTAT Global data—I am able to piece together information about an invention from its multiple patents. For example, if a Japanese firm filed a patent in the U.S. and Japan at the same time, yet only the U.S. application reports inventor locations and only the Japanese application is linked to a firm ID, this patent will still be in my sample.

[6] Note that after the de-duplication procedure in Step 1, both the distribution of patents by office and the share of patents from an office with inventor location information will be different from those in the raw data, because only one of each patent family is kept and the one being kept tend to be from offices with more complete inventor information.

Table A.2: Sample Size and Availability of Inventor Location by Patenting Authority

	All historic patents		Patents filed in 1996-2016	
Patent office	% of obs. (1)	% with location (2)	% of obs (3)	% with location (4)
USA	23.24	71.72	17.03	100.00
CHN	22.14	21.53	32.18	20.61
GER	8.52	59.61	4.22	99.97
KOR	8.33	93.67	10.12	93.39
EPO	6.48	99.72	8.38	99.75
CAN	6.01	29.28	5.37	21.80
PCT	5.18	94.19	7.46	94.64
AUS	4.06	1.01	3.13	0.74
AUT	2.95	70.34	1.81	98.42
TWN	2.89	99.99	4.19	99.99
All others	10.20	56.52	6.10	75.11
Total	100.00	58.19	100.00	64.50

Notes: Columns 1 and 2 report information on all sample patents after Step 1 and 2 of the cleaning process. Column 1 tabulates the fraction of patents in this sample from different patent offices. Column 2 reports, among all patents from an office, the fraction with inventor location available. Columns 3 and 4 reproduce Columns 1 and 2 for patents filed first between 1996 and 2016, the period the empirical analysis focuses on.

most countries from Column 2 to Column 4 is likely due to changing reporting requirements at the patent application stage. That the missing information is concentrated in a small number of countries also reassures that these missing values are due to country-specific requirements, rather than MNCs' self selection into reporting.[7]

4. **Aggregating by firm ID-inventor country-year.** I define the invention time of a patent as the earliest filing year among all patents within the patent family. I then sum across all patents assigned to a firm identifier in a given year to arrive at patent counts by firm identifier-inventor country-year. Note that because MNCs can assign a patent to any of its affiliates regardless of where the invention is performed and which patent office is involved, the result of this aggregation is not necessarily accurate for locations of R&D at the affiliate (firm identifier) level.[8] But after the final step below, it will be accurate for location of R&D at the parent level.

5. **Aggregating to parent firm-inventor country-year level.** I aggregate the R&D output from the previous step to parent firm-inventor country- year. I interpret inventor countries as the location of R&D. For example, if an American firm has 30 patents with inventors located in Japan, I interpret this as output of the American R&D center in Japan. I wish to emphasize that this assignment has nothing to do with whether the patents are from the USPTO or JPO, or whether the assignee on the patent is a U.S. affiliate in Japan or the headquarters in the U.S. The inferred location of R&D depends solely on the reported addresses of the inventors.

A.2 Concerns on Patents as a Measure of Offshore R&D and Validation

As discussed in Section 2.1, using patent data to measure offshore R&D has three advantages: first, the universe of data are readily available at the firm level; second, it is less subject to different definitions of 'research and development' between advanced countries that are pushing the frontier and developing countries that are simply trying to learn what is already known; third, compared to affiliate R&D

[7] The empirical patterns stay virtually the same if the data from China, Australia, and Canada are excluded.

[8] For example, Apple can apply for a patent invented entirely in California through its affiliate in China. It would be wrong to infer from this assignment that Apple China performs R&D in California.

expenditures, the addresses of inventors are less likely to be manipulated by MNCs for tax avoidance.[9]

There might be two concerns on using inventor locations from patents to measure R&D. The first is that firms' self selection into patenting can introduce biases. The second is that when an inventor move across countries, the reported address of a patent might not always accurately reflect where the R&D is carried out. I discuss these two concerns in detail and explain why they are unlikely to bias either the reduced-form or quantitative results.

Self-selection into patenting. Consider the first concern. There are at least two types of selection, both of which are well recognized by the study of R&D using patent data in closed-economy settings (see Griliches, 1998 for an early survey): some R&D efforts might not result in patentable outcomes; firms might choose not to patent a patentable R&D outcome. In the multinational setting considered in this paper, the threat is that such selection might be correlated with the characteristics of the firm or those of the host. For example, if more innovative firms are both more likely to engage in offshore R&D and have a higher propensity to patent their inventions, my measure of offshore R&D would be biased towards these firms.

An advantage of the multinational setting is that I will be able to flexibly control for firm and host characteristics that likely determine the decision to patent an invention in ways that are infeasible in closed-economy settings. Specifically, I control for firm-period fixed effects in all specifications, so any selection at the firm level is absorbed and the coefficients are only identified off within-firm variations. For Facts 1 and 2, where the selection concern is more relevant, I further control for affiliate fixed effects. This will purge out the influence of any time-invariant host-specific factors.[10]

In addition to these two types of selection, in my setting there is another type of selection: firms choose to patent inventions only in hosts in which they either have a manufacturing presence or the intention to launch a product. The idea is, in hosts where firms produce and sell a product, both the likelihood of and potential damage from IPR infringement by local competitors are higher, so they have a stronger incentive to patent. This makes the interpretation of Fact 2 problematic: the observed colocation between production and patenting might have nothing to do with the friction in separating the two. Because my measure of R&D is based *not* on which host country a patent is issued in, but on where the inventor of a patent is located, this concern does not in itself lead to a bias. To the extent it is possible that such selection affects the measured locations of inventors, this concern is addressed in two ways. First, the rich set of fixed effects will absorb all confounding factors that are specific to a host, an affiliate, or a firm. Second, I provide direct evidence that such selection does not affect the measured offshore R&D below.

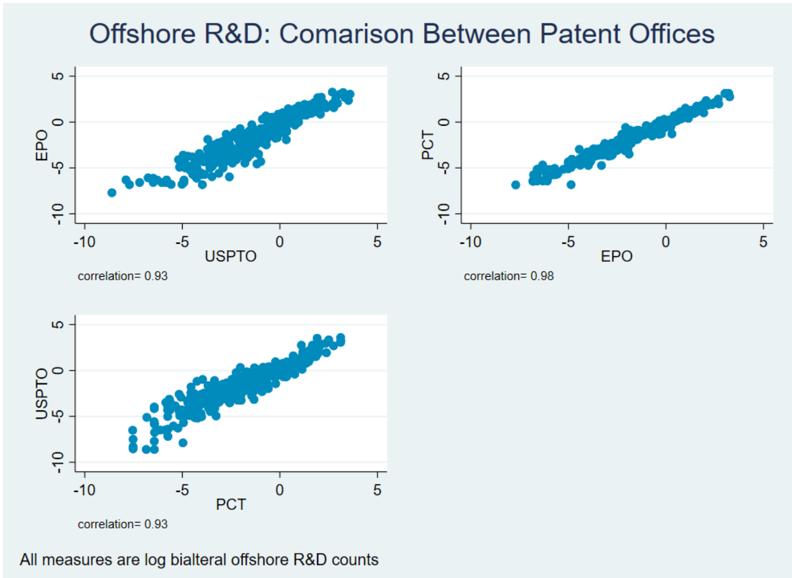
Specifically, if the selection of patent offices affect the measured offshore R&D, then we should see that offshore R&D calculated based on data from different patent authorities differ significantly from one another. Figure A.1 shows the exact opposite: bilateral offshore R&D shares measured using data from the USPTO and two other major international patent offices, the PCT and the EPO, are highly correlated. I conclude that differential selection into patenting in specific hosts do not lead to significant biases in the measured offshore R&D.

In quantitative analysis, I will use the share of inventions in a host by foreign firms as an additional input. For such aggregate shares, I will not be able to address the above concern through controls. Instead, I show directly that my measure is closely correlated with the one based on R&D expenditures in Figure A.2. In the left panel, the horizontal axis is calculated using business enterprise R&D expenditures from the OECD; the vertical axis is the shares calculated based on the patent data, described above. The figure shows a close mapping between the two measures, with a correlation of 0.83, despite that they

[9]Indeed, while official statistics show large FDI flows and significant MNC presence in countries like Bermuda, Panama, and Cayman Islands, my measure shows very few patents are invented in these locations.

[10]Admittedly, there could be time-varying factors. In Fact 1, I control for two usual suspects, the IPR protection index and R&D subsidies and show that they do not affect the main coefficient of interest. In Fact 2, since the variation exploited is at affiliate level, I am able to control for host-industry-period fixed effects, which absorb all time-varying characteristics of a country that might affect patenting in an industry.

Figure A.1: Comparison Among Patents From Individual Patent Offices

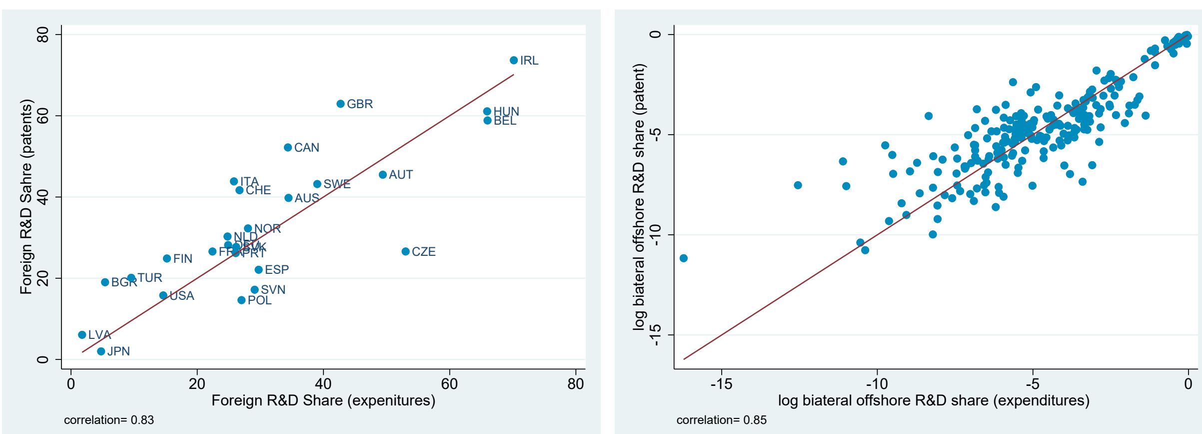


Notes: The figure shows log bilateral offshore R&D measured using three major patent offices (EPO, USPTO, PCT) is closely correlated.

Figure A.2: Aggregate and Bilateral Offshore R&D Measures: Patents v.s. Expenditures

(a) Aggregate Inward Offshore R&D Share

(b) Log Bilateral Offshore R&D Share



Notes: The left panel is the fraction of R&D in a country carried out by foreign firms, measured using two different sources: business enterprise R&D expenditures (horizontal axis) and patents (vertical axis). The right panel plots the log of bilateral offshore R&D shares, in which each dot represents a country pair.

are from two independent sources.[11] The right panel plots log *bilateral* offshore R&D shares measured using R&D expenditures and patents. Again, the two measures are highly correlated. The takeaway from Figure A.2 is that even if one preferred to use R&D expenditures to construct foreign R&D shares, my measure is a good proxy.

Measurement errors due to inventor mobility. Compared to the general population, inventors tend to be more mobile geographically. The second concern on measurement is thus, if, after spending years on a project, an inventor moves to a new country right before the patent application form is submitted,

[11] The expenditures statistics are aggregated from firm-level data from surveys and other administrative sources, and are not publicly available. It is possible that the differences between the two measures are mostly driven by sampling differences. In fact, the correlation between my data and these statistics for offshore production, measured using sales in both sources, is also around 0.8.

then the inventor address would differ from where the underlying R&D was conducted.

The prevalence of such measurement errors is bounded by the propensity of inventors to move across borders *after* becoming an inventor. Recently, Akcigit et al. (2016) shows (in Table 1) that among inventors of USPTO patents, the share of top 5% inventors who have migrated is 3.6%; the share of bottom 95% inventors who have migrated is much lower at 0.7%.[12] Top 5% inventors account for about 30% of total patents. Even if the R&D location of all lifetime patents of those who have moved across borders as an inventor is misclassified, it would amounts to only around 1.6% ($30\% \times 3.6\% + (1 - 30\%) \times 0.7\%$) of all patents. Similar calculation based on patents at the EPO using the numbers reported in Online Appendix Table A11 of Akcigit et al. (2016) finds a smaller number of 1%.

These numbers likely still overstate the extent of measurement errors. One way of seeing this is: assuming a project takes 3 years, then the measurement error would only affect projects initiated during the three-year window prior to the move. In the data, the average duration between first and last patent among all inventors is 12 years, so the three-year window contains a quarter of these inventors' lifetime patents on average, i.e., if all inventions made during this window has wrong inventor addresses, it would amount to $1.6\% \times 25\% \approx 0.4\%$ of all patents. The working career of top inventors is likely longer than average; for them, the three-year window accounts for an even smaller share of their total patents.

To conclude, this back-of-envelop calculation suggests that under reasonable assumptions, misclassified R&D locations due to inventor migration is unlikely to be important.

A.3 Sample Restriction and Descriptive Statistics

Having established the validity of my offshore R&D measure, I describe how the sample is chosen and present descriptive statistics.

Sample period. For best coverage of the financial data, I focus on 1996-2016. To reduce measurement errors associated with patent counts (for example, firms might be continuously doing R&D but the patent application might be discrete), I aggregate the sample into four five-year periods. Within each period, I take the average values of patent and citation counts, and financial statistics.

Countries and their characteristics. I focus on a sample of 37 host countries (but include MNCs whose parent are from other countries in empirical analysis). This sample restriction is made to be consistent with the subsequent quantitative analysis. Specifically, for quantification I will use data on manufacturing output and trade from the 2016 release of the World Intput Output Database. Among the 43 countries in this database, I exclude three countries with population below one million, Cypress, Luxembourg, Malta; I exclude Taiwan, as World Bank and Penn World Table does not report its economic statistics; finally I exclude India and Indonesia due to their poor representation in the Orbis financial database. This results in the 37 countries reported in Table A.1. All empirical patterns remain virtually unchanged if I simply use all countries in the Orbis database.

I combine the firm-level data described previously with time-varying country characteristics. Concretely, I obtain GDP, GDP per capita, and the human capital index from the PWT 9.0.; an updated version of the intellectual property right index from Park (2008); R&D subsidies and the number of researchers from the OECD. All these variables are also averaged over each four-year period. I obtain bilateral distance measures from Mayer and Zignago (2011). Table A.3 summarizes these country characteristics for the last period, 2011-2016.

[12] These numbers are based on disambiguated inventor names from patent records. For an inventor to appear as a mover in this dataset, he or she needs to have filed a patent from one country and then file another patent later from a different country. This definition of mobility excludes people who migrate as a child or a postgraduate student and become inventors only after moving to the new country. However, it is exactly the right measure to use for the specific concern on measurement errors in R&D locations due to the mobility of inventors.

Table A.3: Host Characteristics: Summary Statistics

Variable	Obs	Mean	Std. Dev.	Min	Max
ln (GDP)	37	13.30	1.55	10.37	16.59
ln (GDP per capita)	37	13.30	1.55	10.37	16.59
Human capital idnex	37	3.24	0.38	2.29	3.72
ln (number of researchers)	32	11.38	1.48	8.61	15.07
R&D subsidies	36	0.13	0.12	-0.02	0.44
Intellectual property right index	33	4.31	0.33	3.59	4.88

Notes: This table reports characteristics of the 37 host countries in the sample, averaged over 2011-2016.

Structure of financial and R&D samples separately. In empirical analysis, I use the merged sample between financial and R&D datasets. Section 2.2 presents descriptive data for the matched sample. Tables A.4 gives an overview of these two datasets separately.

Table A.4: Structures of Production and R&D Samples

	Production data		R&D Data		
Period	# of unique firms	# of production facility.	# of unique firms	# of R&D centers (baseline)	# of R&D centers (Liberal)
1	3,615,341	3,643,392	118,951	112,215	140,341
2	7,992,947	8,050,202	138,790	133,430	164,171
3	13,814,682	13,906,270	161,111	158,294	191,439
4	49,264,872	49,389,390	136,321	136,692	162,137
Total	74,687,842	74,989,254	555,173	540,631	658,088
Unique firms	54,535,654		378,859		

Notes: This table summarizes separately the coverage of financial and R&D data over time (before the two are merged).

The left panel of Table A.4 summarizes the structure of the financial data. Columns 1 and 2 are the number of unique firms and unique production affiliates in each period, respectively. The numbers gradually increase as coverage of the database broadens, but Columns 1 and 2 track each other closely, reflecting that the overwhelming majority of firms have only one production affiliate.

The right panel of Table A.4 is the structure of the R&D data. This sample is larger than reported in Table 1 as it includes firms granted a patent but with no available financial information.[13] The R&D sample does not registered as dramatic an expansion as the financial sample. This is unsurprising, as PATSTAT covers close to the universe of world patents from the very beginning. In the fourth period, on average, a firm has 1.189 R&D centers according to the liberal definition, and only 1.003 according to the baseline definition.

A.4 Robustness: Alternative Measures and Sample Restriction

This subsection shows that the facts in the text are robust to alternative measures and sample restrictions.

