Talent, Geography, and Offshore R&D[*]

Jingting Fan[†]

Abstract

I model and quantify the impact of a new dimension of global integration: offshore R&D. In the model, firms match with heterogeneous researchers to develop new product blueprints, and then engage in offshore production and exporting. Cross-country differences in the distributions of firm managerial efficiency and researcher talent generate a "talent-acquisition" motive for offshore R&D, while the frictions impeding offshore production and trade lead to a "market-access" motive. I find empirical support for both motives using firm-level patenting data. I find additional evidence for these motives via counterfactuals using the calibrated model: international differences in endowment distributions and the market access motive collectively account for 90% of the average observed level of offshore R&D. Offshore R&D increases countries' gains from global integration by a factor of 1.2 on average, with much larger increases for developing than for developed countries. Incorporating offshore R&D also has important implications for understanding the welfare impact of traditional forms of global integration, namely trade and offshore production.

Keywords: Gains from openness, FDI, offshore R&D, offshore production, talent-acquisition, market access

JEL Classification: F21 F23 F40 O32

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

1 Introduction

Global integration in the form of international trade and multinational activities is one of the most significant economic phenomena of the past decades. Its impact has become an important topic for policy discussion and academic research. Existing studies on globalization focus on trade and multinationals' *offshore production* activities, but abstract from their *offshore R&D* activities, which also occur at significant levels. Figure 1 plots the share of R&D expenditures in a country incurred by the affiliates of foreign multinationals located in that country as a measure of offshore R&D. Uncolored bars are for 2012, and colored bars are for the first year with available data for each country, dating back to as early as 1985. By this measure, offshore R&D increased in most countries in the past two decades. In 2012, foreign affiliates accounted for more than 30% of R&D expenditures in the median country in the sample.[1]

Figure 1: The Level and Growth of Offshore R&D, 1985-2012
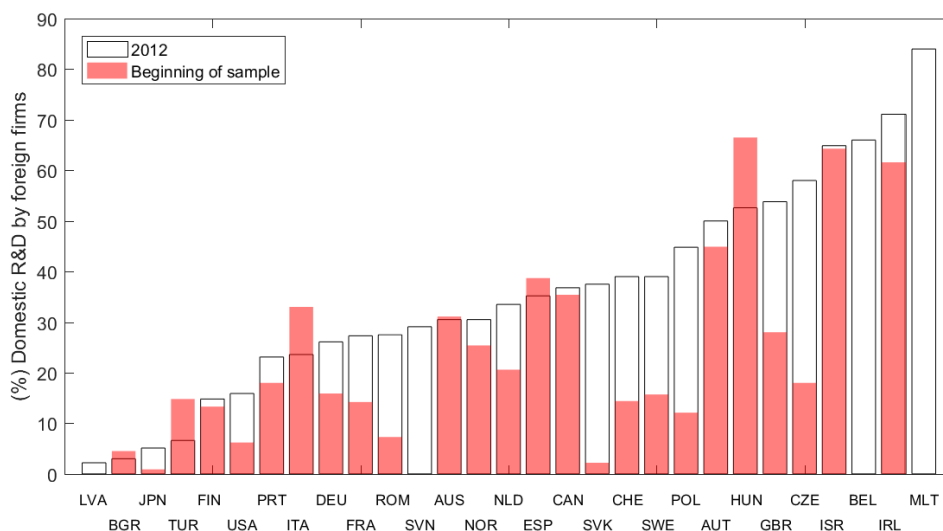
Notes: The measure for offshore R&D in country i is $\frac{\text{Total private R\&D expenditures in country } i \text{ by foreign firms}}{\text{Total private R\&D expenditures in country } i}$. Uncolored bars indicate the value of this variable in 2012; colored bars indicate the value at the beginning of the sample, which differs by country and dates back to as early as 1985. Data source: OECD.

The offshore R&D decisions of multinationals could have important aggregate implications. By determining the location and efficiency of R&D activities, offshore R&D directly affects the income of countries. Moreover, in a world interconnected through trade and offshore production, offshore R&D can affect income indirectly, by shaping countries' specialization in innovation or production.

In this paper I model and quantify the impact of offshore R&D. I address three questions. First, what are the determinants of offshore R&D? Second, how large are the welfare gains of opening up to offshore R&D? Third, how do these gains depend on and interact with the traditional forms

---

[1]In appendix I show that the importance of offshore R&D can also be established using international patent statistics.

of economic integration, namely trade and offshore production?[2]

I develop a unified framework for firms' global R&D and production decisions. In the model, firms differ along two dimensions: innovation efficiency, which governs how effective a firm is in converting researcher input into new product blueprints, and production efficiency, which governs a firm's productivity in converting production labor into output. Researchers differ in their talent. Firms can enter foreign countries (hosts) to perform offshore R&D. In each host, the firm matches with local researchers to develop new varieties. I model R&D as an assignment problem between firms and researchers, in which researcher talent and firm efficiency are complements. This setup deviates from the efficiency units assumption, and implies that quality and quantity of researchers are not perfect substitutes, an important feature of R&D in reality.[3]

I embed this offshore R&D decision into a multi-country general equilibrium model of global production and trade (Arkolakis et al., 2014). Specifically, after a product is developed by an R&D center, whether onshore or offshore, the firm first chooses which countries to sell it to, and then decides where to produce it. A firm from the U.S. therefore can develop a new product in the U.K., produce it in China, and export from there to India. These flexible decisions capture the complex strategies employed by modern multinationals.[4]

The model allows for two motives for offshore R&D commonly cited by firms: "market-access" and "talent-acquisition".[5] The former is straightforward: firms want to produce near their markets to save on trade costs. If separating innovation from production is costly, firms have incentives to offshore their R&D to large markets. The latter motive depends on both firm and host country characteristics. First, it reflects the host country's *relative* abundance of talented inventors, which depends on the abundance of talented inventors—an input supply effect, and the abundance of efficient firms competing for talent—an input demand effect. Second, because of the complementarity in innovation, host relative talent abundance interacts with firm efficiency to reinforce the talent-acquisition motive for high-efficiency firms.