Measure of R&D. In the baseline analysis I measure the intensive margin of R&D by counting the number of patents, weighted by the number of inventors on a patent. According to this measure, if a patent has multiple inventors located in more than one country, I assign each country the fraction of inventors residing in it. This is of course arbitrary. As an alternative, I count each county as having the full patent—for example, if a patent is invented jointly by one person in the U.S. headquarters and one person in Canada, I count the Canadian affiliate and the U.S. headquarters as each having invented a full patent. In what follows I will call this unweighed patent counts.

Second, it is well known that patents differ vastly in their values and that the number of forward citations to a patent is a good proxy for its value—much like the number of forward citations to an

[13] The statistics reported is after excluding firms classified as education institutions and governments, or firms with unknown home countries—most likely individuals or families.

academic article often indicates its importance. To adjust for patent quality, I use the number of citations received by the patents invented at an affiliate as a measure for the invention output of that affiliate.

Finally, in the baseline analysis involving the extensive margin of offshore invention, I define a host country as an R&D center if it has invented at least one full patent. This restriction rules out firms with only a small number of patents collaborated with inventors located in the headquarters. For robustness, I use a more liberal definition, according to which a host country is classified as hosting an R&D center as long as a positive fraction of a patent is invented in it.

Measure of production. In the baseline analysis, I measure affiliate production by sales. Some of the sales are likely due to intermediate products made elsewhere. I use value added for robustness.

Restriction to manufacturing. The baseline analysis includes firms from all industries. Since the quantitative model has a focus on trade, it is more appropriate to interpret it as for manufacturing. I show all baseline estimates remain materially the same when I restrict to manufacturing firms.

Excluding headquarters from regressions. For Fact 1, the baseline regressions include observations that are in the headquarters country of a firm. One might be concerned that the systematic concentration of R&D in countries with talented workers might be driven by headquarters only. Note that this interpretation does not alter the main message that the invention intensity of a firm differs across countries with different talent endowments. Nevertheless, I show that results are similar if headquarters are excluded.

I now explain the results from these robustness exercises.

Table A.5: Fact 1 Robustness

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Dependent var.	$\ln(\frac{\text{patent}}{\text{sales}})$	$\ln(\frac{\text{patent}}{\text{sales}})$	$\ln(\frac{\text{unwgt. patent}}{\text{sales}})$	$\ln(\frac{\text{citation}}{\text{sales}})$	$\ln(\frac{\text{patent}}{\text{VA}})$	R&D Ind. (baseline)		R&D Ind. (liberal)
human capital index	3.705**	3.613**	2.910**	4.185**	3.836***	0.213*	0.189**	0.240**
	(1.547)	(1.383)	(1.364)	(1.706)	(0.818)	(0.110)	(0.092)	(0.100)
ln(GDP per capita)	-0.381	-0.665*	-0.881***	-0.770**	-0.718	0.102***	0.090***	0.085***
	(0.389)	(0.369)	(0.319)	(0.355)	(0.447)	(0.028)	(0.027)	(0.028)
IPR protection	0.309	0.393**	0.383**	0.661***	0.027	0.029	0.024	0.028*
	(0.234)	(0.182)	(0.148)	(0.209)	(0.189)	(0.022)	(0.017)	(0.015)
R&D subsidies	0.695	0.577	0.380	0.261	0.720*	0.036	0.016	0.012
	(0.451)	(0.434)	(0.424)	(0.537)	(0.385)	(0.036)	(0.030)	(0.027)
ln (researchers)	0.249	0.416**	0.402**	0.214	0.217	0.071***	0.064***	0.073***
	(0.182)	(0.176)	(0.162)	(0.181)	(0.290)	(0.022)	(0.019)	(0.020)
log (sales)						0.006***	0.003**	0.004**
						(0.002)	(0.001)	(0.002)
Observations	7533	9672	11464	11464	7585	41396	71226	80253
R^2	0.668	0.679	0.688	0.703	0.653	0.618	0.562	0.603
Within R^2	0.014	0.016	0.013	0.016	0.019	0.008	0.006	0.005
Firm-period FE	Y	Y	Y	Y	Y	Y	Y	Y
Affiliate FE	Y	Y	Y	Y	Y	Y	Y	Y
Sample	mfg. only	excl. HQ	baseline	baseline	baseline	mfg. only	excl. HQ	baseline

Note: This table reports robustness results for the last two columns of Table 3. The dependent variable in Columns 1-5 are the intensive margin invention intensity at an affiliate. Columns 1-2 use the baseline measure, but restrict the sample. Columns 3-5 use different measures for either invention or sales in measuring invention intensity. The dependent variable in Columns 6-8 is the indicator for R&D centers. Columns 6-7 use the baseline measure and restrict the sample; Column 8 uses the liberal definition of R&D centers. Standard errors (in parenthesis) are clustered two way, by firm and by host country. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

Fact 1 robustness. Table A.5 reports the robustness results for Fact 1. I take the last two columns of Table 3—which controls for affiliate fixed effects and firm-period fixed effects—as the baseline for intensive and extensive margin regressions, respectively. Columns 1 to 5 focus on the intensive margin measure of invention intensity. Column 1 restricts the sample to manufacturing only; Column 2 excludes headquarters from the sample; Columns 3 to 5 permute on the dependent variables, the log ratio between invention and production, by changing the measure for either invention output or production.

Throughout, the human capital index is statistically significant and economically sizable, although the number of narrowly defined researchers are not statistically significant in some specifications.

Columns 6 to 8 are robustness for when the dependent variable is an R&D indicator. Column 6 includes only manufacturing firms; Column 7 excludes headquarters; Column 8 uses the liberal definition of R&D centers. Results from all three specifications are qualitatively similar to the baseline.

Fact 2 robustness. Tables A.6 and A.7 report additional results for Fact 2. The specifications in Table A.6 reproduce Columns 1, 4, 5, and 7 of Table 4, using the same measures for R&D and production but restricting the sample to manufacturing firms. Across specifications, the point estimates are generally close to that from the baseline sample.

Table A.7 reports robustness with alternative definitions of invention and production. Columns 1 and 2 show that having an liberally defined R&D center is associated with both the presence and the size of production facilities. The coefficients are close to the baseline estimates (Columns 1 and 4 of Table 4, respectively). Columns 3 through 6 reproduce Columns 5 and 7 of Table 4 using citations and unweighted patent counts to measure the intensive margin of invention, respectively. Finally, Columns 7 to 9 keep the baseline measure of invention but change the dependent variable to value added, which reduces the sample size substantially. The specifications correspond to Columns 4, 5, 7 of Table 4.

Exercises here show that Fact 2 holds across different measures and sub-samples. In Appendix A.5, I show that an alternative identification strategy using the variation in host R&D subsidies and number of researchers lead to qualitatively similar findings.

Table A.6: Fact 2 Robustness: Manufacturing Only

	(1)	(2)	(3)	(4)
Dependent variable	prod. indicator		log (sales)	
R&D Indicator$_{fh,t}$	0.295***	1.042***		
	(0.004)	(0.033)		
ln(patent)$_{fh,t}$			0.324***	0.181***
			(0.014)	(0.046)
ln (distance)$_{fh,t}$		-0.012	-0.395**	
		(0.035)	(0.169)	
common language$_{fh,t}$		0.207***	0.284	
		(0.077)	(0.328)	
contiguity$_{fh,t}$		0.201***	0.227	
		(0.077)	(0.285)	
colonial tie$_{fh,t}$		-0.045	-0.713*	
		(0.070)	(0.382)	
Observations	4156173	61474	13072	6417
R^2	0.735	0.512	0.574	0.969
Within R^2	0.050	0.052	0.099	0.020
Firm-period FE	Y	Y	Y	Y
Host-period FE	Y	Y	Y	-
Home-host FE	Y	Y	Y	-
Host-industry FE	Y	-	-	-
Host-industry-period FE	-	-	-	Y
Affiliate FE	-	-	-	Y

Note: This table reports robustness of Fact 2 using only manufacturing firms. Column 1 corresponds to Column 1 of Table 4; Columns 2 through 4 correspond to Columns 4, 5 and 7 of Table 4. Standard errors (in parenthesis) are clustered by firm. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

Table A.7: Fact 2 Robustness: Alternative Measures for Invention and Production

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
Dependent variable:	prod. indicator			ln(sales)				ln(value added)	
R&D Ind.$_{fh,t}$: liberal	0.275***	0.988***							
	(0.003)	(0.026)							
ln(citation)$_{fh,t}$			0.277***	0.137***					
			(0.012)	(0.040)					
ln(unwgt. patent)$_{fh,t}$					0.382***	0.194***			
					(0.015)	(0.046)			
R&D Ind.$_{fh,t}$: baseline							1.036***		
							(0.029)		
ln(patent)$_{fh,t}$								0.338***	0.107*
								(0.014)	(0.058)
ln (distance)$_{fh,t}$		-0.049*	-0.333**		-0.305**		-0.061*	-0.600***	
		(0.025)	(0.146)		(0.145)		(0.037)	(0.229)	
common language$_{fh,t}$		0.167***	0.513*		0.358		0.234***	0.827**	
		(0.051)	(0.272)		(0.268)		(0.061)	(0.330)	
contiguity$_{fh,t}$		0.096*	0.310		0.187		0.109**	-0.239	
		(0.049)	(0.240)		(0.237)		(0.054)	(0.258)	
colonial tie$_{fh,t}$		0.017	-0.657**		-0.552*		-0.051	-0.554	
		(0.046)	(0.314)		(0.307)		(0.056)	(0.363)	
Observations	7494979	119503	19519	8839	19519	8839	70184	12342	5644
R^2	0.705	0.495	0.559	0.963	0.568	0.963	0.554	0.601	0.930
Within R^2	0.048	0.045	0.066	0.015	0.086	0.020	0.065	0.132	0.005
Firm-period FE	Y	Y	Y	Y	Y	Y	Y	Y	Y
Host-period FE	Y	Y	Y	-	Y	-	Y	Y	-
Home-host FE	Y	Y	Y	-	-	-	Y	Y	-
Host-industry FE	Y	Y	Y	-	-	-	Y	Y	-
Host-industry-period FE	-	-	-	Y	-	Y	-	-	Y
Affiliate FE	-	-	-	Y	-	Y	-	-	Y

Note: This table reports robustness results for Fact 2 using alternative measures of invention and production. Columns 1 and 2 reproduce Columns 1 and 4 of Table 4 using the liberal definition of R&D center; Column 7 reproduces Column 4 of Table 4 using log value added as a measure for production. Columns 3 to 6 reproduces Columns 5 and 7 of Table 4 using alternative intensive measures of invention. Columns 8 and 9 reproduce the same two columns using log value added to measure production. Standard errors (in parenthesis) are clustered by firm. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

Fact 3 robustness. Tables A.8 and A.9 report additional robustness exercises for Fact 3. Tables A.8 uses the baseline measures but restricts the sample to manufacturing. The coefficients are generally similar to the baseline estimates.

Table A.9 uses the same sample as the baseline and changes the measure of invention and production. Columns 1 to 3 show the headquarter effect for invention is robust to the liberal definition of R&D centers and two different measures of invention. Most coefficients are broadly in line with the corresponding ones (Columns 1 and 2) in Table 4. Columns 4 to 6, corresponding to Columns 4 and 6 of Table 5, show the headquarter effect for production are robust when these alternative measures of R&D are used as controls. Finally, Column 7 shows that the result is similar when production is measured using value added.

Table A.8: Fact 3 Robustness: Manufacturing Only

	(1) Affiliate R&D	(2)	(3) Affiliate Production	(4)
	indicator	ln(patent)	indicator	ln (sales)
ln(distance)$_{oh}$	-0.001	-0.140***	-0.004***	-0.230***
	(0.001)	(0.040)	(0.002)	(0.028)
common language$_{oh}$	0.022***	0.226**	0.011	0.005
	(0.006)	(0.093)	(0.009)	(0.066)
contiguity$_{oh}$	0.001	0.048	-0.001	0.137**
	(0.002)	(0.101)	(0.004)	(0.066)
colonial tie	0.006	0.094	0.029***	0.102
	(0.005)	(0.080)	(0.009)	(0.068)
R&D indicator$_{fh,t}$			0.379***	1.181***
			(0.022)	(0.037)
Observations	4045403	28244	4045403	54208
R^2	0.149	0.297	0.368	0.465
Within R^2	0.004	0.010	0.069	0.063
Firm-period FE	Y	Y	Y	Y
Host-industry FE	Y	Y	Y	Y
Host-period FE	Y	Y	Y	Y

Note: This table reproduces Columns 1-2 and 5-6 of Table 5 for manufacturing firms. Measures of R&D and production are the same as in the baseline. Standard errors (in parenthesis) are clustered by country pair. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

Table A.9: Fact 3: Alternative Measures

	(1) Headquarter Effect for R&D	(2)	(3)	(4) Headquarter Effect for Production	(5)	(6)	(7)
Dependent var.	R&D ind. (liberal)	ln(unwgt. patent)	ln(citation)	ln(sales)			ln(VA)
ln(distance)$_{oh}$	-0.003***	-0.119***	-0.078**	-0.247***	-0.102**	-0.131**	-0.189***
	(0.001)	(0.028)	(0.031)	(0.027)	(0.047)	(0.051)	(0.045)
common language$_{oh}$	0.030***	0.222***	0.224***	0.086	0.059	0.076	0.025
	(0.006)	(0.063)	(0.070)	(0.060)	(0.082)	(0.084)	(0.070)
contiguity$_{oh}$	0.005*	0.109*	0.183***	0.175***	-0.005	0.009	0.208***
	(0.002)	(0.062)	(0.065)	(0.057)	(0.093)	(0.098)	(0.066)
colonial tie$_{oh}$	0.001	0.025	-0.015	0.123*	0.046	0.060	0.156**
	(0.005)	(0.055)	(0.067)	(0.068)	(0.079)	(0.083)	(0.074)
R&D ind. (liberal)$_{fh,t}$				1.096***			
				(0.028)			
ln(unwgt. patent)$_{fh,t}$					0.392***		
					(0.018)		
ln(citation)$_{fh,t}$						0.284***	
						(0.015)	
R&D ind. (baseline)$_{fh,t}$							1.185***
							(0.030)
Observations	7295102	45364	45364	103131	16189	16189	60553
R^2	0.155	0.324	0.455	0.445	0.502	0.490	0.504
Within R^2	0.006	0.016	0.010	0.054	0.090	0.069	0.079
Firm-period FE	Y	Y	Y	Y	Y	Y	Y
Host-industry FE	Y	Y	Y	Y	Y	Y	Y
Host-period FE	Y	Y	Y	Y	Y	Y	Y

Note: This table reports the robustness results of Fact 3 to alternative measures. Columns 1 to 3 replicate Columns 1 and 2 of Table 5 with different measures of R&D. Columns 4 to 6 replicate Column 6 of Table 5 with different measures of R&D. Column 7 replicate Column 6 of Table 5 using value added as a measure for production. Standard errors (clustered at country-pair level) in parenthesis. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

A.5 IV Estimates for Fact 2

The second fact, the colocation of invention and production, plays an important role in quantitative analysis as it pins down the market access motive. The most demanding specification for this fact controls for firm-period, host-industry-period, and affiliate fixed effects, which rules out the following confounding factors: shocks to host countries or all affiliates of a firm that drive the colocation; idiosyncratic match quality between a firm and a host, which encourages both invention and production; changes in the comparative advantage of a country that affect the entire industry. A remaining source of threat is time changes in the idiosyncratic match quality between firms and hosts. Such changes need to be firm specific—otherwise it will be absorbed by host-industry-period fixed effects. One example of such shocks is development of a new technology in a country that is useful to only a few firms within an industry, but affects both production and invention of these firms directly.

To address this concern, I use an alternative identification strategy. Under the assumption that controlling for other time-varying country characteristics, changes in a host's R&D environment affect affiliate production only through affiliate R&D, proxies of R&D environment can serve as as instrumental variables. I use three instruments: R&D subsidies, the IPR protection index, and the number of researchers in the country. The first two are policies set at national level without regard to individual foreign firms; the last one depends largely on the supply factors. These variables might be correlated with other country-level determinants of affiliate production, which motivates me to control for host size, GDP per capita, and the general human capital measure. Finally, I also include affiliate and other fixed effects, so identification comes only from time variation within a host.

Columns 1 and 2 of Table A.10 report the baseline 2SLS results and the corresponding first stage regression. The first stage shows that R&D subsidies, the number of researchers, and IPR protection all have positive effects on affiliate R&D. The robust F statistic is above 10, the conventional rule of thumb for detecting weak IV. The 2SLS estimate suggests that a one-percent increase in affiliate R&D increases production by 0.46%, which is similar to the baseline estimate.

The lower panel of the table reports additional diagnostic statistics. Recent studies (c.f. Lee, McCrary, Moreira and Porter, 2020) suggest that 2SLS inference based on standard t-statistics might not be conservative enough. I report tests from two tests that are robust to weak IV, both of which are able to reject that the coefficient is zero. Finally, the over-identification test suggests that we cannot reject that all three IVs give the same estimate.

One might be concerned that IPR protection can affect affiliate production directly—in host counties experiencing an improvement in the protection of IPR, the risk of IP theft is lower, so firms might be more willing to move production there. Columns 3 and 4 allow IPR protection to be endogenous and use the remaining two IVs. The results are similar. The first-stage is slightly weaker, but the weak-IV robust tests are both able to reject that the key coefficient is zero.

I conduct additional robustness exercises: Columns 5 and 6 focus on manufacturing firms; Columns 7 through 10 use two alternative measures of invention—unweighted patent counts and citations; Columns 11 and 12 use value added, instead of sales, to measure firm production. All these robustness exercises give similar results.

This IV strategy has weaknesses: despite other time-varying controls, it is still possible that R&D subsidies and the number of researchers in a country are correlated with unobserved changes in country characteristics that directly affect affiliate production. But to the extent that this is the main concern, it is directly addressed in the baseline specifications with host-industry-period fixed effects. In this sense, while both identification strategies are imperfect, they exploit orthogonal variations and thus complement each other. That both approaches give qualitatively similar estimates is reassuring.

Table A.10: Fact 2: IV Estimates

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)
	Baseline IV		Endogenous IPR		Mfg. Only		R&D Measure 2		R&D Measure 3		Prod. Measure 2	
	2SLS	1st stage	2SLS	1st stage	2SLS	1st Stage	2SLS	1st stage	2SLS	1st stage	2SLS	1st stage
$\ln(\text{patent})_{fh,t}$	0.466***		0.516**		0.599***						0.673*	
	(0.141)		(0.240)		(0.160)						(0.354)	
$\ln(\text{unwgt. patent})_{fh,t}$							0.533***					
							(0.166)					
$\ln(\text{citation})_{fh,t}$									0.451**			
									(0.183)			
R&D subsidies		0.374		0.374		0.436*		0.181		0.060		0.304
		(0.228)		(0.228)		(0.265)		(0.210)		(0.329)		(0.186)
ln (researchers)		0.635***		0.635***		0.514***		0.607***		0.409**		0.562***
		(0.170)		(0.170)		(0.182)		(0.137)		(0.202)		(0.170)
IPR protection		0.459**	-0.062	0.459**		0.441**		0.431***		0.703***		0.139
		(0.191)	(0.231)	(0.191)		(0.220)		(0.137)		(0.187)		(0.169)
Observations	11464	11464	11464	11464	7533	7533	11464	11464	11464	11464	7585	7585
Time-varying controls	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y
Firm-period FE	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y
Home-Host FE	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y
Affiliate FE	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y
1st Stage F												
K-P F-stat		10.380		7.448		8.216		14.703		12.324		7.304
Weak IV robustness inference												
Anderson-Rubin test p-value	<0.01		0.028		<0.01		<0.01		<0.01		<0.01	
Stock-Wright test p-value	0.025		0.086		0.036		0.025		0.025		0.072	
Over identification test												
Hansen J statistic p-value	0.461		0.227		0.389		0.565		0.299		0.266	

Note: This reports the estimates for colocation between invention and production, using time variation in host R&D subsidies, the number of researchers, and the IPR protection index as instrument variables for changes in affiliate invention. All specification includes firm-period, home-home, and affiliate fixed effects and the following time varying controls: ln(GDP), ln(GDP per capita), and the human capital index. Columns 1 and 2 use all three IVs and the baseline sample; Columns 3 and 4 allows the IPR protection index to be endogenous; Columns 5 and 6 restrict to manufacturing industry only; Column 7 to 10 use two alternative measures of affiliate invention: unweighted patent counts and citation; Columns 11 and 12 use value added to measure affiliate production as the outcome variable. Standard errors (in parenthesis) are clustered two way, by host country and by firm. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

A.6 Additional Evidence on the Independence among R&D Centers

The quantitative model developed in Section 3 assumes that offshore R&D centers belonging to the same parent develop differentiated varieties and thus operate independently. This subsection provides evidence in support of this assumption.

Specifically, I investigate whether the R&D decision of an MNC in host i responds to changes in R&D-related policies or other factors in the headquarters o and in other countries $i' \neq i$ where the firm has a presence. The idea is as follows: Section A.5 of this appendix shows that R&D in host i responds to R&D-related shocks in country i. If the coordination among R&D centers is important, in response to the expansion of R&D in i, firms will adjust R&D in other hosts.[14]

Table A.11 reports the results. Columns 1 and 2 regress the extensive and intensive margin measures of offshore R&D on the characteristics of the headquarter country, with headquarters themselves excluded from the sample. I control for affiliate and host-period fixed effects, so the coefficients are identified off time variation among affiliates from different countries. None of the predictors of R&D—R&D subsidies, IPR protection, the number of researchers—has a statistically significant impact on affiliate R&D. This statistical insignificance is not due to the lack of variation in these characteristics: in fact, as the first-stage regressions reported in Table A.10 shows, the same variables have economically sizeable

[14] I focus on the response to shocks, rather than on the cross-sectional relation, because the model implies that the affiliates of the same parent inherit correlated innovation efficiency, hence their invention output might be correlated even if each of them carry out R&D independently.

and statistically significant impacts on R&D in the same host.

Columns 3 to 6 regress affiliate R&D on the average characteristics of all other countries in which the firm has a presence. Because firms operate in different sets of countries, the variation is at firm level and we can control for home-period fixed effects. Columns 3 and 4 include headquarters in the regression sample; Columns 5 and 6 exclude headquarters. Both set of regressions find that affiliate R&D do not respond to shocks affecting firm's R&D in either the home or other host countries.

Although I can not rule out that firms coordinate among their affiliates on R&D entirely, this piece of evidence suggests that such coordination is not a first-order feature of my data.

Table A.11: Evidence on R&D Center Independence

	(1)	(2)	(3)	(4)	(5)	(6)
	R&D and HQ shocks		R&D and shocks in other countries			
			incl. HQ		excl. HQ	
	R&D ind.	ln(patent)	R&D ind.	ln(patent)	R&D ind.	ln(patent)
Headquarter country char.:						
R&D subsidies	-0.000	0.122				
	(0.001)	(0.144)				
IPR protection	-0.001	0.007				
	(0.001)	(0.130)				
ln (researcher)	-0.000	0.006				
	(0.001)	(0.118)				
Average char. of other countries where firm has an affiliate:						
R&D subsidies			-0.039	-0.109	-0.072	-0.145
			(0.037)	(0.182)	(0.046)	(0.390)
IPR protection			0.007	0.008	0.012	0.111
			(0.009)	(0.041)	(0.011)	(0.086)
ln (researcher)			0.002	0.009	0.002	0.006
			(0.001)	(0.006)	(0.001)	(0.010)
Observations	3290430	21216	112731	34012	97773	21966
R^2	0.571	0.740	0.724	0.768	0.666	0.707
Within R^2	0.000	0.001	0.000	0.003	0.000	0.003
Time-varying controls	Y	Y	Y	Y	Y	Y
Affiliate FE	Y	Y	Y	Y	Y	Y
Host-period FE	Y	Y	Y	Y	Y	Y
Home-period FE	-	-	Y	Y	Y	Y

Note: Columns 1 and 2 regress measures of affiliate R&D on time-varying characteristics of the home country. Additional time-varying controls are ln(GDP), ln(GDP per capita), and the human capital index of the home country. Headquarters themselves are excluded from the regression. Columns 3 to 6 regress measures of affiliate R&D on the average characteristics among all other countries in which the firm has an affiliate. Additional time-varying controls are the average value of ln(GDP), ln(GDP per capita), and the human capital index of these countries. Columns 3 and 4 include headquarters as one of the affiliates; Columns 5 and 6 exclude the headquarters themselves. Affiliate fixed effects are controlled for throughout. Standard errors in parenthesis. Standard errors for Columns 1 and 2 are clustered two way, by home country and by firm; standard errors in Columns 3-5 are clustered by firm. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

B Theory

This section provides additional details on model aggregation, the full definition of the competitive equilibrium, and the proof of Proposition 1.

B.1 Aggregation

This subsection derives a few results under Assumption 1 for aggregation. For convenience I first introduce the following lemma, which has been proved in Arkolakis et al. (2018) and is only included for this appendix to be self-contained.

Lemma B.1. *Suppose $\boldsymbol{\eta} = (\eta)_{h=1}^{N}$ is a random variable with the following CDF:*

$$F(\boldsymbol{x}) \equiv Prob(\eta_1 \leq x_1, ..., \eta_N \leq x_N) = \begin{cases} 1 - (\sum_{m=1}^{N} \frac{1}{N} x_m^{-\theta}), \forall m \in \{1, ...N\},\ x_m \geq 1 \\ 0,\ \exists m \in \{1, ..., N\},\ x_m < 1 \end{cases}$$

Define $\zeta \equiv \max_m A_m \eta_m$, where $A_m,\ m = 1, ..., N$ are positive constants. Then the following holds:

1. *The CDF for ζ is*

$$Prob(\zeta \leq x) = \begin{cases} 1 - \tilde{A}^{\theta} x^{-\theta}, & \text{if } x \geq \bar{A} \\ 0, & \text{if } x < \bar{A}. \end{cases} \tag{B.1}$$

   where $\tilde{A} = \left(\frac{1}{N}\sum_m A_m^{\theta}\right)^{\frac{1}{\theta}}$ and $\bar{A} = \max_m A_m$

2. *The conditional expectation of ζ above x is:*

$$\mathbb{E}[\zeta | \zeta \geq x] = \frac{\theta}{\theta - 1} x,\ \forall x \geq \bar{A}.$$

3. *The conditional probability of the maximum value of $\sum_{m'} A_{m'} \eta_{m'}$ realizing at m is*

$$Prob(m = \arg\max_{m'} A_{m'} \eta_{m'} \Big| \zeta \geq x) = \frac{A_m^{\theta}}{\sum_m A_m^{\theta}},\ \forall x > \bar{A},\ \forall m = 1, 2, ..., N. \tag{B.2}$$

   Moreover, the distribution of ζ conditional on the maximum value realizing at m is:

$$Prob(\zeta \geq x' | m = \arg\max_{m'} A_{m'} \eta_{m'} \wedge \zeta \geq x\Big) = (\frac{x'}{x})^{-\theta},\ \forall x' \geq x > \bar{A}, \tag{B.3}$$

   which is independent of m.

Proof. 1. $\forall x$, following the definition of ζ

$$\begin{aligned} Prob(\zeta \leq x) &= Pr(A_1 \eta_1 \leq x, ..., A_N \eta_N \leq x) \\ &= Prob(\eta_1 \leq \frac{x}{A_1}, ..., \eta_N \leq \frac{x}{A_N}) \\ &\quad \text{(using the definition of F)} \\ &= \begin{cases} 1 - \left(\sum_m \frac{1}{N} \big[(\frac{x}{A_m})^{-\theta}\big]\right) = 1 - \tilde{A}^{\theta} x^{-\theta}, \text{ if } x \geq \max_m A_m \equiv \bar{A} \\ 0, \quad \text{if } x < \bar{A}. \end{cases} \end{aligned}$$

2. From Equation (B.1), for $x \geq \bar{A}$, from part 1, $\forall x' \geq x$

$$Pr(\zeta > x' | \zeta > x) = \frac{(\frac{x'}{\bar{A}})^{-\theta}}{(\frac{x}{\bar{A}})^{-\theta}} = \left(\frac{x'}{x}\right)^{-\theta}.$$

Therefore, the conditional distribution of ζ above $\forall x \geq \bar{A}$ is Pareto with tail parameter θ and scale parameter x. Thus we have

$$\mathbb{E}[\zeta | \zeta > x] = \frac{\theta}{\theta - 1} x.$$

3. For $x > \bar{A}$,

$$Pr(m = \arg\max_{m'} A_{m'}\eta_{m'} \wedge A_m \eta_m \geq x) = \int_x^{\infty} Prob(A_{m'}\eta_{h'} \leq u, \forall m' \neq m \big| A_m \eta_m = u) f_m(u) du,$$

where $f_m(u)$ is the marginal density of $A_m \eta_m$. For $u \geq x > \bar{A}$, the integrand in the above function is:

$$\begin{aligned} Prob(A_{m'}\eta_{m'} \leq u, \forall m' \neq m \big| A_m \eta_m = u) f_m(u) &= \frac{\partial Pr(A_1\eta_1 \leq u, A_m\eta_m \leq C, ...A_N\eta_N \leq u)}{\partial C}\Big|_{C=u} \\ &= \frac{A_m^{\theta}}{N} \theta u^{-\theta - 1}. \end{aligned}$$

Therefore,

$$Pr(m = \arg\max_{m'} A_{m'}\eta_{m'} \wedge \zeta \geq x) = \frac{A_m^{\theta}}{N} x^{-\theta}.$$

And

$$\begin{aligned} Pr(m = \arg\max_{m'} A_{m'}\eta_{m'} \Big| \zeta \geq x) &= \frac{Pr(m = \arg\max_{m'} A_{m'}\eta_{m'} \wedge \zeta \geq x)}{Pr(\zeta \geq x)} \\ &= \frac{\frac{A_m^{\theta}}{N} x^{-\theta}}{\left(\frac{1}{N}\sum_{m'} A_{m'}^{\theta}\right) x^{-\theta}} \\ &= \frac{A_m^{\theta}}{\sum_{m'} A_{m'}^{\theta}}, \end{aligned}$$

Note that $\forall x' \geq x > \bar{A}$

$$\begin{aligned} Prob(\zeta \geq x' | m = \arg\max_{m'} A_{m'}\eta_{m'} \wedge \zeta \geq x\Big) &= \frac{Pr(m = \arg\max_{m'} A_{m'}\eta_{m'} \wedge \zeta \geq x' \wedge \zeta \geq x)}{Pr(m = \arg\max_{m'} A_{m'}\eta_{m'} \wedge \zeta \geq x)} \\ &= \frac{\frac{A_m^{\theta}}{N} x'^{-\theta}}{\frac{A_m^{\theta}}{N} x^{-\theta}} \\ &= (\frac{x'}{x})^{-\theta}. \end{aligned}$$

□

Deriving Equations (1), (2), (7) and (8). With Lemma B.1, I derive expressions for a few aggregate objects. For convenience, define a new random variable $\zeta_{oid} \equiv \max_m \frac{T_m \phi_{oim}^P}{W_m^m \tau_{md}} \cdot \eta_m$, then from the first part of the Lemma, the distribution of ζ_{oid} is given by:

$$H_{oid}(x) \equiv Prob(\zeta_{oid} \leq x) = \begin{cases} 1 - (\frac{\tilde{\zeta}_{oid}}{x})^\theta, \ x \geq \overline{\zeta_{oid}} \\ 0, \ x < \overline{\zeta_{oid}} \end{cases}$$

where $\tilde{\zeta}_{oid} \equiv \left(\sum_m \frac{1}{N}(\frac{T_m\phi^P_{oim}}{W^l_m\tau_{md}})^\theta\right)^{\frac{1}{\theta}}$ and $\overline{\zeta_{oid}} \equiv \max_m \frac{T_m\phi^P_{oim}}{W^l_m\tau_{md}}$. Here $\tilde{\zeta}_{oid}$ and $\overline{\zeta_{oid}}$ correspond to $\tilde{A}$ and $\overline{A}$ in Equation (B.1), respectively.

Noting that $p_{oid}(z^P, \boldsymbol{\zeta}) = \frac{\sigma}{\sigma-1} \cdot \frac{1}{\zeta_{iod}} \cdot \frac{1}{z^P}$, the probability that the product is manufactured in m is simply the probability that the best realization of ζ_{oid} realizes in m. From the third part of the lemma, this probability is

$$\psi_{oimd} \equiv Prob(m = \max_m \frac{T_m\phi^P_{oim}}{W^l_m\tau_{md}} \cdot \eta_m | \zeta_{oid} > x) = \frac{\frac{1}{N}(\frac{T_m\phi^P_{oim}}{W^l_m\tau_{md}})^\theta}{\frac{1}{N}\sum_{m'}(\frac{T_m\phi^P_{oim'}}{W^l_{m'}\tau_{m'd}})^\theta}, \ \forall x \geq \bar{\zeta}_{oid} \tag{B.4}$$

Because the conditional distribution of ζ_{oid} is the same regardless of which country ends up with the maximum value for $\frac{T_m\phi^P_{oim}}{W^l_m\tau_{md}} \cdot \eta_m$ (Equation (B.3)), the above choice probability is also equal to the share of sales produced in m.