Empirically, these two motives imply that market size and relative talent abundance increase offshore R&D into a host country, with the latter having a stronger effect for more efficient firms. Moreover, as most models with firm heterogeneity would predict, more efficient firms enter more countries for offshore R&D, and innovate more in each of them. I test these predictions using firm-level patenting data from the United States Patent and Trademark Office (USPTO). I define a patent as an output of offshore R&D in country A by a firm from country B, if its inventor is in country

---

[2]Throughout this paper, I use the term offshore production to refer to cases in which a product is produced in a location different from where it is developed. This is related to the term "multinational production" used in recent studies (Ramondo, 2014; Ramondo and Rodrıguez-Clare, 2013; Irarrazabal et al., 2013; Arkolakis et al., 2014; and Tintelnot, forthcoming).

[3]The output distribution of researchers is highly skewed. Akcigit et al. (2015) shows that the average top 1% inventor has 1019 lifetime citations, while the median inventor has only 11.

[4]DuPont offers a good example. Headquartered in Delaware, U.S., it has major R&D centers located in the U.S., Brazil, China, Switzerland, Korea, Germany, and Japan. Moreover, it has production facilities in 19 countries, from which it serves around 90 countries.

[5]According to firm-level surveys (see, for example, Thursby and Thursby, 2006), the quality of research personnel and host country market potential are the two most important factors firms consider, when choosing where to build their offshore R&D centers.

A and its owner is in country B. Empirical exercises based on this measure confirm the model predictions. In addition to these direct model implications, the assumption of complementarity between firm efficiency and researcher talent also implies that efficient firms match with talented inventors. I provide evidence for complementarity in the appendix.

I proceed to examine the quantitative implications of the model by calibrating it to 25 countries and a composite of 22 other countries. I parameterize each country's distribution of firm efficiency using the World Management Survey developed by Bloom et al. (2012), and its talent distribution using the international cognitive test score database developed by Hanushek and Woessmann (2012). I determine other parameters by matching various statistics of the firm size distribution in the U.S. and the intensities of bilateral international activities, including trade, offshore production, and offshore R&D. The model reasonably matches several non-targeted patterns in the data, while an otherwise similar model without complementarity between firm efficiency and researcher talent does not.

I quantify the importance of international differences in the distributions of firm efficiency and researcher talent in explaining the observed level of offshore R&D. I eliminate the incentives of offshore R&D arising from these distribution differences by first giving each individual country the management distribution of U.S. (the highest in the world), and then the talent distribution of Brazil (the lowest in the world). The former reduces the average level of offshore R&D by around three quarters, whereas the latter reduces this average by around one third. So differences in the distributions of talent and management efficiency are an important driving force for offshore R&D.

I further examine how a country's access to foreign markets through exporting, and to foreign producers through offshore production, affect its attractiveness as a destination for R&D. While both consumer and producer access increase the return to innovation in partial equilibrium, I find that they have opposite general equilibrium effects: consumer access reduces inward offshore R&D, while producer access increases it. Therefore, increasing access to foreign markets through reducing exporting costs would not necessarily help a country in attracting R&D-intensive FDI. Country specialization in innovation or production is the key to understanding this result. When a country loses access to foreign consumers through exporting, its competitiveness in production weakens, which lowers wages and makes it more attractive as a host for offshore R&D centers. As a result, it specializes more in innovation, and firms do R&D there and offshore their production to other countries. Such specialization is not possible without offshore production, so when both consumer and producer access are shut down, the average offshore R&D across countries decreases to less than half of the benchmark level.

Together, these two sets of experiments suggest that the talent-acquisition and market-access motives in the model are strong enough to account for the observed level of offshore R&D on average. I further examine the normative implications of offshore R&D. Under a special case, I derive an analytic expression for the model-implied gains from openness, which augments the expression in Arkolakis et al. (2014) with an additional term that captures the importance of foreign

companies in domestic R&D.[6] The expression makes it clear that offshore R&D represents a new channel for countries to benefit from global integration. I use the calibrated model to evaluate the quantitative relevance of this channel. The average welfare gains from offshore R&D, defined analogously to the gains from trade, are around 2.5% of real income. Compared to a restricted version of the model with only trade and offshore production, the welfare gains from openness in the full model with offshore R&D are larger by a factor of 1.2. Importantly, this amplification is substantially larger for emerging countries than for developed countries, mainly because a larger share of R&D in emerging countries is carried out by foreign affiliates. Overlooking this channel therefore will not only result in underestimating the gains from globalization, but also bias the assessment of the relative size of the welfare gains across countries.

Existing quantitative studies on multinational activities do not separately model offshore R&D and offshore production, even though they are very different activities that can be targeted by specific policies.[7] Is this an innocuous assumption for policy simulations? To answer this question, I compare the effects of policies designed to promote these two multinational activities, focusing on China and India as an example.[8] First, I reduce the inward offshore R&D costs in these two countries; second, I reduce inward offshore production costs. I find that, in the first experiment, China and India reap most of the benefits, whereas in the second experiment, developed countries also benefit significantly. The gains are small for developed countries in the first experiment because offshore R&D liberalization weakens the comparative advantage of China and India in production, which reduces the welfare gains from global specialization for everyone. In the second experiment, in contrast, the changes are more aligned with countries' comparative advantage. This comparison highlights the different implications for other countries of liberalization in offshore R&D and production. Such differences are especially relevant for studying multilateral investment agreements.

Offshore R&D also has implications for the welfare gains from other types of economic openness. To make this point, I perform an experiment with the same unilateral reductions in inward offshore production costs as in the previous experiment, but in a restricted version of the model without offshore R&D. Compared to the previous experiment, this experiment leads to substantially higher welfare gains for developed countries, and lower welfare gains for India and China. The distribution of profit from innovation is the key to the difference. More offshore production in China and India increases wages and reduces the profits from performing R&D there. With the equilibrium level of offshore R&D, the profit decreases are shared among domestic and foreign firms in these two countries; without offshore R&D, all the losses would be borne by domestic firms. This experiment shows that it is important to model offshore R&D, even if one's goal is to

---

[6]Gains from openness is defined as the change in real income as a country moves from complete isolation to the observed equilibrium.

[7]For example, countries can grant tax credits or open their borders specifically to R&D intensive FDI. An example is the U.K. "patent box," which reduced the corporate tax rate on revenues from R&D by 10 p.p.

[8]This policy evaluation is interesting in its own right because these two emerging giants are becoming popular destinations for offshore activities. Related to this trend, their governments are attempting to attract more foreign companies, especially R&D intensive ones, by cutting red tape and speeding up the approval process.

evaluate the effects of offshore production.

The rest of this paper is organized as follows. I review related literature in the next section. In Section 3 I describe the theoretical framework, and derive testable implications. In Section 4 I test these implications empirically. I then calibrate the model in Section 5, and perform counterfactual exercises in Section 6. I conclude and discuss future directions in Section 7.

2 Related Literature

This paper is related to the recent literature that quantifies the gains from globalization, especially studies on the aggregate implications of technological transfer through multinational activities (see, among others, McGrattan and Prescott, 2009; Burstein and Monge-Naranjo, 2009; Ramondo and Rodrıguez-Clare, 2013; Arkolakis et al., 2014; Irarrazabal et al., 2013; Tintelnot, forthcoming; Alviarez, 2016; and Holmes et al., 2015).[9] Within this literature, the most closely related paper is Arkolakis et al. (2014), which studies the welfare gains from trade and offshore production. The present paper differs in two aspects. First, rather than treating innovation efficiency of a country as a single exogenous parameter, I decompose it into two measurable components, firm innovation efficiency and researcher talent, and examine the role of each in shaping a country's comparative advantage in innovation. Second, I allow firms to perform offshore R&D by mobilizing their managerial capacity abroad, so a country's comparative advantage in innovation is endogenous. I show that this channel has quantitatively important implications for both the gains from openness, and the effect of specific policy changes.

This paper is also related to the literature explaining the pattern of FDI, dating at least as far back as the theoretical work by Helpman (1984) and Markusen (1984) (for horizontal and vertical FDI, respectively). More recently, researchers have examined the determinants of M&A FDI (Nocke and Yeaple, 2007; Nocke and Yeaple, 2008; and Head and Ries, 2008), and have incorporated firm heterogeneity into the model (Helpman et al., 2004).[10] This paper contributes to this literature in two ways. Theoretically, I outline a rich model of R&D, which can be viewed as a model of FDI with two-tiered vertical linkage: one between headquarters and R&D centers, and one between R&D centers and production sites. This structure allows the model to capture the complex strategies frequently seen in modern multinationals, in a way that existing two-country models of offshore R&D cannot (Gersbach and Schmutzler, 2011). Empirically, I test model predictions using patenting information at the USPTO for firms from multiple countries, and show that relative talent abundance and firm efficiency matter for offshore R&D in a way that is consistent with the model.[11]

---

[9]See Antràs and Yeaple (2014) for a recent review of the literature on multinational corporations. Also see Costinot and Rodriguez-Clare (2014) for a review of quantitative studies on the aggregate implications of international trade, which encompasses the bulk of the research on the gains from globalization.

[10]Studies have also examined empirically the impact on FDI flows of various factors, including skill endowments (Yeaple, 2003), institutions (Alfaro et al., 2008), and taxes and corruption (Wei, 2000).

[11]The empirical results complement the management literature on firms' incentives in doing offshore R&D, most of which are either based on firms' self-reported incentives or focus on firms in/from a single region. See, for example,

In terms of modeling, this paper is related to a number of studies that use an assignment framework to understand international trade and offshoring.[12] I apply matching framework to innovation decisions in a model of multinational production and trade, and quantify the effects of complementarity between firms and researchers. In doing so, I develop a computational algorithm that can solve the matching function efficiently in the presence of multiple countries and when endogenous offshore R&D decisions lead to discontinuities in innovation efficiency distributions. This setup and computational algorithm could have applications in other contexts.[13] In addition, while existing studies document positive assortative matching in general labor markets and the market for managers, this paper is, to my knowledge, the first to document positive assortative matching between inventors and firms.[14]

This paper's focus on international cooperation in R&D is shared by several recent papers (Kerr and Kerr, 2014; Kerr et al., 2016; and Branstetter et al., 2013). These papers discuss international cooperation either among inventors from different countries, or between inventors and firms from different countries, made possible by international migration or multinational activities. This paper contributes to this literature by developing a model of offshore R&D, testing its specific predictions, and quantifying the aggregate implications of offshore R&D.