For later use, I calculate the following:

$$\int_{\boldsymbol{\eta}} \mathbb{1}(p_{oid}(z^P, \boldsymbol{\zeta}) < \hat{p}_d) \cdot p_{oid}(z^P, \boldsymbol{\eta})^{1-\sigma} dF(\boldsymbol{\eta}) \tag{B.5}$$

$$= (\frac{\sigma}{\sigma-1})^{1-\sigma}(\frac{1}{z^P})^{1-\sigma} \int_{\frac{\sigma}{\sigma-1}\frac{1}{\hat{p}_d}\frac{1}{z^P}}^{\infty} \zeta_{oid}^{\sigma-1} dH_{oid}(\zeta_{oid})$$

(under the assumption that $\frac{\sigma}{\sigma-1}\frac{1}{\hat{p}_d}\frac{1}{z^P} > \bar{\zeta}_{oid}$, which is implied by Assumption 1.b)

$$= (\frac{\sigma}{\sigma-1})^{1-\sigma}(\frac{1}{z^P})^{1-\sigma} \int_{\frac{\sigma}{\sigma-1}\frac{1}{\hat{p}_d}\frac{1}{z^P}}^{\infty} \zeta_{oid}^{\sigma-1} d[1 - (\frac{\tilde{\zeta}_{oid}}{\zeta_{oid}})^\theta]$$

(assuming $\theta > \sigma - 1$)

$$= \frac{\theta}{\theta - (\sigma-1)} \hat{p}_d^{\theta+1-\sigma} (\frac{\sigma-1}{\sigma}\tilde{\zeta}_{oid} z^P)^\theta$$

(plugging in $\hat{p}_d$)

$$= \frac{\theta}{\theta - (\sigma-1)} (\frac{\sigma-1}{\sigma})^\theta P_d^{\theta+1-\sigma} (\frac{\sigma W^h_d c^M_d}{X_d})^{\frac{\theta+1-\sigma}{1-\sigma}} (\tilde{\zeta}_{oid} z^P)^\theta$$

I define $\bar{r}_{oid}(z^P)$ to be the expected revenue. Combine Equation (B.5) with the definition of $\bar{r}_{oid}(z^P)$ to obtain:

$$\begin{aligned} \bar{r}_{oid}(z^P) &= \frac{X_d}{P_d^{1-\sigma}} \int_{\boldsymbol{\eta}} \mathbb{1}(p_{oid}(z^P, \boldsymbol{\zeta}) < \hat{p}_d) \cdot p_{oid}(z^P, \boldsymbol{\eta})^{1-\sigma} dF(\boldsymbol{\eta}) \\ &= \frac{\theta}{\theta - (\sigma-1)} (\sigma-1)^\theta \sigma^{1-\frac{\theta\sigma}{\sigma-1}} X_d^{\frac{\theta}{\sigma-1}} P_d^\theta (W^h_d c^M_d)^{\frac{\theta+1-\sigma}{1-\sigma}} (\tilde{\zeta}_{oid} z^P)^\theta \end{aligned} \tag{B.6}$$

Define the expected marketing cost incurred for a variety as $\bar{c}^M_{oid}(z^P)$, then similar steps give:

$$\begin{aligned}\bar{c}^M_{oid}(z^P) &= c^M_d W^h_d \int_{\boldsymbol{\eta}} \mathbb{1}(p_{oid}(z^P,\boldsymbol{\zeta}) < \hat{p}_d) dF(\boldsymbol{\eta}) \\ &= c^M_d W^h_d \int_{\frac{\sigma}{\sigma-1}\frac{1}{\hat{p}_d}\frac{1}{z^P}}^{\infty} dH_{oid}(\zeta_{oid}) \\ &= c^M_d W^h_d (\tilde{\zeta}_{oid}\frac{\sigma-1}{\sigma}\hat{p}_d z^P)^\theta \\ &= (\frac{\sigma-1}{\sigma})^\theta \sigma^{\frac{\theta}{1-\sigma}} (c^M_d W^h_d)^{\frac{\sigma-1-\theta}{\sigma-1}} X_d^{\frac{\theta}{\sigma-1}} P_d^\theta (\tilde{\zeta}_{oid} z^P)^\theta.\end{aligned} \tag{B.7}$$

The operational profit is simply the difference between markup and marketing cost, given by:

$$\begin{aligned}\bar{\pi}_{oid}(z^P) &= \frac{1}{\sigma}\frac{X_d}{P_d^{1-\sigma}} \int_{\boldsymbol{\eta}} \mathbb{1}(p_{oid}(z^P,\boldsymbol{\zeta}) < \hat{p}_d)\cdot p_{oid}(z^P,\boldsymbol{\eta})^{1-\sigma} dF(\boldsymbol{\eta}) - c^M_d w^H_d \int_{\boldsymbol{\eta}} \mathbb{1}(p_{oid}(z^P,\boldsymbol{\zeta}) < \hat{p}_d) dF(\boldsymbol{\eta}) \\ &= \frac{1}{\sigma}\bar{r}_{oid}(z^P) - \bar{c}^M_{oid}(z^P) \\ &= \frac{(\sigma-1)^{1+\theta}}{\theta-(\sigma-1)} \sigma^{\frac{\sigma\theta}{1-\sigma}} (c^M_d w^H_d)^{\frac{\sigma-1-\theta}{\sigma-1}} X_d^{\frac{\theta}{\sigma-1}} P_d^\theta (\tilde{\zeta}_{oid} z^P)^\theta.\end{aligned} \tag{B.8}$$

Equations (B.6), (B.7), (B.8) immediately imply:

$$\begin{aligned}\frac{\bar{c}^M_{oid}(z^P)}{\bar{r}_{oid}(z^P)} &= \frac{\theta-(\sigma-1)}{\theta\sigma} \\ \frac{\bar{\pi}_{oid}(z^P)}{\bar{r}_{oid}(z^P)} &= \frac{\sigma-1}{\theta\sigma}.\end{aligned}$$

I now derive the aggregate price index P_d and trade flows X_{oid}

$$\begin{aligned}P_d^{1-\sigma} &= \sum_o\sum_i R_{oi}\int_{\mathbb{Z}^P}\int_{\mathbb{Z}^R} v_{oi}(z^P,z^R)[\int_{\boldsymbol{\eta}} \mathbb{1}(p_{oid}(z^P,\boldsymbol{\zeta}) < \hat{p}_d)\cdot p_{oid}(z^P,\boldsymbol{\eta})^{1-\sigma} dF(\boldsymbol{\eta})]\cdot g_{oi}(z^P,z^R)dz^P dz^R \\ &= \sum_o\sum_i R_{oi}\int_{\mathbb{Z}^P}\int_{\mathbb{Z}^R} v_{oi}(z^P,z^R)[\frac{\theta}{\theta-(\sigma-1)}(\sigma-1)^\theta\sigma^{1-\frac{\theta\sigma}{\sigma-1}} X_d^{\frac{\theta+1-\sigma}{\sigma-1}} P_d^{\theta+1-\sigma}(W^h_d c^M_d)^{\frac{\theta+1-\sigma}{1-\sigma}}(\tilde{\zeta}_{oid}z^P)^\theta]\cdot g_{oi}(z^P,z^R)dz^P dz^R \\ &= \frac{\theta}{\theta-(\sigma-1)}(\frac{\sigma-1}{\sigma})^\theta P_d^{\theta+1-\sigma}(\frac{\sigma W^h_d c^M_d}{X_d})^{\frac{\theta+1-\sigma}{1-\sigma}} \sum_o\sum_i \tilde{\zeta}^\theta_{oid} R_{oi}\int_{\mathbb{Z}^P}(z^P)^\theta[\int_{\mathbb{Z}^R} v_{oi}(z^P,z^R)\cdot g_{oi}(z^P,z^R)dz^R]dz^P \\ &= \frac{\theta}{\theta-(\sigma-1)}(\frac{\sigma-1}{\sigma})^\theta P_d^{\theta+1-\sigma}(\frac{\sigma W^h_d c^M_d}{X_d})^{\frac{\theta+1-\sigma}{1-\sigma}} \sum_o\sum_i \tilde{\zeta}^\theta_{oid}\int_{\mathbb{Z}^P}(z^P)^\theta V_{oi}(z^P)dz^P.\end{aligned}$$

$$\begin{aligned}X_{oid} &= P_d^{\sigma-1}X_d\cdot R_{oi}\int_{\mathbb{Z}^P}\int_{\mathbb{Z}^R} v_{oi}(z^P,z^R)[\int_{\boldsymbol{\eta}} \mathbb{1}(p_{oid}(z^P,\boldsymbol{\zeta}) < \hat{p}_d)\cdot p_{oid}(z^P,\boldsymbol{\eta})^{1-\sigma} dF(\boldsymbol{\eta})]\cdot g_{oi}(z^P,z^R)dz^P dz^R \\ &= P_d^{\sigma-1}X_d\cdot\frac{\theta}{\theta-(\sigma-1)}(\frac{\sigma-1}{\sigma})^\theta P_d^{\theta+1-\sigma}(\frac{\sigma W^h_d c^M_d}{X_d})^{\frac{\theta+1-\sigma}{1-\sigma}}\tilde{\zeta}^\theta_{oid}\cdot R_{oi}\int_{\mathbb{Z}^P}(z^P)^\theta[\int_{\mathbb{Z}^R} v_{oi}(z^P,z^R)\cdot g_{oi}(z^P,z^R)dz^R]dz^P \\ &= \frac{\theta}{\theta-(\sigma-1)}(\frac{\sigma-1}{\sigma})^\theta(\frac{X_d}{P_d^{1-\sigma}})^{\frac{\theta}{\sigma-1}}(\sigma W^h_d c^M_d)^{\frac{\theta+1-\sigma}{1-\sigma}}\tilde{\zeta}^\theta_{oid}\int_{\mathbb{Z}^P}(z^P)^\theta V_{oi}(z^P)dz^P.\end{aligned}$$

B.2 Definition of Equilibrium

Definition 1. *Given the fundamentals, a competitive equilibrium of the model is characterized by a set of decision rules, prices, and allocations, such that* $\forall o, i, d = 1, .., N$ *the following holds:*

1. *Firms' production, market entry, and pricing decisions for each individual variety are optimal, which implies that the following holds* $\forall z^P \in \mathbb{Z}^P$*:*

$$\begin{aligned}
\bar{r}_{oid}(z^P) &= \frac{\theta(\sigma-1)^\theta \sigma^{1-\frac{\theta\sigma}{\sigma-1}}}{\theta-(\sigma-1)} X_d^{\frac{\theta}{\sigma-1}} P_d^\theta (W_d^h c_d^M)^{\frac{\theta+1-\sigma}{1-\sigma}} (\tilde{\zeta}_{oid} z^P)^\theta \\
\bar{c}_{oid}^M(z^P) &= \frac{\theta-(\sigma-1)}{\theta\sigma} \bar{r}_{oid}(z^P) \\
\bar{\pi}_{oid}(z^P) &= \frac{1}{\sigma}\bar{r}_{oid}(z^P) - \bar{c}_{oid}^M(z^P) = \frac{\sigma-1}{\theta\sigma}\bar{r}_{oid}(z^P) \\
\bar{\pi}_{oi}(z^P) &= \sum_d \bar{\pi}_{oid}(z^P), \\
&\textit{where } \tilde{\zeta}_{oid} \equiv [\sum_m \frac{1}{N} (\frac{T_m \phi_{oim}^p}{W_m^l \tau_{md}})^\theta]^{\frac{1}{\theta}}.
\end{aligned} \tag{B.9}$$

2. *Firms' R&D and offshore R&D decisions satisfy the following:*

$$\begin{aligned}
v_{oi}(z^P, z^R) &= z^{R\,\frac{1}{1-\gamma}} \Big(\frac{\gamma\bar{\pi}_{oi}(z^P)}{W_i^h}\Big)^{\frac{\gamma}{1-\gamma}} \\
\pi_{oi}^R(z^P, z^R) &= (\gamma^{\frac{\gamma}{1-\gamma}} - \gamma^{\frac{1}{1-\gamma}}) \Big(\frac{1}{W_i^h}\Big)^{\frac{\gamma}{1-\gamma}} (\bar{\pi}_{oi}(z^P) z^R)^{\frac{1}{1-\gamma}} \\
\pi_{oi}^R(z^R) &= \int_{\mathbf{Z}^P} \pi_{oi}^R(z^P, z^R) g^P(z^P|z^R) dz^P \\
\pi_{oi}^R(\hat{z}_{oi}^R \phi_{oi}^R) &= c_i^R W_i^h
\end{aligned} \tag{B.10}$$

3. *The distribution of R&D center innovation efficiency in each host is consistent with firms' offshore R&D decisions and the endowment distribution of the origin countries:*

$$\begin{aligned}
R_{oi} &= E_o \cdot \Big(1 - G_o^E(\hat{z}_{oi}^R)\Big) \\
g_{oi}^R(z^R) &= \frac{1}{R_{oi}} \mathbb{1}(z^R > \hat{z}_{oi}^R \phi_{oi}^R) \cdot E_o \cdot g_o^E(\frac{z^R}{\phi_{oi}^R}) \cdot \frac{1}{\phi_{oi}^R} \\
g_{oi}(z^P, z^R) &= g^P(z^P|z^R) g_{oi}^R(z^R).
\end{aligned} \tag{B.11}$$

4. *Firm decisions are consistent with aggregate trade flows*

$$V_{oi}(z^P) = R_{oi} \int_{\mathbb{Z}^R} v_{oi}(z^P, z^R) \cdot g_{oi}(z^P, z^R) dz^R \tag{B.12}$$

$$X_{oid} = \theta(\frac{\sigma}{\sigma-1})^{-\theta} \frac{1}{\theta-(\sigma-1)} (\sigma W_d^h c_d^M)^{\frac{\theta-(\sigma-1)}{1-\sigma}} \left(\frac{X_d}{P_d^{1-\sigma}}\right)^{\frac{\theta}{\sigma-1}} \tilde{\zeta}_{oid}^{\theta} \int_{\mathbf{Z}^P} (z^P)^{\theta} V_{oi}(z^P) dz^P$$

$$X_{oimd} = \psi_{oimd} X_{oid}$$

$$Y_{om} = \frac{\sigma-1}{\sigma} \sum_{i,d} X_{oimd}$$

$$\Pi_{oi} = \frac{1-\gamma}{\sigma} \left(\frac{\sigma-1}{\theta}\right) \sum_{m,d} X_{oimd}$$

$$C_{od}^M = \frac{1}{\sigma} \left(\frac{\theta-(\sigma-1)}{\theta}\right) \sum_{i,m} X_{oimd}$$

$$I_{oi} = \frac{\gamma}{1-\gamma} \Pi_{oi}$$

$$C_{oi}^R = \mathbb{1}(o \neq d) \cdot E_o \cdot [1 - G_o^R(\hat{z}_{od}^R)] \cdot c_{oi}^R W_i^h,$$

and the aggregate price indices

$$P_d^{1-\sigma} = \theta(\frac{\sigma}{\sigma-1})^{-\theta} \frac{1}{\theta-(\sigma-1)} \left(\frac{\sigma W_d^h c_d^M}{X_d}\right)^{\frac{\theta-(\sigma-1)}{1-\sigma}} P_d^{\theta-(\sigma-1)} \sum_o \sum_i \tilde{\zeta}_{oid}^{\theta} \int_{\mathbf{Z}^P} (z^P)^{\theta} V_{oi}(z^P) dz^P \tag{B.13}$$

5. *Workers' occupation choices are optimal:*

$$W_d^h \hat{\alpha}_d = W_d^l \tag{B.14}$$

$$L_d^h = L_d \cdot \int_{\alpha > \hat{\alpha}_d} \alpha \, dA_d(\alpha)$$

$$L_d^l = L_d \cdot A_d(\hat{\alpha}_d).$$

6. *Labor markets clear*

$$W_d^h L_d^h = \sum_o I_{od} + \sum_o C_{od}^M + \sum_o C_{od}^R \tag{B.15}$$

$$W_d^l L_d^l = \sum_o Y_{od}$$

7. *Total income equals total expenditures*

$$X_d = W_d^h L_d^h + W_d^l L_d^l + \sum_i (\Pi_{di} - C_{di}^R) \tag{B.16}$$

Note that once $\{\hat{\alpha}_d, W_d^l, W_d^h : d \in 1, ..., N\}$, $\{X_d : d \in 1, ..., N\}$, and $\{P_d : d \in 1, ..., N\}$ are known, all other endogenous variables appearing in the above definition can be calculated sequentially with Equations (B.9), (B.10), (B.11), and (B.12). The competitive equilibrium can therefore be viewed as a fixed point in $\{\hat{\alpha}_d, W_d^l, W_d^h : d \in 1, ..., N\}$, $\{X_d : d \in 1, ..., N\}$, and $\{P_d : d \in 1, ..., N\}$, so that Equations (B.13), (B.14) , (B.15), and (B.16) hold, with endogenous objects in these four set of equations defined implicitly as functions of wages, prices and expenditures by Equations (B.9), (B.10), (B.11), and (B.12).

B.3 Proof of Proposition 1

Proof. I proceed in four steps. The first two steps express the real wage for low-skill workers as flow variables. The third step derives the relationship between real wage and real income. The fourth and final step combines results from the first three steps to derive $\frac{X_d}{P_d}$ as a function of several ratios and calculate the gains from openness by setting some of the ratios to their autarky values, which is usually 1 or a model constant. I then compare my expression of the gains from openness to that in the literature.

Step 1: Expressing real wage for low-skill workers as flow variables

The first step derives $\frac{W_d^l}{P_d}$, and is a bit tedious. To give a broad direction, I will express P_d as a function of the share of consumption expenditures of country d spent on goods *invented* in country d. Intuitively, P_d measures competitiveness of the product market in country d; for any given level of domestic invention, if a higher share of income is spent on these inventions, then it must be because country d has a poor access to varieties invented elsewhere, and P_d will thus be high. In a similar vein, I will express W_d^l as a function of the share of consumption expenditures of country d spent on the goods *produced* in country d. All else equal, if this share is higher, then the wage of country d is likely lower.

It will prove convenient to define K_{oi} as follows:

$$K_{oi} \equiv \int_{\mathbf{Z}^P} (z^P)^\theta V_{oi}(z^P) dz^P. \tag{B.17}$$

Loosely speaking, K_{oi} is the productivity-adjusted number of varieties invented in i by firms from o.