3 A Model of Offshore R&D and Production

This section sets up the model and describes firms' global innovation and operation decisions.

3.1 Environment

There are N countries in the model, indexed by $i = 1, 2, ...N$. Country i is endowed with L_i^R measure of researchers, who differ in their talent, $\theta \in \Theta$, distributed according to $H_i(\theta)$, and L_i^P measure of homogenous production workers.[15] Researchers work with R&D centers to develop new differentiated varieties. Production workers manufacture these varieties and perform operational tasks for R&D centers (in the form of fixed costs). Country i is also endowed with E_i measure of heterogeneous firms with different innovation efficiencies, $\tilde{z}^R \in \tilde{\mathbb{Z}}^R$, distributed according to $G_i^E(\tilde{z}^R)$. Firms build R&D centers in different countries, which then recruit local researchers to develop new varieties. I use R_i to denote the measure of R&D centers in country i. In equilibrium R_i is an endogenous outcome determined by firms' offshore R&D decisions.

Ambos (2005), Shimizutani and Todo (2008), and Ito and Wakasugi (2007). See also Hall (2011) for a review.

[12]See, among others, Grossman and Maggi (2000), Yeaple (2005), Costinot and Vogel (2010), and Antràs et al. (2006).

[13]Roys and Seshadri (2014) quantifies a general equilibrium model of team production based on Antràs et al. (2006) in a closed-economy setting. Their model fixes the team size each of manager exogenously, so wages do not play an allocative role. In the present paper, wages determine team size and firm size distribution.

[14]Existing research mostly focuses on the match between workers in general and firms (see for example, Abowd et al., 1999 and the references thereto). More recently, research has focused on the match between firms/projects and CEOs (Terviö, 2008, among others).

[15]The talent distribution in a country reflects the quality of the education system, education choice, as well as cultural traits such as openness to innovation. By taking the talent distribution as given, this paper abstracts from the effect of international integration on these factors.

The representative consumer in country i decides how much to spend on each variety, according to the following preference:

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

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:

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

where $P_i^{1-\sigma} = \int_{\Omega_i} p_i(\omega)^{1-\sigma} d\omega$ is the ideal demand price index aggregated over $p_i(\omega)$, the price of variety ω in country i.

3.2 Firm Decisions: Overview

This subsection overviews firms' decisions. In the model, firms operate in multiple countries, and make sequential decisions on R&D, production, and exporting. I will use the following indexing conventions throughout this paper: o denotes a firm's headquarters, that is, the country where a firm originates; i denotes the country where a product is developed—the location of the R&D center; l denotes the country where the product is manufactured; and d denotes the destination country where it 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_i^R in country i production labor. An R&D center's innovation efficiency depends on that of its parent.[16] Motivated by evidence on spatial frictions in knowledge transfers within firms (see, for example, Irarrazabal et al., 2013; Keller and Yeaple, 2013), I assume that firms can only transfer part of their innovation management efficiency to offshore R&D centers. Letting $\phi_{oi}^R \leq 1$ be the proportion of innovation efficiency that can be transferred, the innovation management efficiency for an R&D center in country i operated by a country o firm is $z^R = \tilde{z}^R \phi_{oi}^R$. This efficiency governs how many varieties can be developed by a given number of researchers.

Innovative firms are not always the most efficient in carrying out manufacturing. To allow for this heterogeneity, 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. To capture positive correlation between innovation efficiency and production efficiency, the distribution from which z^P is drawn increases in z^R in the sense of first-order stochastic domi-

---

[16]This assumption follows a long tradition in the theory of multinationals, see, for example, Helpman, 1984; Helpman et al. (2004); and Nocke and Yeaple (2008). Empirically, Guadalupe et al. (2012) documents an increase in innovation and adoption of foreign technology upon acquisition by foreign companies.

Figure 2: Firm's Two-tiered Decisions

(a) Offshore R&D Decisions

Innovation Efficiency $\tilde{z}^R$

Home Country o

Fixed Cost c^R

Host Country i_1

Retained innovation efficiency: $\tilde{z}^R \boldsymbol{\Phi}^R_{o\,i_1}$

Draw production efficiency: z^P from $G^P(z^P | \tilde{z}^R \boldsymbol{\Phi}^R_{o\,i_1})$

Fixed Cost c^R

Host Country i_2

Retained innovation efficiency: $\tilde{z}^R \boldsymbol{\Phi}^R_{o\,i_2}$

Draw production efficiency: z^P from $G^P(z^P | \tilde{z}^R \boldsymbol{\Phi}^R_{o\,i_2})$

(b) Offshore Production and Export

Host Country i_1

Fixed Marketing cost c^M

Market d_1

Draw $(\boldsymbol{\eta}_1, \boldsymbol{\eta}_2, \boldsymbol{\eta}_3, \dots, \boldsymbol{\eta}_N)$, $\boldsymbol{\eta}_l \in F_l(x) = (e^{-T_l x^{-\delta}})$

Loss in product quality $1 - \boldsymbol{\Phi}^P_{i_1 l}$

Production site l

Shipping cost $\tau_{l d_1}$

Unit Production Cost: $\dfrac{w^P_l \tau_{l d_1}}{z^P \boldsymbol{\Phi}^P_{i_1 l} \boldsymbol{\eta}_l}$

9

nance. I use $G^P(z^P|z^R)$ to denote the CDF for production efficiency draws, with $g^P(z^P|z^R)$ being the corresponding probability density function (PDF).[17] This offshore R&D module is illustrated in Figure 2a. As the figure indicates, firms can open multiple R&D centers in different countries, but at most one R&D center in each country.

Given the production and innovation efficiency of affiliated R&D centers, (z^P, z^R), firms recruit researchers in each center to develop new differentiated varieties, and decide which countries to sell their products to. To sell products to destination country d, a per-variety fixed marketing cost of c_d^M in terms of country d production labor needs to be paid.

As Figure 2b indicates, firms can potentially manufacture products developed by their R&D centers in a third country l, where they do not necessarily perform R&D, and then export to destination countries. By separating production from R&D (offshore production), firms can take advantage of cheaper production labor and save on shipping fees. However, geographic separation makes it difficult for R&D centers to communicate with production plants, reducing production efficiency. I use $\phi_{il}^P \leq 1$ to denote the fraction of productivity that a firm can transfer from its R&D center in country i to production site in country l. For an R&D center with production efficiency z^P, the preserved *plant-level* offshore productivity in country l is $z^P \phi_{il}^P$. I further assume that there is a stochastic element, η_l, idiosyncratic to a production site and a variety, which enters productivity multiplicatively, so the *variety-level* productivity in l is $z^P \phi_{il}^P \eta_l$. The cost of producing and delivering one unit of product is $\frac{w_l^P \tau_{ld}}{z^P \phi_{il}^P \eta_l}$, which takes into account the cost of production labor, w_l^P, and shipping fee, τ_{ld}.

In the model, firms perform offshore R&D for several reasons. First, if a country is relatively abundant in talented inventors, foreign firms might want to enter to make full use of their skills. Anecdotes abound about MNCs establishing offshore R&D centers in order to tap into the local talent pool. Google, for instance, recently announced a plan to train two million Android developers in India within the next three years. According to a survey of 200 R&D executives (Thursby and Thursby, 2006), MNCs rank being close to highly qualified R&D personnel as the most important factor for the location choice of R&D centers in their home countries and other developed countries, and as the second most important factor, right after growth potential, for their new R&D centers in emerging economies.

The aforementioned production and trade decisions also imply that firms might choose to perform R&D in places close to major destination markets, or places with good access to countries with cheap production labor, in order to produce and distribute their products more efficiently.

By allowing for both offshore R&D and production decisions, the model captures the complexity of multinationals' global strategies. This stands in contrast to existing quantitative studies of multinationals that do not allow for offshore R&D. Such restriction might not be important if R&D activities performed by foreign affiliates are simply product adaption to local markets, a

---

[17]Under this assumption, the production management efficiency is specific to each R&D center. R&D centers with different innovation management efficiencies affiliated with the same parent will draw from different distributions. An alternative interpretation of this production management efficiency is the quality of products developed by an R&D center.

"by-product" of offshore production. Figure 3 demonstrates that, while U.S. multinationals' R&D expenditures in host countries increase strongly with host income, there is no such relationship for their overall employment. So product adaption unlikely to be the whole story, and offshore R&D is not simply a by-product of offshore production.

Figure 3: Overseas R&D and Employment by U.S. Multinationals

(a) U.S. Multinational Affiliate R&D (b) U.S. Multinational Affiliate Employment
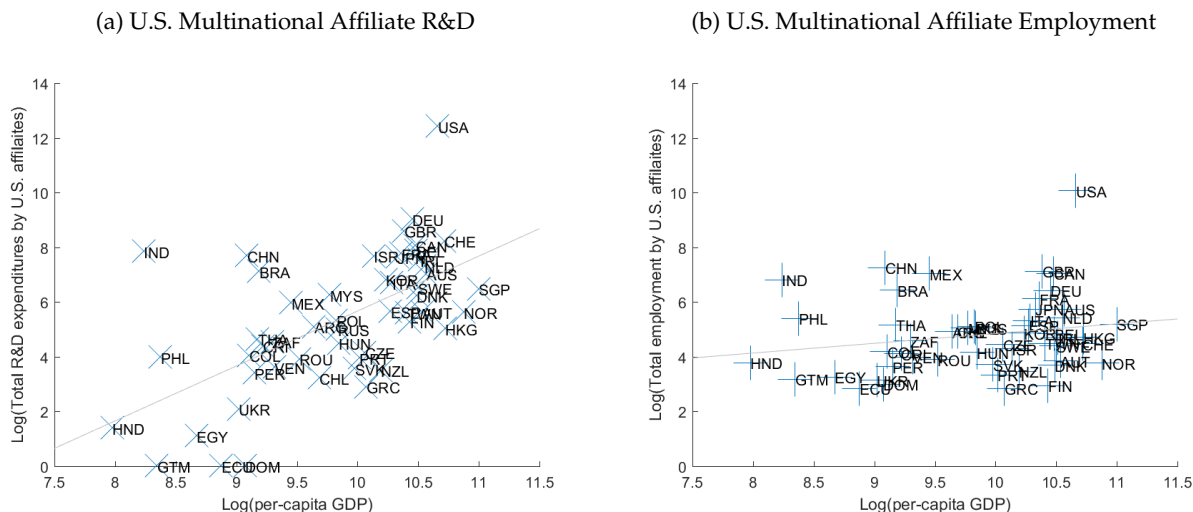
Notes: The left panel plots the log of total R&D expenditures by U.S. multinationals in each host country against host income. The right panel plots the total employment of U.S. multinationals against host income. Data source: Bureau of Economic Analysis.