From Equation (B.13) we have the following:

$$\begin{aligned} P_d^{-\theta} &= \theta\left(\frac{\sigma}{\sigma-1}\right)^{-\theta} \frac{1}{\theta-(\sigma-1)} \left(\frac{\sigma W_d^h c_d^M}{X_d}\right)^{\frac{\theta-(\sigma-1)}{1-\sigma}} \sum_o \sum_i \tilde{\zeta}_{oid}^\theta \int_{\mathbf{Z}^P} (z^P)^\theta V_{oi}(z^P) dz^P \\ &= \theta\left(\frac{\sigma}{\sigma-1}\right)^{-\theta} \frac{1}{\theta-(\sigma-1)} \left(\frac{\sigma W_d^h c_d^M}{X_d}\right)^{\frac{\theta-(\sigma-1)}{1-\sigma}} \sum_o \sum_i \tilde{\zeta}_{oid}^\theta \cdot K_{oi} \end{aligned}$$

Similarly, from Equation (B.12) we have:

$$X_{oid} = \theta\left(\frac{\sigma}{\sigma-1}\right)^{-\theta} \frac{1}{\theta-(\sigma-1)} (\sigma W_d^h c_d^M)^{\frac{\theta-(\sigma-1)}{1-\sigma}} \left(\frac{X_d}{P_d^{1-\sigma}}\right)^{\frac{\theta}{\sigma-1}} \tilde{\zeta}_{oid}^\theta K_{oi}$$

Define $\lambda_{oid}^E = \frac{\sum_m X_{oimd}}{X_d}$, which denotes the share of expenditure in country d spent on goods invented in i by firms from country o, I obtain:

$$\lambda_{oid}^E = \frac{X_{oid}}{X_d} = \frac{\tilde{\zeta}_{oid}^\theta K_{oi}}{\sum_{o,i} \tilde{\zeta}_{oid}^\theta K_{oi}} = \frac{\tilde{\zeta}_{oid}^\theta K_{oi}}{P_d^{-\theta}} \cdot \theta\left(\frac{\sigma}{\sigma-1}\right)^{-\theta} \frac{1}{\theta-(\sigma-1)} \left(\frac{\sigma W_d^h c_d^M}{X_d}\right)^{\frac{\theta-(\sigma-1)}{1-\sigma}},$$

which implies

$$P_d^{-\theta} = \frac{\theta\left(\frac{\sigma}{\sigma-1}\right)^{-\theta} \frac{1}{\theta-(\sigma-1)} \left(\frac{\sigma W_d^h c_d^M}{X_d}\right)^{\frac{\theta-(\sigma-1)}{1-\sigma}} \cdot \tilde{\zeta}_{oid}^\theta K_{oi}}{\lambda_{oid}^E}.$$

Specializing this equation to $o = i = d$ gives

$$P_d^{-\theta} = \frac{\theta\left(\frac{\sigma}{\sigma-1}\right)^{-\theta} \frac{1}{\theta-(\sigma-1)} \left(\frac{\sigma W_d^h c_d^M}{X_d}\right)^{\frac{\theta-(\sigma-1)}{1-\sigma}} \cdot \tilde{\zeta}_{ddd}^\theta K_{dd}}{\lambda_{ddd}^E} \tag{B.18}$$

Now consider W_d^l. Define $\lambda_{dd}^T = \frac{\sum_{o,i} X_{oidd}}{X_d}$ as the fraction of country d's spending on goods produced in d, regardless where it is invented and which the headquarter countries of the firms are.

Noticing $\lambda_{dd}^T = \sum_{o,i} \psi_{oidd}\lambda_{oid}^E$ and recalling $\psi_{oimd} = \frac{\frac{1}{N}(\frac{T_m\phi_{oim}^P}{W_m^l\tau_{md}})^\theta}{\tilde{\zeta}_{oid}^\theta}$, we have

$$\lambda_{dd}^T = \sum_{o,i} \frac{\frac{1}{N}(\frac{T_d\phi_{oid}^P}{W_d^l\tau_{dd}})^\theta}{\tilde{\zeta}_{oid}^\theta}\lambda_{oid}^E \tag{B.19}$$

$$\implies (W_d^l)^\theta = \frac{1}{\lambda_{dd}^T}\sum_{o,i}\frac{\frac{1}{N}(T_d\phi_{oid}^P)^\theta}{\tilde{\zeta}_{oid}^\theta}\lambda_{oid}^E$$

Combining Equations (B.18) and (B.19) gives

$$P_d^{-\theta}\cdot(W_d^l)^\theta = \frac{\theta(\frac{\sigma}{\sigma-1})^{-\theta}\frac{1}{\theta-(\sigma-1)}\left(\frac{\sigma W_d^h c_d^M}{X_d}\right)^{\frac{\theta-(\sigma-1)}{1-\sigma}}}{\lambda_{dd}^T}[\sum_{o,i}\frac{\frac{1}{N}(T_d\phi_{oid}^P)^\theta}{\tilde{\zeta}_{oid}^\theta}\lambda_{oid}^E]\frac{\tilde{\zeta}_{ddd}^\theta K_{dd}}{\lambda_{ddd}^E} \tag{B.20}$$

$$= \frac{\theta(\frac{\sigma}{\sigma-1})^{-\theta}\frac{1}{\theta-(\sigma-1)}\left(\frac{\sigma W_d^h cM}{X_d}\right)^{\frac{\theta-(\sigma-1)}{1-\sigma}}}{\lambda_{dd}^T}[\sum_{o,i}\frac{1}{N}(T_d\phi_{oid}^P)^\theta\frac{\tilde{\zeta}_{ddd}^\theta}{\tilde{\zeta}_{oid}^\theta}\lambda_{oid}^E]\frac{K_{dd}}{\lambda_{ddd}^E}$$

The term $[\sum_{o,i}\frac{1}{N}(T_d\phi_{oid}^P)^\theta\frac{\tilde{\zeta}_{ddd}^\theta}{\tilde{\zeta}_{oid}^\theta}\lambda_{oid}^E]$ broadly captures the importance of country d as a production location, which can be derived as a function of flows as follows:

$$\frac{X_{oidd}}{X_d} = \lambda_{oid}^E\cdot\frac{\frac{1}{N}(\frac{T_d\phi_{oid}^P}{W_d^l\tau_{dd}})^\theta}{\tilde{\zeta}_{oid}^\theta}$$

$$\implies \frac{X_{oidd}}{X_d}\tilde{\zeta}_{ddd}^\theta = \lambda_{oid}^E\cdot\frac{\tilde{\zeta}_{ddd}^\theta}{\tilde{\zeta}_{oid}^\theta}\cdot\frac{1}{N}(\frac{T_d\phi_{oid}^P}{W_d^l\tau_{dd}})^\theta$$

$$\implies \frac{X_{oidd}}{X_d}\tilde{\zeta}_{ddd}^\theta(W_d^l)^\theta = \lambda_{oid}^E\cdot\frac{\tilde{\zeta}_{ddd}^\theta}{\tilde{\zeta}_{oid}^\theta}\cdot\frac{1}{N}(T_d\phi_{oid}^P)^\theta$$

$$\implies \frac{X_{oidd}}{X_d}\frac{T_d^\theta}{\psi_{dddd}} = \lambda_{oid}^E\cdot\frac{\tilde{\zeta}_{ddd}^\theta}{\tilde{\zeta}_{oid}^\theta}\cdot\frac{1}{N}(T_d\phi_{oid}^P)^\theta$$

$$\implies \frac{T_d^\theta}{\psi_{dddd}}\sum_{o,i}\frac{X_{oidd}}{X_d} = \sum_{o,i}\lambda_{oid}^E\cdot\frac{\tilde{\zeta}_{ddd}^\theta}{\tilde{\zeta}_{oid}^\theta}\cdot\frac{1}{N}(T_d\phi_{oid}^P)^\theta$$

$$\implies T_d^\theta\frac{\sum_m X_{ddmd}}{X_{dddd}}\cdot\lambda_{dd}^T = \sum_{o,i}\lambda_{oid}^E\cdot\frac{\tilde{\zeta}_{ddd}^\theta}{\tilde{\zeta}_{oid}^\theta}\cdot\frac{1}{N}(T_d\phi_{oid}^P)^\theta$$

Plugging this into Equation (B.20), I obtain:

$$
\begin{aligned}
P_d^{-\theta} \cdot (W_d^l)^\theta &= \frac{\theta(\frac{\sigma}{\sigma-1})^{-\theta}\frac{1}{\theta-(\sigma-1)}\left(\frac{\sigma W_d^h c_d^M}{X_d}\right)^{\frac{\theta-(\sigma-1)}{1-\sigma}}}{\lambda_{dd}^T} \cdot T_d^\theta \cdot \frac{\sum_m X_{ddmd}}{X_{dddd}} \cdot \lambda_{dd}^T \cdot \frac{K_{dd}}{\lambda_{ddd}^E} \\
&= \theta(\frac{\sigma}{\sigma-1})^{-\theta}\frac{1}{\theta-(\sigma-1)}\left(\frac{\sigma W_d^h c_d^M}{X_d}\right)^{\frac{\theta-(\sigma-1)}{1-\sigma}} \cdot T_d^\theta \cdot \frac{\sum_m X_{ddmd}}{X_{dddd}} \cdot \frac{K_{dd}}{\lambda_{ddd}^E} \\
&\quad (\text{noting that } \lambda_{ddd}^E \equiv \frac{\sum_m X_{oimd}}{X_d} = \frac{\sum_{o,m} X_{odmd}}{X_d} \cdot \frac{\sum_m X_{ddmd}}{\sum_{o,m} X_{odmd}}) \\
&= T_d^\theta\theta(\frac{\sigma}{\sigma-1})^{-\theta}\frac{1}{\theta-(\sigma-1)}\left(\frac{\sigma W_d^h c_d^M}{X_d}\right)^{\frac{\theta-(\sigma-1)}{1-\sigma}} \cdot \frac{\sum_m X_{ddmd}}{X_{dddd}} \cdot \frac{X_d}{\sum_{o,m} X_{odmd}} \cdot \frac{\sum_{o,m} X_{odmd}}{\sum_m X_{ddmd}} \cdot K_{dd} \\
&\quad (\text{noting that } T_d^\theta, \theta, \sigma, c_d^M \text{ are constants }) \\
\implies (\frac{W_d^l}{P_d}) &\propto (\frac{W_d^h}{X_d})^{\frac{\theta-\sigma+1}{(1-\sigma)\theta}} \cdot (\frac{X_{dddd}}{\sum_m X_{ddmd}})^{-\frac{1}{\theta}} \cdot (\frac{\sum_{o,m} X_{odmd}}{X_d})^{-\frac{1}{\theta}} \cdot (\frac{\sum_m X_{ddmd}}{\sum_{o,m} X_{odmd}})^{-\frac{1}{\theta}} \cdot (K_{dd})^{\frac{1}{\theta}}.
\end{aligned}
$$

We can already see that the real wage for low-skill workers is a function of $\frac{X_{dddd}}{\sum_m X_{ddmd}}, \frac{\sum_{o,m} X_{odmd}}{X_d}, \frac{\sum_m X_{ddmd}}{\sum_{o,m} X_{odmd}}$. They capture the importance of foreign locations for production, the importance of varieties developed outside the country, and the importance of foreign firms in domestic R&D, respectively. Note also that all these ratios equal to one in autarky. Define $\widehat{x}$ be the ratio between the baseline variable x and its value in autarky x', we have:

$$
\widehat{\frac{W_d^l}{P_d}} = (\widehat{\frac{W_d^h}{X_d}})^{\frac{\theta-\sigma+1}{(1-\sigma)\theta}}(\frac{X_{dddd}}{\sum_m X_{ddmd}})^{-\frac{1}{\theta}} \cdot (\frac{\sum_{o,m} X_{odmd}}{X_d})^{-\frac{1}{\theta}} \cdot (\frac{\sum_m X_{ddmd}}{\sum_{o,m} X_{odmd}})^{-\frac{1}{\theta}} \cdot (\widehat{K_{dd}})^{\frac{1}{\theta}} \tag{B.21}
$$

Step 2: deriving $\widehat{K_{dd}}$.

Now I derive $\widehat{K_{dd}} \equiv \frac{K_{dd}}{K'_{dd}}$, where K'_{dd} is the autarky value of K_{dd}. Assume that the total number of high-skill workers in country i working directly on variety development is L_i^R and let L_{oi}^R be those working at R&D centers from o: $L_i^R = \sum_o L_{oi}^R$. Note that L_i^R is endogenous variable even though the total number of high-skill workers is assumed to be exogenous for this proposition.

$$
\begin{aligned}
L_{oi}^R &= R_{oi}\int_{\mathbf{Z}^P}[\int_{\mathbf{Z}^R} h_{oi}(z^P, z^R) g_{oi}^R(z^R) dz^R] \cdot g_{oi}^p(z^P) dz^P \\
&= R_{oi}\int_{\mathbf{Z}^P}[\int_{\mathbf{Z}^R}\left(\frac{\gamma\overline{\pi}_{oi}(z^P)\cdot z^R}{w_i^h}\right)^{\frac{1}{1-\gamma}} g_{oi}^R(z^R) dz^R] \cdot g_{oi}^P(z^P) dz^P \\
&= \left(\frac{\gamma}{W_i^h}\right)^{\frac{1}{1-\gamma}} \cdot R_{oi} \cdot (\overline{\pi}_{oi}^P)^{\frac{1}{1-\gamma}}[\int_{\mathbf{Z}^R}(z^R)^{\frac{1}{1-\gamma}} g_{oi}^R(z^R) dz^R] \cdot [\int_{\mathbf{Z}^P}(z^P)^{\frac{\theta}{1-\gamma}} g_{oi}^P(z^P) dz^P],
\end{aligned}
$$

where I define $\overline{\pi}_{oi}^P$ as the component in $\overline{\pi}_{oi}(z^P)$ that is independent of z^P: $\overline{\pi}_{oi}^P \equiv \frac{\overline{\pi}_{oi}(z^P)}{(z^P)^\theta}$.

Summing across all origin countries:

$$
\begin{aligned}
L_i^R &= \left(\frac{\gamma}{W_i^h}\right)^{\frac{1}{1-\gamma}} \sum_o R_{oi} \cdot (\overline{\pi}_{oi}^P)^{\frac{1}{1-\gamma}}[\int_{\mathbf{Z}^R}(z^R)^{\frac{1}{1-\gamma}} g_{oi}^R(z^R) dz^R] \cdot [\int_{\mathbf{Z}^P}(z^P)^{\frac{\theta}{1-\gamma}} g_{oi}^P(z^P) dz^P] \qquad \text{(B.22)} \\
\implies \left(\frac{\gamma}{W_i^h}\right)^{\frac{1}{1-\gamma}} &= \frac{L_i^R}{\sum_o R_{oi} \cdot (\overline{\pi}_{oi}^P)^{\frac{1}{1-\gamma}}[\int_{\mathbf{Z}^R}(z^R)^{\frac{1}{1-\gamma}} g_{oi}^R(z^R) dz^R] \cdot [\int_{\mathbf{Z}^P}(z^P)^{\frac{\theta}{1-\gamma}} g_{oi}^P(z^P) dz^P]}
\end{aligned}
$$

With this, now I proceed to characterizing $\widehat{K}_{oi}$. By Equation (B.17):

$$
\begin{aligned}
K_{oi} &= \int_{\mathbf{Z}^P} (z^P)^{\theta} V_{oi}(z^P) dz^P \\
&= \int_{\mathbf{Z}^P} (z^P)^{\theta} R_{oi} [\int_{\mathbb{Z}^R} v_{oi}(z^P, z^R) g_{oi}(z^P, z^R) dz^R] dz^P \\
&\text{(under the assumption that } z^R \text{ and } z^P \text{ are independent)} \\
&= R_{oi} \cdot \int_{\mathbf{Z}^P} (z^P)^{\theta} [\int_{\mathbb{Z}^R} v_{oi}(z^P, z^R) g^R_{oi}(z^R) dz^R] g^P_{oi}(z^P) dz^P \\
&= R_{oi} \cdot (\overline{\pi}^P_{oi})^{\frac{\gamma}{1-\gamma}} \cdot \Big(\frac{\gamma}{W^h_i}\Big)^{\frac{\gamma}{1-\gamma}} [\int_{\mathbb{Z}^R} z^{R\frac{1}{1-\gamma}} g^R_{oi}(z^R) dz^R] \cdot [\int_{\mathbf{Z}^P} (z^P)^{\theta} \cdot (z^P)^{\frac{\theta\gamma}{1-\gamma}} g^P_{oi}(z^P) dz^P] \\
&= \Big(\frac{\gamma}{W^h_i}\Big)^{\frac{\gamma}{1-\gamma}} \cdot R_{oi} \cdot (\overline{\pi}^P_{oi})^{\frac{\gamma}{1-\gamma}} \cdot [\int_{\mathbb{Z}^R} z^{R\frac{1}{1-\gamma}} g^R_{oi}(z^R) dz^R] \cdot [\int_{\mathbf{Z}^P} \cdot (z^P)^{\frac{\theta}{1-\gamma}} g^P_{oi}(z^P) dz^P] \\
&\text{(plugging in Equation (B.22))} \\
&= \frac{(L^R_i)^{\gamma} \cdot R_{oi} \cdot (\overline{\pi}^P_{oi})^{\frac{\gamma}{1-\gamma}} \cdot [\int_{\mathbb{Z}^R} z^{R\frac{1}{1-\gamma}} g^R_{oi}(z^R) dz^R] \cdot [\int_{\mathbf{Z}^P} \cdot (z^P)^{\frac{\theta}{1-\gamma}} g^P_{oi}(z^P) dz^P]}{(\sum_o R_{oi} \cdot (\pi^P_{oi})^{\frac{1}{1-\gamma}} [\int_{\mathbf{Z}^R} (z^R)^{\frac{1}{1-\gamma}} g^R_{oi}(z^R) dz^R] \cdot [\int_{\mathbf{Z}^P} (z^P)^{\frac{\theta}{1-\gamma}} g^P_{oi}(z^P) dz^P])^{\gamma}} \\
&= (L^R_i)^{\gamma} \cdot \Big(R_{oi} \cdot [\int_{\mathbb{Z}^R} z^{R\frac{1}{1-\gamma}} g^R_{oi}(z^R) dz^R] \cdot [\int_{\mathbf{Z}^P} \cdot (z^P)^{\frac{\theta}{1-\gamma}} g^P_{oi}(z^P) dz^P]\Big)^{1-\gamma} \\
&\times \frac{\Big(R_{oi} \cdot (\overline{\pi}^P_{oi})^{\frac{1}{1-\gamma}} \cdot [\int_{\mathbb{Z}^R} z^{R\frac{1}{1-\gamma}} g^R_{oi}(z^R) dz^R] \cdot [\int_{\mathbf{Z}^P} \cdot (z^P)^{\frac{\theta}{1-\gamma}} g^P_{oi}(z^P) dz^P]\Big)^{\gamma}}{(\sum_o R_{oi} \cdot (\pi^P_{oi})^{\frac{1}{1-\gamma}} [\int_{\mathbf{Z}^R} (z^R)^{\frac{1}{1-\gamma}} g^R_{oi}(z^R) dz^R] \cdot [\int_{\mathbf{Z}^P} (z^P)^{\frac{\theta}{1-\gamma}} g^P_{oi}(z^P) dz^P])^{\gamma}} \\
&= (L^R_i)^{\gamma} \cdot \Big(R_{oi} \cdot [\int_{\mathbb{Z}^R} z^{R\frac{1}{1-\gamma}} g^R_{oi}(z^R) dz^R] \cdot [\int_{\mathbf{Z}^P} \cdot (z^P)^{\frac{\theta}{1-\gamma}} g^P_{oi}(z^P) dz^P]\Big)^{1-\gamma} \cdot \Big(\frac{L^R_{oi}}{\sum_o L^R_{oi}}\Big)^{\gamma} \\
&= (L^R_i)^{\gamma} \cdot \Big(R_{oi} \cdot [\int_{\mathbb{Z}^R} z^{R\frac{1}{1-\gamma}} g^R_{oi}(z^R) dz^R] \cdot [\int_{\mathbf{Z}^P} \cdot (z^P)^{\frac{\theta}{1-\gamma}} g^P_{oi}(z^P) dz^P]\Big)^{1-\gamma} \cdot \Big(\frac{I_{oi}}{\sum_o I_{oi}}\Big)^{\gamma},
\end{aligned}
$$

where the last equation follows from that all researchers in a country are paid the same wage and the definition of I_{oi}.

Specializing the above equation to $o = i = d$, noting that because domestic firms always do R&D locally, $\Big(R_{dd} \cdot [\int_{\mathbb{Z}^R} z^{R\frac{1}{1-\gamma}} g^R_{dd}(z^R) dz^R] \cdot [\int_{\mathbf{Z}^P} \cdot (z^P)^{\frac{\theta}{1-\gamma}} g^P_{dd}(z^P) dz^P]\Big)^{1-\gamma}$ is a constant that does not respond to economic shocks and that in autarky, $\frac{I_{dd}}{\sum_o I_{od}} = 1$, we have

$$
\begin{aligned}
\widehat{K}_{dd} &= (\widehat{L^R_d})^{\gamma} \cdot \Big(\frac{I_{dd}}{\sum_o I_{od}}\Big)^{\gamma} \\
&= \Big(\frac{L^R_d}{L^{R'}_d}\Big)^{\gamma} \cdot \Big(\frac{I_{dd}}{\sum_o I_{od}}\Big)^{\gamma} \\
&= \Big(\frac{L^R_d / L^h_d}{L^{R'}_d / L^h_d}\Big)^{\gamma} \cdot \Big(\frac{I_{dd}}{\sum_o I_{od}}\Big)^{\gamma},
\end{aligned}
\tag{B.23}
$$

in which $L^{R'}_d$ is the number of high-skill workers in country d working on variety development in autarky. Since in autarky a fixed share of income is given to marketing and moreover, there is no fixed overhead for offshore R&D (as domestic firms do not pay the fixed R&D center setup cost), it follows from Equation (B.12) that $L_d^{R'}/L_d^{h'} = \frac{\gamma(\sigma-1)}{\theta-(1-\gamma)(\sigma-1)}$. The only item in $\widehat{K}_{dd}$ that is yet to be characterized is thus L_d^R/L_d^h. To this end, note that

$$\begin{aligned}
\frac{L_d^R}{L_d^h} &= \frac{\sum_o I_{od}}{\sum_o I_{od} + \sum_o C_{od}^M + \sum_{o\neq d} C_{od}^R} \\
&\quad (\text{using } C_{od}^M = \frac{1}{\sigma}\cdot\Big(\frac{\theta-(\sigma-1)}{\theta}\Big)\sum_{i,m} X_{oimd}) \\
&= \frac{\sum_o I_{od}}{\sum_o I_{od} + \frac{\theta-(\sigma-1)}{\theta\sigma}X_d + \sum_{o\neq d} C_{od}^R}
\end{aligned}$$

Note also that for $o \neq d$, given the assumption of $G_o^R(\tilde{z}^R)$ being a Pareto distribution, I can write the total fixed cost paid by firms from o opening up R&D centers in d as a function of total profit generated by these R&D centers. Letting $\hat{z}_{od}^R$ be the cutoff for firms to conduct offshore R&D in d:

$$\begin{aligned}
C_{od}^R &= c_{od}^R W_d^h E_o \int_{\hat{z}_{od}^R}^{\inf} dG_o^R(\tilde{z}^R) \\
&= c_{od}^R W_d^h E_o \int_{\hat{z}_{od}^R}^{\inf} d[1-(\frac{x}{\underline{Z}_o^R})^{-\kappa_R}] \\
&= c_{od}^R W_d^h E_o \int_{\hat{z}_{od}^R}^{\inf} d[1-(\frac{x}{\underline{Z}_o^R})^{-\kappa_R}] \\
&= c_{od}^R W_d^h E_o (\frac{\hat{z}_{oi}^R}{\underline{Z}_o^R})^{-\kappa_R}.
\end{aligned}$$

The indifference condition at the cutoff implies:

$$\begin{aligned}
&\pi_{oc}(\hat{z}_{od}^R) = \int \pi_{od}(\hat{z}_{od}^R\phi_{od}^R, z^P)dG_{od}^P(z^P) \\
&= (\hat{z}_{od}^R\phi_{od}^R)^{\frac{1}{1-\gamma}} \cdot \underbrace{(\gamma^{\frac{\gamma}{1-\gamma}} - \gamma^{\frac{1}{1-\gamma}})\Big(\frac{1}{W_d^h}\Big)^{\frac{\gamma}{1-\gamma}}[\sum_{d'}\frac{(\sigma-1)^{1+\theta}}{\theta-(\sigma-1)}\sigma^{\frac{\sigma\theta}{1-\sigma}}(c_{d'}^M W_{d'}^H)^{\frac{\sigma-1-\theta}{\sigma-1}} X_{d'}^{\frac{\theta}{\sigma-1}} P_{d'}^{\theta}(\tilde{\zeta}_{oid'})^{\theta}]^{\frac{1}{1-\gamma}}\int_{\mathbb{Z}^P}(z^P)^{\frac{\theta}{1-\gamma}}dG_{od}^P(z^P)}_{\equiv Z_{od}} \\
&\equiv (\hat{z}_{od}^R\phi_{od}^R)^{\frac{1}{1-\gamma}} \cdot Z_{od} \\
&= c_{od}^R W_d^h,
\end{aligned}$$

in which Z_{oi} is introduced to shorten the expressions. This gives:

$$(\hat{z}_{od}^R)^{\frac{1}{1-\gamma}}(\phi_{od}^R)^{\frac{1}{1-\gamma}} = \frac{c_{od}^R W_d^h}{Z_{od}}$$

Total variable profit from od is

$$\begin{aligned}
\Pi_{od} &= E_o \int_{\hat{z}_{od}\phi_{od}^R}^{\infty} \pi_{od}(z^R) dG_o^R(\tilde{z}^R) \qquad \text{(B.24)}\\
&= E_o \cdot Z_{od} \cdot \int_{\hat{z}_{od}^R}^{\infty} (z_{oi}^R \phi_{od}^R)^{\frac{1}{1-\gamma}} dG_o^R(\tilde{z}^R) \\
&(\text{using that} \int_{\hat{z}_{od}^R}^{\infty} (z_{od}^R \phi_{od}^R)^{\frac{1}{1-\gamma}} dG_o^R(\tilde{z}^R) = (\phi_{od}^R)^{\frac{1}{1-\gamma}} \frac{\kappa^R (\underline{Z}_o^R)^{\kappa^R}}{\kappa^R - \frac{1}{1-\gamma}} (\hat{z}_{od}^R)^{\frac{1}{1-\gamma}-\kappa^R}) \\
&= E_o \cdot Z_{od} \cdot (\phi_{od}^R)^{\frac{1}{1-\gamma}} \cdot \frac{\kappa^R (\underline{Z}_o^R)^{\kappa^R}}{\kappa^R - \frac{1}{1-\gamma}} \cdot (\hat{z}_{od}^R)^{\frac{1}{1-\gamma}-\kappa^R} \\
&(\text{plugging } Z_{od} = \frac{c_{od}^R W_d^h}{(\hat{z}_{od}^R)^{\frac{1}{1-\gamma}} (\phi_{od}^R)^{\frac{1}{1-\gamma}}}) \\
&= E_o \cdot c_{od}^R W_d^h \cdot (\frac{\hat{z}_{od}^R}{\underline{Z}_o^R})^{-\kappa^R} \cdot \frac{\kappa^R}{\kappa^R - \frac{1}{1-\gamma}} \\
&= C_{od}^R \cdot \frac{\kappa^R}{\kappa^R - \frac{1}{1-\gamma}}
\end{aligned}$$

i.e., we obtain the fraction of fixed RD cost as a share of profit as: $\frac{\kappa^R - \frac{1}{1-\gamma}}{\kappa^R}$, which is assumed to be below 1 for profit to be integrable.

Plugging this into Equation (B.23), I obtain

$$\begin{aligned}
\widehat{K}_{dd} &= (\frac{I_{dd}}{\sum_o I_{od}})^{\gamma} \cdot \Big(\frac{\sum_o I_{od}}{\sum_o I_{od} + \frac{\kappa^R - \frac{1}{1-\gamma}}{\kappa^R} \sum_{o \neq d} \Pi_{od} + \frac{\theta-(\sigma-1)}{\theta\sigma} X_d} \cdot \frac{\theta - (1-\gamma)(\sigma-1)}{\gamma(\sigma-1)}\Big)^{\gamma} \qquad \text{(B.25)}\\
&= (\frac{I_{dd}}{\sum_o I_{od}})^{\gamma} \cdot \Big(\frac{1}{1 + \frac{\kappa^R - \frac{1}{1-\gamma}}{\kappa^R} \cdot \frac{1-\gamma}{\gamma} \cdot \frac{\sum_{o \neq d} I_{od}}{\sum_o I_{od}} + \frac{\theta-(\sigma-1)}{\theta\sigma} \frac{X_d}{\sum_o I_{od}}} \cdot \frac{\theta - (1-\gamma)(\sigma-1)}{\gamma(\sigma-1)}\Big)^{\gamma},
\end{aligned}$$

in other words, $\widehat{K}_{dd}$ is only a function of 1) the fraction of R&D done by firms from d, $\frac{I_{dd}}{\sum_o I_{od}}$, and 2), the fraction of variable R&D expenses in income: $\frac{\sum_o I_{od}}{X_d}$.

Step 3: deriving $\widehat{\frac{W_d^h}{X_d}}$ and $\widehat{\frac{W_d^l}{X_d}}$.

I now deriving ratios between baseline and autarky values for $\frac{W_d^h}{X_d}$ and $\frac{W_d^l}{X_d}$. Note that because the supply of high- and low-skill workers are exogenous, $\widehat{\frac{W_d^h}{X_d}} = \widehat{\frac{W_d^h L_d^h}{X_d}}$ and $\widehat{\frac{W_d^l}{X_d}} = \widehat{\frac{W_d^l L_d^l}{X_d}}$, i.e., we only need to derive the change in income share of high- and low-skill workers.

From Equations (B.15) and (B.24), we have $\frac{W_d^h L_d^h}{X_d} = \frac{\sum_o I_{od} + \frac{\kappa^R - \frac{1}{1-\gamma}}{\kappa^R} \cdot \frac{1-\gamma}{\gamma} \cdot \sum_{o \neq d} I_{od} + \frac{\theta-(\sigma-1)}{\theta\sigma} X_d}{X_d}$, In autarky, this ratio collapses to $\frac{\theta-(1-\gamma)(\sigma-1)}{\sigma\theta}$, so we have

$$\widehat{\frac{W_d^h}{X_d}} = \frac{\sigma\theta}{\theta - (1-\gamma)(\sigma-1)} \cdot [\frac{\sum_o I_{od}}{X_d} + \frac{\kappa^R - \frac{1}{1-\gamma}}{\kappa^R} \cdot \frac{1-\gamma}{\gamma} \cdot \frac{\sum_{o \neq d} I_{od}}{X_d} + \frac{\theta - (\sigma-1)}{\theta\sigma}]. \qquad \text{(B.26)}$$

From Equation (B.15), we have $\frac{W_d^l L_d^l}{X_d} = \frac{Y_d}{X_d}$. In autarky, this ratio is simply $\frac{\sigma-1}{\sigma}$, so

$$\frac{\widehat{W_d^l}}{X_d} = \frac{\sigma}{\sigma-1}\frac{Y_d}{X_d}. \tag{B.27}$$

Step 4: Putting all together

Combining Equations (B.21), (B.25), (B.26), and (B.27), I obtain:

$$\begin{aligned}
\frac{\widehat{X_d}}{P_d} &= (\frac{\widehat{W_d^l}}{P_d}) \times (\frac{\widehat{X_d}}{W_d^l}) \qquad\qquad (B.28)\\
&= (\frac{\widehat{W_d^h}}{X_d})^{\frac{\theta-\sigma+1}{(1-\sigma)\theta}} \times (\frac{X_{dddd}}{\sum_m X_{ddmd}})^{-\frac{1}{\theta}} \times (\frac{\sum_{o,m} X_{odmd}}{X_d})^{-\frac{1}{\theta}} \times (\frac{\sum_m X_{ddmd}}{\sum_{o,m} X_{odmd}})^{-\frac{1}{\theta}} \times (\widehat{K_{dd}})^{\frac{1}{\theta}} \times (\frac{\widehat{X_d}}{W_d^l})\\
&= (\frac{X_{dddd}}{\sum_m X_{ddmd}})^{-\frac{1}{\theta}} \times (\frac{\sum_{o,m} X_{odmd}}{X_d})^{-\frac{1}{\theta}} \times (\frac{\sum_m X_{ddmd}}{\sum_{o,m} X_{odmd}})^{-\frac{1}{\theta}} \times (\frac{I_{dd}}{\sum_o I_{od}})^{\frac{\gamma}{\theta}}\\
&\quad \times \Big(\frac{1}{1+\frac{\kappa^R-\frac{1}{1-\gamma}}{\kappa^R}\cdot\frac{1-\gamma}{\gamma}\cdot\frac{\sum_{o\neq d} I_{od}}{\sum_o I_{od}} + \frac{\theta-(\sigma-1)}{\theta\sigma}\frac{X_d}{\sum_o I_{od}}} \cdot \frac{\theta-(1-\gamma)(\sigma-1)}{\gamma(\sigma-1)}\Big)^{\frac{\gamma}{\theta}}\\
&\quad \times \Big(\frac{\sigma\theta}{\theta-(1-\gamma)(\sigma-1)}\cdot[\frac{\sum_o I_{od}}{X_d} + \frac{\kappa^R-\frac{1}{1-\gamma}}{\kappa^R}\cdot\frac{1-\gamma}{\gamma}\cdot\frac{\sum_{o\neq d} I_{od}}{X_d} + \frac{\theta-(\sigma-1)}{\theta\sigma}]\Big)^{\frac{\theta-\sigma+1}{(1-\sigma)\theta}} \times \frac{X_d}{Y_d} \times \frac{\sigma-1}{\sigma}\\
&\equiv (\frac{X_{dddd}}{\sum_m X_{ddmd}})^{-\frac{1}{\theta}} \times (\frac{\sum_{o,m} X_{odmd}}{X_d})^{-\frac{1}{\theta}} \times (\frac{\sum_m X_{ddmd}}{\sum_{o,m} X_{odmd}})^{-\frac{1}{\theta}} \times (\frac{I_{dd}}{\sum_o I_{od}})^{\frac{\gamma}{\theta}} \times \frac{X_d}{Y_d} \times f(\frac{\sum_o I_{od}}{X_d}, \frac{I_{dd}}{\sum_o I_{od}}),
\end{aligned}$$

in which $f(\frac{\sum_o I_{od}}{X_d}, \frac{I_{dd}}{\sum_o I_{od}})$ is defined to be a function of only model elasticities and two measures, $\frac{\sum_o I_{od}}{X_d}$ and $\frac{I_{dd}}{\sum_o I_{od}}$.[15]