Importantly, I assume that different varieties developed by a firm, either in the same or in different R&D centers, are differentiated from each other and from varieties developed by all other firms. Such an assumption is consistent with how R&D is organized in many multinational firms. General Electric, for example, organizes its ten research labs by scientific disciplines in five countries (the U.S., Germany, India, China, and Brazil).[18] [19] This assumption implies that firms make offshore R&D decisions for each country independently and that R&D centers affiliated with the same firm operate as if they are independent from each other.

Given this independence, in the remainder of this section, I first consider the production and trade decision of a firm, after a variety has been developed. I then describe the innovation decision of each R&D center, and firms' decisions to build offshore R&D centers. Finally, I characterize the market for researchers and analyze the welfare gains from openness under a special case.

---

[18] Alternatively, this assumption can be interpreted as capturing M&A FDI. More than 70% of FDI flows in the data are in the form of mergers and acquisitions (Nocke and Yeaple, 2008). One explanation for this observation is that, by transferring know-how and managerial capacity to targets, acquiring firms can improve the operating efficiency of the targets. The differentiated-variety assumption adopted in the present paper is consistent with this perspective of FDI—multinationals transfer their managerial technology to newly acquired foreign R&D centers, and increase the efficiency of these R&D centers in carrying out their independent product development.

[19] This assumption treats R&D at headquarters and R&D in offshore centers symmetrically. Recently, Bilir and Morales (2016) estimates the effects of R&D on productivity for multinational firms. They find that R&D at headquarters have stronger spillover effects to foreign affiliates than R&D at affiliates to other affiliates. The current model cannot account for this finding. But an extension of the 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 finding.

3.3 Production and Trade

Consider a variety developed by an R&D center (z^P, z^R) in country i, which can potentially be produced in any country by production labor using a linear production technology. For each variety, an R&D center obtains a vector of N idiosyncratic productivity draws, one for each potential production site, denoted $\boldsymbol{\eta} = (\eta_1, \eta_2, .., \eta_N)$. I assume that η_l is independent across countries, and follows a Frechet distribution: $F(x) \equiv \text{Prob}(\eta \leq x) = \exp(-\Lambda_l x^{-\delta})$, where Λ_l governs the mean of the draws for country l, and δ governs the dispersion of the draws across varieties and countries. The productivity for a variety in country l is: $z^P \phi_{il} \eta_l$.

Letting w_l^p denote the wage rate for each unit of production labor in country l, the cost of serving country d by producing in country l is $c_{ild} = \frac{w_l^p \tau_{ld}}{z^P \phi_{il}^p \eta_l}$, where τ_{ld} is the iceberg shipping cost from l to d. Given the monopolistic competition market structure, the price for a variety sold in country d, *if* produced in country l, is

$$p_{ild} = \frac{\sigma}{\sigma - 1} \frac{w_l^p \tau_{ld}}{z^P \phi_{il}^p \eta_l}.$$

Conditional on serving destination market d, a firm chooses the lowest cost production location for each of its varieties. Because there are no fixed costs in offshore production, all countries are potentially production sites. The price of this variety in country d is simply the lowest one among all possible choices:

$$p_{id}(\boldsymbol{\eta}) = \min_l \{ \frac{\sigma}{\sigma - 1} \frac{w_l^p \tau_{ld}}{z^P \phi_{il}^p \eta_l} \}.$$

For each variety and each destination market, production will take place in one country. However, since each R&D center develops a continuum of varieties, in equilibrium, a firm will serve each destination through all countries in the world.[20] For tractability, I assume that each R&D center needs to decide first which destination markets to enter and pays the fixed marketing cost before knowing the idiosyncratic country-specific productivity draws, so firms make destination market entry decisions based on expected profits. The expected per-variety profit from market d for the R&D center from country i, defined as $\pi_i^d(z^P)$, is

$$\pi_i^d(z^P) = \frac{1}{\sigma} (\frac{\sigma}{\sigma - 1})^{1-\sigma} \Gamma(\frac{\delta + 1 - \sigma}{\delta}) P_d^{\sigma-1} X_d (\frac{1}{z^P})^{1-\sigma} \Psi_{id}^{\frac{\sigma-1}{\delta}} - c_d^M w_d^p,$$

where Γ is the Gamma function, and $\Psi_{id} = \sum_l \Lambda_l (\frac{w_l^p \tau_{ln}}{\phi_{il}^p})^{-\delta}$. The first term in this expression is calculated from $\frac{1}{\sigma} P_d^{\sigma-1} X_d \int \min_l(p_{ild}(\boldsymbol{\eta})^{1-\sigma}) d\tilde{F}(\boldsymbol{\eta})$, with $\tilde{F}(\boldsymbol{\eta})$ being the distribution of $\boldsymbol{\eta} = (\eta_1, \eta_2, ..., \eta_N)$.