This completes the proof of Proposition 1. This proof holds when z^R and z^P are independent. But accommodate any arbitrary distributions for $G_{oi}^P(z^P)$ and flexible Pareto distributions for $G_o^R(\tilde{z}^R)$. $\square$

B.4 The Gains from Openness in the Literature

I compare Equation (B.28) to the gains from openness in Ramondo and Rodríguez-Clare (2013) and Arkolakis et al. (2018), both of which feature trade and offshore production, but not offshore R&D. Their formulas are given by the following:

$$GO_d = \overbrace{\underbrace{[(\frac{X_{ddd}}{X_d})^{\alpha}(\frac{\sum_m X_{dmd}}{X_d})^{\beta}]}_{\text{Direct Effect (Ramondo et al., 2015)}} \times [(\frac{X_d}{Y_d})^{\delta}]}^{\text{Arkolakis et al. (2018)}} - 1, \tag{B.29}$$

where α, β, and δ are model elasticities that differ due to exact setups of models. Because Ramondo and Rodríguez-Clare (2013) and Arkolakis et al. (2018) assume R&D takes place at the headquarters, X_{oimd} is not defined. Instead, the gains from openness is a function of X_{omd}, defined on headquarter (R&D) country o, manufacturing location m, and consumption location d. The direct effect in Equation

[15] $f(\frac{\sum_o I_{od}}{X_d}, \frac{I_{dd}}{\sum_o I_{od}})$ collects the remaining terms and can be rearranged to be:

$$f(\frac{\sum_o I_{od}}{X_d}, \frac{I_{dd}}{\sum_o I_{od}}) = [1 + \frac{(1-\gamma)\kappa^R - 1}{\gamma\kappa^R}(1 - \frac{I_{dd}}{\sum_o I_{od}}) + \frac{\theta-(\sigma-1)}{\theta\sigma}\frac{X_d}{\sum_o I_{od}}]^{\frac{1-\gamma}{\theta}-\frac{1}{\sigma-1}}\left(\frac{\sum_o I_{od}}{X_d}\right)^{\frac{1}{\theta}-\frac{1}{\sigma-1}}\underbrace{[\theta-(1-\gamma)(\sigma-1)]^{\frac{1}{\sigma-1}-\frac{1-\gamma}{\theta}}\gamma^{-\frac{\gamma}{\theta}}(\sigma-1)^{1-\frac{\gamma}{\theta}}\theta^{\frac{1}{\theta}-\frac{1}{\sigma-1}}\sigma^{\frac{1}{\theta}-\frac{1}{\sigma-1}-1}}_{\text{constant}}$$

(B.29) captures the effect of openness on the real wage, whereas the indirect effect takes into account the difference between the real wage and real income.

To compare this to my model, note that when ϕ^P_{oim} is independent of i (e.g., when $s = 0$ in my quantification), where a product is invented has no bearing on its production location. In this case, the first four terms of Equation (B.28) is

$$(\frac{\sum_i X_{didd}}{\sum_{m,i} X_{dimd}})^{-\frac{1}{\theta}} \times (\frac{\sum_{o,m} X_{odmd}}{X_d})^{-\frac{1}{\theta}} \times (\frac{I_{dd}}{\sum_o I_{o,d}})^{-\frac{1}{\theta}} \times (\frac{I_{dd}}{\sum_o I_{od}})^{\frac{\gamma}{\theta}}.$$

$(\frac{\sum_i X_{didd}}{\sum_{m,i} X_{dimd}})^{-\frac{1}{\theta}}$ captures the importance of importing good produced in other countries, as does the first term in Equation (B.29). $(\frac{\sum_{o,m} X_{odmd}}{X_d})^{-\frac{1}{\theta}}$ in the above expression captures the importance of access to goods invented in other countries and corresponds to the second term in Equation (B.29). $(\frac{I_{dd}}{\sum_o I_{o,d}})^{-\frac{1}{\theta}}$ is new to the literature. It captures the importance of foreign firms in R&D at host d, which to the first order is captured by the share in total R&D expenditures of foreign firms. $(\frac{I_{dd}}{\sum_o I_{o,d}})^{\frac{\gamma}{\theta}}$, on the other hand, captures an offsetting effect from the labor market—foreign R&D centers compete with domestic firms for high-skill workers and crowd out the R&D of domestic firms. The importance of this channel is also captured by the foreign R&D share, only with the elasticity scaled by γ, the researcher share in innovation.

Finally, the indirect effect in my model ($\frac{X_d}{Y_d} \times f(\frac{\sum_o I_{od}}{X_d}, \frac{I_{dd}}{\sum_o I_{od}})$), which takes into account changes in income composition due to openness, also differs from that in Equation (B.29). Aside from elasticities, the main difference is that whereas in Equation (B.29) the indirect effect only depends on $\frac{X_d}{Y_d}$, in my model it also depends on the share of R&D in total income and the foreign share of domestic R&D. Both differences arise because endogenous offshore R&D redistributes the innovation rent between headquarters and R&D hosts.

C Quantification

This section provides additional details on quantification.

C.1 Additional Data for Quantification

Sample countries. The model economy consists of the same 37 countries as in the empirical section. See section A.3 of this appendix for a discussion on the sample selection. Among these countries, Ireland is frequently dubbed as a 'tax-haven' country, in which reported financial statistics might be unreliable. I include Ireland in the sample because this allows me to not take a stand on how to redirect the foreign linkages that run through it. That said, if these links are simply assumed to be non-existence, measured openness for other countries will be similar.

Country-specific openness measures. In disciplining host specific barriers to inward MNC activities, I use three targets: the share of production by foreign firms, the share of innovation by foreign R&D centers, and the share of foreign R&D centers in total R&D center counts. These targets are calculated from the firm-level data described in Section A.3, with the following modifications. First, given my interpretation of the model as for manufacturing, I focus only on manufacturing firms in calculating these ratios. Second, all firms in the model carry out some R&D, but not all firms in the data are granted patents. Instead of using the joint sample to construct the ratios, I use the full financial sample to calculate inward offshore production measure, and the full R&D sample to calculate the two (extensive and intensive margins) inward offshore R&D measures. Third, for the three countries for which the financial data have relatively low coverage (Mexico, Turkey, U.S.), instead of aggregating the firm-level data, I use the aggregate shares from Ramondo et al. (2015). Finally, recent research has found that patents granted by the Chinese patent office to local firms are systematically less likely to be global patents than the patents they grant to foreign firms, suggesting differential treatments based on where firms are from (Holmes et al., 2015). To avoid biases arising from potential discriminatory treatments, I exclude all patents issued by the Chinese patent authority when calculating the inward offshore R&D ratio for China. Table C.1 reports the three openness measures.

Endowment distributions. The calibration uses the World Management Survey (Bloom et al., 2012) and an internationally comparable cognitive ability score database (Hanushek and Woessmann, 2012). I take the exponent of the innovation management score so that its distribution has a right tail that resembles the firm size distribution. I compute the mean, standard deviation, and skewness of the exponent of scores in each country, which then serve as an input to the calibration. The distribution statistics for cognitive test scores are from Hanushek and Woessmann (2012). These statistics include the average cognitive score for high school students in a country, the share of students achieving 'top' performance, and the share of student achieving 'basic' performance. Thresholds for 'top' and 'basic' performance are defined based on a common set of standards so these shares are comparable internationally.

A few countries in the sample are not included in the World Management Survey. I impute their management distribution statistics by regressing each statistics on income, R&D share, and geographic-region fixed effects, where geographic regions are at sub-continent level. The inclusion of income is motivated by the finding in Bloom et al. (2012) that management knowhow explains a substantial share of cross-country income differences; the inclusion of geographic-region fixed effects is meant to capture management practice differences driven by culture. The R^2 of these regressions are all above 0.85.

Table C.1 reports these statistics for all countries.

C.2 Parameterization

Bilateral trade costs. I assume that the iceberg trade costs are symmetric and that trade cost with own country is 1, i.e., $\tau_{mm} = 1, \forall m$ and $\tau_{md} = \tau_{dm}, \forall m \neq d$. Under these assumptions, the approach in Head and Ries (2001) generalizes to my setting.

To this end, using Equation (B.12):

$$X_{oimd} = X_{oid}\psi_{oimd} \equiv B_d \frac{1}{N}\left(\frac{T_m \phi^P_{oim}}{W^l_m \tau_{md}}\right)^\theta \cdot B_{o,i},$$

where B_d and $B_{o,i}$ are functions of the equilibrium objects in d and in (o,i) introduced to shorten notations. Using this expression, we have:

$$\begin{aligned}
&\frac{\sum_{o,i} X_{oimd}}{\sum_{o,i} X_{oimm}} \cdot \frac{\sum_{o,i} X_{oidm}}{\sum_{o,i} X_{oidd}} \\
&= \frac{B_d(\frac{1}{W^l_m \tau_{md}})^\theta \sum_{o,i} \phi^P_{oim} B_{o,i}}{B_m(\frac{1}{W^l_m \tau_{mm}})^\theta \sum_{o,i} \phi^P_{oim} B_{o,i}} \times \frac{B_m(\frac{1}{W^l_d \tau_{dm}})^\theta \sum_{o,i} \phi^P_{oid} B_{o,i}}{B_d(\frac{1}{W^l_d \tau_{dd}})^\theta \sum_{o,i} \phi^P_{oid} B_{o,i}} \\
&= (\tau_{md})^{-2\theta}.
\end{aligned}$$

Notice that although the flow items such as X_{oimd} are not observable, $\sum_{o,i} X_{oidm}$ is simply the total sales from d to m, which is observable.

Slightly abusing notations, I write $\tau_{md} = (\frac{\sum_{o,i} X_{oimd}}{\sum_{o,i} X_{oimm}} \cdot \frac{\sum_{o,i} X_{oidm}}{\sum_{o,i} X_{oidd}})^{-\frac{1}{2\theta}} = (\frac{X_{md}}{X_{mm}} \times \frac{X_{dm}}{X_{dd}})^{-\frac{1}{2\theta}}$, where X_{md} denotes the sales from m to d as in the gravity literature. I obtain these sales for the aggregated manufacturing sector from the World Input Output Database.

Relating production efficiency to innovation efficiency. To discipline the relationship between firms' innovation and production management efficiency, I use micro data from the World Management

Table C.1: Calibration Targets: Country Characteristics

	Income and Openness					Innovation Mgt. Dist.			Talent Dist.		
ISO	$\frac{X_m}{P_m}$	$\frac{\sum_i V_{oi}}{\sum_{o,i} V_{oi}}$	$\frac{\sum_{o\neq m} Y_{om}}{\sum_o Y_{om}}$	$\frac{\sum_{o\neq i} V_{oi}}{\sum_o V_{oi}}$	$\frac{\sum_{o\neq i} R_{oi}}{\sum_o R_{oi}}$	mean	std	skew.	mean	% basic	% top
AUS	0.81	1.03	60.87	39.78	33.38	6.43	3.64	1.88	5.09	93.84	11.24
AUT	0.76	0.59	63.09	45.49	41.70	6.91	4.14	2.15	5.09	93.11	9.74
BEL	0.81	0.45	82.31	58.85	54.32	7.14	4.48	2.27	5.04	93.13	9.38
BGR	0.29	0.05	32.02	19.02	17.04	6.80	3.93	1.50	4.79	76.53	8.30
BRA	0.26	1.10	19.17	57.58	46.73	5.26	3.33	2.34	3.64	33.85	1.09
CAN	0.73	1.20	35.35	52.21	47.55	8.40	6.09	2.01	5.04	94.84	8.33
CHE	0.90	1.51	36.21	41.68	44.57	7.40	4.88	2.41	5.14	91.85	13.36
CHN	0.17	19.84	23.93	41.52	21.93	5.94	2.74	1.85	4.94	93.48	8.34
CZE	0.52	0.28	62.14	26.59	22.89	6.47	3.44	1.33	5.11	93.07	12.22
DEU	0.77	6.85	46.54	28.16	29.94	8.21	5.25	2.20	4.96	90.60	10.52
DNK	0.80	0.40	63.62	38.37	38.26	7.28	5.46	2.75	4.96	88.78	8.75
ESP	0.74	1.15	49.89	22.09	16.93	5.29	3.46	2.21	4.83	85.88	7.93
EST	0.46	0.03	53.49	28.48	34.27	6.11	3.71	2.14	5.19	97.32	9.46
FIN	0.73	0.80	28.05	24.86	18.43	6.84	4.79	2.51	5.13	95.78	12.39
FRA	0.82	3.89	42.68	26.58	40.34	6.43	4.25	2.52	5.04	92.62	8.49
GBR	0.70	1.85	87.73	62.95	53.21	7.36	5.04	2.24	4.95	92.88	8.79
GRC	0.59	0.09	24.51	50.08	80.00	5.63	3.70	1.89	4.61	79.77	4.24
HRV	0.51	0.03	28.34	53.37	60.87	5.45	3.21	1.93	4.70	83.35	4.76
HUN	0.48	0.08	47.70	61.09	38.01	6.56	3.57	1.37	5.05	94.11	10.28
IRL	1.13	0.69	77.34	73.67	68.78	7.14	6.73	3.86	4.99	91.37	9.40
ITA	0.79	1.23	33.85	43.84	29.03	6.47	4.15	2.17	4.76	87.54	5.45
JPN	0.63	11.76	6.65	1.99	17.42	7.83	5.57	1.82	5.31	96.67	16.76
KOR	0.60	4.87	9.71	6.72	14.33	7.06	4.42	1.92	5.34	96.16	17.84
LTU	0.55	0.04	46.93	20.30	16.76	6.70	4.59	2.44	4.78	89.07	2.97
LVA	0.45	0.02	38.21	6.07	31.25	5.94	3.45	2.05	4.80	86.95	4.99
MEX	0.34	0.33	17.80	55.01	64.30	6.90	4.43	1.66	4.00	48.93	0.88
NLD	0.80	2.01	93.21	30.27	29.58	7.34	4.79	2.37	5.11	96.54	9.16
NOR	1.34	0.39	46.67	32.27	22.60	8.81	7.72	3.54	4.83	89.44	5.61
POL	0.51	0.47	49.96	14.61	24.63	7.25	4.60	1.73	4.85	83.76	9.86
PRT	0.55	0.21	46.40	26.20	40.78	5.38	2.99	1.94	4.56	80.27	3.16
ROU	0.41	0.06	64.47	52.75	47.65	6.63	3.68	1.41	4.56	78.05	4.56
RUS	0.42	2.32	35.17	11.63	13.63	6.82	3.96	1.51	4.92	88.35	8.05
SVK	0.53	0.06	67.86	27.71	44.79	6.89	4.06	1.54	5.05	90.55	11.16
SVN	0.51	0.08	36.61	17.17	19.96	5.40	3.14	1.90	4.99	93.89	6.12
SWE	0.79	1.13	48.39	43.20	34.43	7.06	4.17	1.99	5.01	93.94	8.76
TUR	0.54	0.73	6.34	20.12	23.86	5.86	2.58	2.09	4.13	58.23	3.92
USA	1.00	32.38	15.29	15.77	13.62	10.94	8.15	2.15	4.90	91.82	7.33

Note: This table reports the country-level statistics used as targets in parameterization. The first set of targets are on the income and openness of countries, corresponding to Panel B of Table 7. These columns are: real income (U.S. normalized to 1), the contribution (%) of a country to the world world R&D (based on the origin of firms), the share (%) of domestic production by foreign firms, the share (%) of domestic R&D by foreign R&D centers, and the share (%) of foreign R&D centers among all active R&D centers in a host. '*Innovation Mgt. Dist.*' refers to the sample distribution statistics constructed from the World Management Survey as described in Section C.1. '*Talent Dist.*' refers to the talent distribution statistics from Hanushek and Woessmann (2012), in which '*Mean*' is the mean score for a country, and '*% basic*' and '*% topc*' are shares of students achieving 'basic' and 'top' performance, respectively. The performance standards are common across countries..

Survey to estimate the following equation:

$$\text{Prob}(z^P \in H | z^R) = \frac{\exp(\delta_0 + \delta_1 \times z^R)}{1 + \exp(\delta_0 + \delta_1 \times z^R)}. \tag{C.1}$$

This dataset covers around 11338 firms from 34 countries. I classify a firm as being a H type, if its production management scores falls in the top 5% in the sample. Table C.2 presents summary statistics on this score and the indicator for H type. Table C.3 presents results from a logit regression of Equation (C.1). Column 1 uses no fixed effects whereas Column 2 includes country fixed effects. Both specifications find positive and statistically significant coefficient, consistent with strong correlation between innovation efficiency and production efficiency. Based on the estimates, I set $\delta_1 = 0.21$ and $\delta_0 = -5$.

For illustration, Figure C.1 plots the parameterized distribution for the U.S. About 12% of the American firms end up being a H type.