This expected profit increases in the production efficiency of an R&D center, z^P, so there exists a threshold $\hat{z}_{id}^p$ such that R&D centers from i will expend marketing costs and enter country d if

---

[20]This result implies that the model cannot capture the extensive-margin of firms' offshore *production* decisions. This is not necessarily an important drawback, as the focus of this paper is on offshore R&D and its interaction with offshore production in the aggregate. In the next section I show the model predictions on firms' offshore R&D decisions are supported empirically.

and only if their production efficiency is above this threshold. This cutoff is given by:

$$\pi_i^d(\hat{z}_{id}^P) = 0. \tag{1}$$

A firm makes an independent entry decision for each destination market. The per-variety expected profit for a firm with production efficiency draw z^P, taking into account its potential entry into all destination markets, is

$$\pi_i(z^P) = \sum_d \mathbb{I}_{z^P \geq \hat{z}_{id}^p} \pi_i^d(z^P). \tag{2}$$

3.4 Innovation and the Market for Researchers

R&D centers choose the talent of researchers, θ, and their quantity, $l(\theta)$, to develop new differentiated varieties. Let y be the measure of differentiated varieties developed:

$$y = f(z^R, \theta)l(\theta)^\gamma,$$

where γ measures the return to the number of researchers, and $f(z^R, \theta)$ captures how firm innovation efficiency and researcher talent affect innovation output. I assume that $\gamma < 1$, implying decreasing returns to scale in the number of researchers. This assumption has several interpretations. First, it can be thought of as a reduced-form approximation to a model in which R&D requires supervision from the top management, but managerial time is limited in a company. In such a context, hiring more researchers results in less supervision time for each of them, reducing researcher productivity.[21] An alternative is to think of innovation output as a function of both accumulated knowledge capital and researcher input. In a static model in which the distribution of knowhow and accumulated knowledge is given, the research output features decreasing returns to researcher input. Finally, decreasing returns to scale might stem from increases in coordination costs, free-riding, and disagreement among researchers as teams expand.[22]

Given $\pi_i(z^P)$, the per-variety expected profit, the optimization problem for the R&D center is

$$\pi_i^R(z^P, z^R) = \max_{\theta \in \Theta, l(\theta)} [\pi_i(z^P)f(z^R, \theta)l(\theta)^\gamma - w_i(\theta)l(\theta)],$$

where $w_i(\theta)$ is the wage for a researcher with talent θ. As is clear from the equation, the production efficiency of a firm affects innovation incentives because it determines the profit for each variety. I make the following assumption about f:

Assumption 1 *f is twice continuously differentiable and increasing in its arguments, i.e., $f_1, f_2 > 0$. Further, f is log-supermodular, i.e., $\frac{\partial^2 \log f(z^R, \theta)}{\partial z^R \partial \theta} > 0$.*

---

[21]See Antràs et al. (2006) for an analysis of the effects of offshoring in a model in which managers can only supervise a fixed number of workers.

[22]Such coordination costs have been documented empirically. For example, Haas and Choudhury (2015) finds that, while total patenting increases with the number of members in a team, the increase is smaller than the increase in the team size—there is decreasing returns in the number of researchers in a team.

The assumption that $f_1, f_2 > 0$ simply means that more efficient firms and more able researchers are more productive in innovation. The log-supermodularity assumption implies strong complementarity between researcher ability and firm efficiency. Under this assumption, more productive firms have a comparative advantage in working with more able researchers.[23] R&D activities require cooperation between researchers, and a large amount of managerial and monetary resources. Moreover, after a product prototype is developed, testing and marketing costs are big hurdles to clear before the product can reach consumers. A well-managed firm can do all of these tasks better, so it is especially profitable for them to work with talented researchers. The model captures this idea with the log-supermodularity of f.

The setup here deviates from the efficiency units assumption. A researcher with high talent is more valuable than multiple researchers with lower talent. Similarly, a firm with high innovation efficiency is more productive in R&D than multiple firms with lower efficiencies. These implications are in line with a few observations in the literature. First, as mentioned earlier, the *quality* of research talent is one of the top considerations when firms choose where to build their offshore R&D centers, along with the *cost* of research labor.[24] Second, it is well documented that there are a large number of small and less productive firms in developing countries, the prevalence of which can account for an important fraction of cross-country income differences (Hsieh and Klenow, 2009). Management efficiency might be a source of performance differences between firms (Bloom et al., 2013). To the extent that many developing countries have a large number of very small firms, they might not necessarily the lack a sufficient *stock* of management efficiency. The model here is consistent with view that it is not necessarily lack of management efficiency stock, but rather the lack of exceptional firms like Apple and Google, that explains the low incomes in developing countries.[25] Finally, the complementarity also implies that the same inventor will be paid more to work in a more efficient firm. This is consistent with the finding that larger and more productive firms pay a wage premium (see, for example, Schank et al., 2007), and the evidence on positive assortative matching between firms and inventors I provide in the appendix.

I now characterize the market for researchers. Let $T_i(z^P, z^R) : (\mathbb{Z}_i^P, \mathbb{Z}_i^R) \to \Theta$ be the optimal choice of θ for an R&D center characterized by (z^P, z^R). We have the following lemma:

Lemma 1 *T_i is continuous and strictly increasing in z^R. Moreover, T_i is independent of z^P.*

Proof   See appendix.

The proof of Lemma 1 is an extension of assortative matching results in the literature (see, for example, Grossman and Helpman, 2014; Grossman et al., 2015; Sampson, 2014) to the case with an additional source of heterogeneity, namely the production efficiency. Because high z^R R&D

---

[23]The log-supermodularity assumption has been adopted in a growing literature in international trade which uses assignment models to study questions such as the determinants of specialization and the impacts of trade integration, as reviewed recently by Costinot et al. (2015). The framework here is similar to the one in Grossman et al. (2015).

[24]Branstetter et al. (2013) conducts interviews with foreign-affiliated R&D centers in China. The interview responses stress the scale and quality of the research talent in China, rather than its cost.