Table C.2: Firm Management Score Summary Statistics

Variable	Obs	Mean	Std. Dev.	Min	Max
z^R	11338	6.68	4.92	1	54.6
$\mathbb{1}(z^P \in G_H^P)$	11340	.051	0.22	0	1

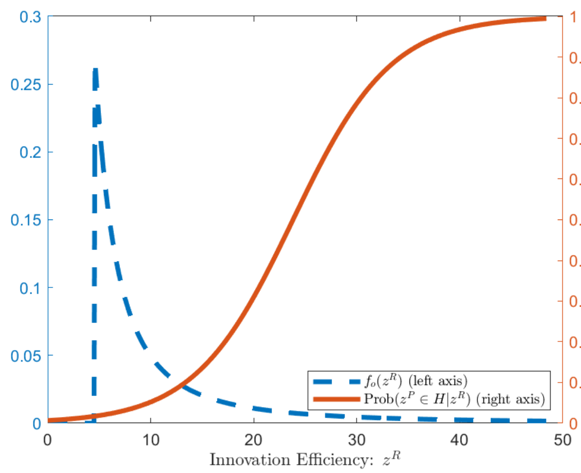
Notes: This table presents summary statistics for firm-level innovation management scores and the indicator for whether a firm is in the top 5% production efficiency.

Table C.3: Estimates for δ_0 and δ_1

	(1)	(2)
	$\mathbb{1}(z^P \in G_H^P)$	
z^R	0.213***	0.210***
	(0.00719)	(0.00797)
cons	-4.921***	-
	(0.0950)	-
N	11338	10637
pseudo R^2	0.251	0.281
country FE		yes

Notes: This table presents results from a Logit regression of the high production efficiency indicator $\mathbb{1}(z^P \in G_H^P)$, on firms' innovation efficiency, z^R. The high production efficiency indicator takes a value of 1 if the production management score of a firm is in the top 5% in the world. The second column controls for country fixed effects. Standard errors are in parentheses. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Figure C.1: The U.S. Firm Knowhow Distribution



Notes: The horizontal axis is the support of the innovation efficiency for firms from the U.S. The dashed line (left axis) is the probability density function of the innovation efficiency. The solid line (right axis) is the probability that a firm with a given z^R obtains a draw from $G_H^P(z^P)$, i.e., the value of Equation (C.1).

Geographic friction parameters. Table C.4 reports the 22 targets used to pin down 17 geographic parameters. The left panel is the data, corresponding to regressions discussed in Section 2.3. The brackets are the 95% confidence intervals of these coefficients. The right panel reports the corresponding regression coefficients estimated with the exact same specifications using model-simulated data. The coefficients in general match the data counterparts closely. I highlight using underline the three coefficients that are outside the 95% confidence interval of the original coefficient. In all three cases, they are not too far off.

Table C.4: Calibration Targets: Geographic Friction Parameters

	A. Data					**B. Model**				
	Headquarter Effect (Table 5)			Colocation (Tables 5 and 4)		Headquarter Effect			Colocation	
Dependent var.	R&D indicator	log (R&D)	log(sales)	log(sales)	log(sales)	R&D indicator	log (patents)	log(sales)	log(sales)	log(sales)
$\log(\text{dist})_{oh}$	-0.002	-0.129	-0.282	-0.253		-0.0001	-0.118	-0.271	-0.271	
	[-0.003,-0.000]	[-0.197,-0.062]	[-0.337,-0.227]	[-0.291,-0.214]						
Common language$_{oh}$	0.020	0.258	0.162	0.094		0.018	0.285	0.116	0.0953	
	[0.011,0.029]	[0.117,0.399]	[0.030,0.294]	[0.021,0.168]						
Contiguity$_{oh}$	0.002	0.106	0.185	0.174		0.003	0.0741	0.176	0.172	
	[-0.001,0.006]	[-0.036,0.248]	[0.059,0.311]	[0.103,0.245]						
Colonial tie$_{oh}$	0.002	0.029	0.153	0.129		**0.009**	0.0117	0.121	0.111	
	[-0.006,0.009]	[-0.102,0.160]	[-0.002,0.308]	[0.057,0.201]						
R&D center indicator				1.198	1.042				**1.106**	**1.122**
				[1.147,1.259]	[0.991,1.092]					
$\log(\text{dist}_{fh,t})$					-0.0235					-0.0048
					[-0.073,0.025]					
Common language$_{fh,t}$					0.220					0.210
					[0.120,0.319]					
Contiguity$_{fh,t}$					0.143					0.141
					[0.046,0.239]					
Colonial tie$_{fh,t}$					0.090					0.153
					[0.001,0.179]					
Firm FE	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y
Host FE	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y
Home-host FE	N	N	N	N	Y	N	N	N	N	Y

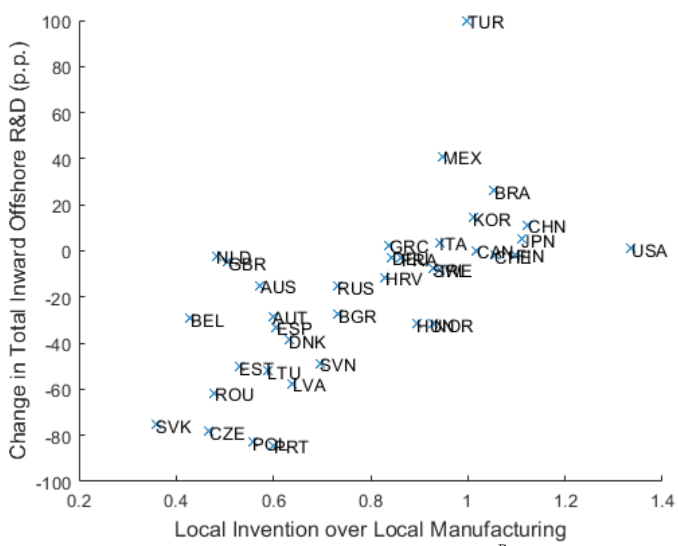
Notes: This table reports moments that pin down geographic parameters. Panel A is a collection of reduced-form regressions, reported in Section 2.3 of the text. The bracket under each coefficient is its 95% confidence interval. The right panel reports results from regressions using model-simulated data. Coefficients highlighted by underscores are outside the 95% confidence interval.

C.3 Additional Counterfactual Experiments

A prediction of the general equilibrium model is that an increase in the manufacturing productivity of foreign firms in a host has an ambiguous impact on inward offshore R&D. On the one hand, due to the colocation effect, an increase in foreign production can draw in additional R&D; on the other hand, the general equilibrium force will push the country to specialize in production, which increases the wage and deters more foreign R&D centers from entry. The strength of the general equilibrium force depends on the share of domestic production devoted to locally invented products. It will be stronger, if most production in a host is for varieties invented elsewhere, in which case an increase in production efficiency of foreign firms will lead to a substantial increase in the total production. Conversely, if a country currently does not produce varieties invented elsewhere, then the increase in offshore production will be relatively small and the partial equilibrium effect will dominate.

To see which force is stronger, I consider a set of experiments, in which I increase the openness of each country to inward production by increasing ϕ_m^P by 10 p.p. Figure C.2 plot the results. The vertical axis is the percentage point change in inward offshore R&D shares. About half of the countries have a negative value—for them, the general equilibrium effect dominates. The horizontal axis is the share of domestic invention over domestic production in a host, which captures the relative importance of local invention for local production. As anticipated, countries producing more relative to their own invention—those in the left part of the figure—experiences bigger decreases in offshore R&D.

Figure C.2: Inward Production Efficiency and Inward R&D



Notes: The figure plots the change in offshore R&D in response to a 10 p.p. increase in ϕ_m^P in each country m. The vertical axis is the percentage point change in inward offshore R&D; the horizontal axis is the baseline equilibrium value of $\frac{\sum_{o,m',d} X_{omm'd}}{\sum_{o,i,d} X_{oimd}}$, i.e., the total sales of all varieties invented in m over total sales of all goods produced in m.

C.4 Numerical Implementation

Table 7 of the text summarizes the three categories of parameters and the corresponding moments that identify them. I design a nested fixed point algorithm based on the feature of this problem to pin down the parameters efficiently. Before explaining the algorithm, it is useful to review the conditions characterizing the competitive equilibrium.

As discussed in Section B of this Appendix, Equations (B.13), (B.14), (B.15), and (B.16) jointly characterize a fixed point system in wages and occupation choice $\{W_d^h, W_d^l, \hat{a}_d : d = 1, ...N\}$, prices $\{P_d : d = 1, ...N\}$, and aggregate expenditures $\{X_d : d = 1, ...N\}$. Furthermore, the moments that identify (subject to normalization) manufacturing TFP, $\{T_m | m = 1, ..., N\}$, measure of firms, $\{E_o | o = 1, .., N\}$, and host fixed effects in offshore activities $\{\phi_m^P, \phi_i^R, \phi_i^{cR} | i, m = 1, .., N\}$ are characterized by the following system of equations:

$$
\begin{aligned}
T_m &: \frac{X_m}{P_m} = \widehat{\text{Real GDP}}_m, \ m = 1, ...N \\
E_o &: \frac{\sum_i I_{oi}}{\sum_{o,i} I_{oi}} = \widehat{\frac{\sum_i I_{oi}}{\sum_{o,i} I_{oi}}}, \ o = 1, ...N \\
\phi_m^P &: \frac{\sum_{o \neq m} Y_{om}}{\sum_o Y_{om}} = \widehat{\frac{\sum_{o \neq m} Y_{om}}{\sum_o Y_{om}}}, \ m = 1, ...N \\
\phi_i^R &: \frac{\sum_{o \neq i} I_{oi}}{\sum_o I_{oi}} = \widehat{\frac{\sum_{o \neq i} I_{oi}}{\sum_o I_{oi}}}, \ i = 1, ...N \\
\phi_i^{cR} &: \frac{\sum_{o \neq i} R_{oi}}{\sum_o R_{oi}} = \widehat{\frac{\sum_{o \neq i} R_{oi}}{\sum_o R_{oi}}}, \ i = 1, ...N
\end{aligned}
\tag{C.2}
$$

The the right-hand sides of these equations are the data. The left-hand side are their model counterparts. I write in front of each equation a fundamental variable (e.g., T_m) to stress that the model predictions are a function of these fundamentals.

Together with Equations (B.13), (B.14), (B.15), and (B.16), Equation (C.2) characterizes a fixed point in the model fundamentals and endogenous outcomes, such that: 1) the solution to the fixed point problem is a competitive equilibrium; 2), the solution to the fixed point problem ensures that the model matches the data exactly as specified in (C.2). I implement the following algorithm.

1. Choose z_H^P and κ^P.
   (a) Choose the 17 parameters governing bilateral frictions and colocation pattern: $\{\overrightarrow{\beta^{P,om}}, \overrightarrow{\beta^{P,im}}, \overrightarrow{\beta^R}, \overrightarrow{\beta^{cR}}\}$ and s
      i. Solve Equations (B.13), (B.14), (B.15), (B.16), and (C.2) jointly for the following: $\{W_d^h, W_d^l, \hat{a}_d : d = 1, ...N\}$, $\{P_d : d = 1, ...N\}$,$\{X_d : d = 1, ...N\}$, $\{T_m | m = 1, ..., N\}$, $\{E_o | o = 1, .., N\}$, and $\{\phi_m^P, \phi_i^R, \phi_i^{cR} | i, m = 1, .., N\}$.
      ii. Simulate 5e4 firms. I assign the number of firms from a country to be proportional to its size and draw $\tilde{z}^R$ for these firms from its calibrated knowhow distribution.
      iii. Solve for the optimal offshore R&D and production decision of these firms. Then estimate the same specifications as in the left panel of Table C.4 using the model-simulated data.
      iv. Evaluate the objective function $f = \sum_{k=1}^{22} (\frac{x_k - \widehat{x}_k}{\hat{\sigma}_k})^2$, where $x_k, k = 1, ..., 22$ is a model-based regression coefficient, $\widehat{x}_k$ is the empirical estimate, and $\hat{\sigma}_k$ is the standard error of $\hat{x}_k$.
   (b) If the choice of $\{\overrightarrow{\beta^{P,om}}, \overrightarrow{\beta^{P,im}}, \overrightarrow{\beta^R}, \overrightarrow{\beta^{cR}}\}$ and s minimize f defined above, proceed to step 2, otherwise return to Step 1.(a) and try a different set of parameter values.
2. Compare the model-based firm size distribution to its data counterparts (Panel A of Table 7). If they are close enough, exit; otherwise return to Step 1.

Some additional details on implementing the above algorithm. First, in searching over the space of the 17 geographic parameters, I try multiple starting points using two algorithms implemented by Knitro, interior point and active-set. Both algorithms give similar results.

Second, in solving for the fixed point problem in Step 1.(a).i, for a given set of fundamentals and aggregate prices and wages, I evaluate the endogenous objects in Equations (B.13), (B.14), (B.15), (B.16), and (C.2). These objects can be found by sequentially calculating Equations (B.9), (B.10), (B.11), and (B.12). Most of these expressions are analytical and hence can be directly evaluated. For the ones that cannot be analytically evaluated, I approximate their values numerically as follows. First, the cutoff for offshore R&D, $\hat{z}^R_{oi}$ is given by the implicit function in Equation (B.10). I analytically integrate over $\pi_{oi}(z^P, z^R)$ to obtain function $\pi^R_{oi}(z^R)$,[16] and then determine the cutoff $\hat{z}^R_{oi}$ as the indifference point for offshore R&D $\pi^R_{oi}(\tilde{z}^R \phi^R_{oi}) = c^R_{oi} W^h_i$ using the Brent method. Second, $V_{oi}(z^P)$ in Equation (B.12) is a an integration of $v_{oi}(z^P, z^R)$ over $\mathbb{Z}^R$ space; moreover, $V_{oi}(z^P)$ itself becomes an integrand for X_{oid} and P_d. For this step analytical integration is not available. I approximate for $V_{oi}(z^P)$ numerically using an adaptive Cash-Karp algorithm. The number of such numerical approximations for each evaluation of Equations (B.13), (B.15), (B.16), (B.14), and (C.2) increases quadratically with the number of countries in the sample; solving the equation systems and then finding the best-fit geographic parameters requires thousands of such evaluations. Step 1.(a).i is implemented in C++ to speed up the computation.

References

Akcigit, Ufuk, Salomé Baslandze, and Stefanie Stantcheva, "Taxation and the international mobility of inventors," *American Economic Review*, 2016, *106* (10), 2930–81.

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

Bloom, Nicholas, Christos Genakos, Raffaella Sadun, and John Van Reenen, "Management Practices Across Firms and Countries," *The Academy of Management Perspectives*, 2012, *26* (1), 12–33.

Cravino, Javier and Andrei A Levchenko, "Multinational Firms and International Business Cycle Transmission," *The Quarterly Journal of Economics*, 2017, *132* (2), 921–962.

Griliches, Zvi, "Patent Statistics as Economic Indicators: a Survey," in "R&D and productivity: the econometric evidence," University of Chicago Press, 1998, pp. 287–343.

Hanushek, Eric A and Ludger Woessmann, "Do Better Schools Lead to More Growth? Cognitive Skills, Economic Outcomes, and Causation," *Journal of Economic Growth*, 2012, *17* (4), 267–321.

Head, Keith and John Ries, "Increasing Returns Versus National Product Differentiation as an Explanation for the Pattern of US-Canada Trade," *American Economic Review*, 2001, *91* (4), 858–876.

Holmes, Thomas J, Ellen R McGrattan, and Edward C Prescott, "Quid pro quo: Technology capital transfers for market access in China," *The Review of Economic Studies*, 2015, *82* (3), 1154–1193.

Kalemli-Ozcan, Sebnem, Bent Sorensen, Carolina Villegas-Sanchez, Vadym Volosovych, and Sevcan Yesiltas, "How to Construct Nationally Representative Firm Level Data from the Orbis Global Database: New Facts and Aggregate Implications," December 2019.

Lee, David L, Justin McCrary, Marcelo J Moreira, and Jack Porter, "Valid t-ratio Inference for IV," *arXiv preprint arXiv:2010.05058*, 2020.

Mayer, Thierry and Soledad Zignago, "Notes on CEPII?s distances measures: The GeoDist database," 2011.

Park, Walter G, "International Patent Protection: 1960–2005," *Research Policy*, 2008, *37* (4), 761–766.

Ramondo, Natalia and Andrés Rodríguez-Clare, "Trade, Multinational Production, and the Gains from Openness," *Journal of Political Economy*, 2013, *121* (2), 273–322.

_ , _ , and Felix Tintelnot, "Multinational Production: Data and Stylized Facts," *The American Economic Review*, 2015, *105* (5), 530–536.

[16]Under the assumption that z^P is drawn from two Pareto distributions probabilistically, this integration is analytical:

$$\begin{aligned}\pi^R_{oi}(z^R) &= \int_{\mathbf{Z}^P} \pi^R_{oi}(z^P, z^R) g^P(z^P|z^R) dz^P \\ &= Prob(z^P_H|z^R) \cdot \int_{\mathbf{Z}^P} \pi^R_{oi}(z^P, z^R) dG_H(z^P) + Prob(z^P_L|z^R) \cdot \int_{\mathbf{Z}^P} \pi^R_{oi}(z^P, z^R) dG_L(z^P)\end{aligned}$$

Note from Equation (B.9) that $\overline{\pi}_{oi}(z^P)$ is linear in a power function of z^P, and thus so is $\pi^R_{oi}(z^P, z^R) = (\gamma^{\frac{\gamma}{1-\gamma}} - \gamma^{\frac{1}{1-\gamma}})\left(\frac{1}{w^S_i}\right)^{\frac{\gamma}{1-\gamma}}\left(\overline{\pi}_{oi}(z^P) \cdot z^R\right)^{\frac{1}{1-\gamma}}$. With G^P_H and G^P_L being Pareto distributions, $\pi^R_{oi}(z^R)$ thus can be written as a function of z^R in an analytical form.