[25]Roys and Seshadri (2014) builds a model of matching between heterogeneous entrepreneurs and workers, enriched with human capital accumulation, to show that the model can account for the differences in life cycle dynamics between firms in rich and poor countries, and can explain a substantial share of income differences between countries.

centers enjoy a higher marginal productivity increase from hiring better researchers, they have a comparative advantage in working with high-ability researchers, leading to assortative matching. Since z^P enters firms' innovation output multiplicatively in the form of $\pi_i(z^P)$, higher z^P does not affect the type of researchers hired by an R&D center, but only their quantity. In the following I will write the matching function simply as $T_i(z^R)$, omitting the argument z^P.

Given the equilibrium $w_i(\theta)$, the demand of an R&D center for researchers, if it chooses researchers with talent θ, is

$$l_i(z^P, z^R) = \left(\frac{\gamma \pi_i(z^P) f(z^R, \theta)}{w_i(\theta)}\right)^{\frac{1}{1-\gamma}}.$$ (3)

The corresponding measures of invention and profit are therefore:

$$y_i(z^P, z^R) = \left(\gamma \frac{\pi_i(z^P)}{w_i(\theta)}\right)^{\frac{\gamma}{1-\gamma}} f(z^R, \theta)^{\frac{1}{1-\gamma}},$$ (4)

$$\pi_i^R(z^P, z^R) = \left(\gamma^{\frac{\gamma}{1-\gamma}} - \gamma^{\frac{1}{1-\gamma}}\right) w_i(\theta)^{-\frac{\gamma}{1-\gamma}} [\pi_i(z^P) f(z^R, \theta)]^{\frac{1}{1-\gamma}}.$$ (5)

In equilibrium, firms choose the type of researchers to maximize profit. This requires the improvement in marginal output from higher-quality researchers to be exactly offset by their higher wages. We can obtain this equation by differentiating Equation (5) with respect to θ:

Lemma 2 $w_i(\theta)$ *satisfies the following relationship:*

$$\frac{w_i'(\theta)}{w_i(\theta)} = \frac{f_2(z^R, \theta)}{\gamma f(z^R, \theta)}\Big|_{\theta = T_i(z^R)}.$$ (6)

Proof See appendix.

The formal proof of Lemma 2 establishes the differentiability of $w_i(\theta)$. The proof is similar to that in Sampson (2014) and is delegated to the appendix.

Since researchers are heterogeneous, labor market clearing requires that the total demand equals total supply for each type. Let $\underline{\theta}_i$ and $\bar{\theta}_i$ be the lower and upper limits of the support for the researcher talent distribution, and let $\underline{z}_i^R$ and $\bar{z}_i^R$ denote the lower and upper limit of the support for the innovation efficiency distribution, respectively. To derive the researcher market clearing conditions for each type, I start with an aggregate version: for all $\underline{\theta}_i < \theta < \bar{\theta}_i$, the number of researchers with talent lower than θ is equal to the total demand for researchers with talent below θ. Formally,

$$L_i^R \int_{\underline{\theta}_i}^{T_i(z^R)} dH_i(\theta) = R_i \int_{\underline{z}_i^R}^{z^R} \left[\int_{\mathbf{Z}^P} l_i(z^P, z) g_i^P(z^P|z) dz^P\right] g_i^R(z) dz$$

$$= R_i \int_{\underline{z}_i^R}^{z^R} \left(\frac{\gamma f(z, T_i(z))}{w_i(T_i(z))}\right)^{\frac{1}{1-\gamma}} \left[\int_{\mathbf{Z}^p} \pi_i(z^p)^{\frac{1}{1-\gamma}} g_i^P(z^P|z) dz^P\right] g_i^R(z) dz,$$

where R_i is the measure of R&D centers in country i and $g_i^R(z)$ is their PDF, both of which are determined in equilibrium by firms' offshore R&D decisions. On the left of this equation is the

total number of researchers with talent below $T_i(z^R)$, and on the right side is the corresponding total demand.

Differentiating this equation with respect to z^R, we have the following equation:[26]

$$L_i^R T'(z^R) h_i(T_i(z^R)) = R_i \left(\frac{\gamma f(z^R, T_i(z^R))}{w_i(T_i(z^R))} \right)^{\frac{1}{1-\gamma}} \int_{\mathbf{Z}^P} g_i^R(z^R) \pi_i(z^P)^{\frac{1}{1-\gamma}} g^P(z^P | z^R) dz^P \tag{7}$$

Equation (7) then characterizes the market clearing condition for each researcher type. Equations 6 and 7, together with two boundary values,

$$T_i(\underline{z}_i^R) = \underline{\theta}_i, \quad T_i(\bar{z}_i^R) = \bar{\theta}_i, \tag{8}$$

determine the matching function $T_i(z^R)$ and the wage schedule $w_i(\theta)$. In summary, we have the following results:

Proposition 1 *Under Assumption 1, 1) Firms with higher innovation efficiency hire strictly better researchers. Firms with the same innovation efficiency but different production efficiencies hire the same type of researchers in different quantities. 2) The researcher labor market is characterized by Equations 6, 7, and 8.*

How does the output of R&D centers with different innovation efficiencies depend on the talent distribution of a country? Since a change in the talent distribution affects the entire matching function, characterizing the effect of a general change is difficult. I consider changes to the talent distribution that can be ranked by the following criterion:

Definition 1 *Consider $h(\theta)$ and $\tilde{h}(\theta)$, probability density functions for the talent distribution. $\tilde{h}(\theta)$ is more talent abundant relative to $h(\theta)$, if $\tilde{h}(\theta_2)h(\theta_1) \geq \tilde{h}(\theta_1)h(\theta_2), \forall \theta_1 < \theta_2$.*

This notion of factor abundance, which is stronger than first-order stochastic dominance, is introduced by Costinot and Vogel (2010) to characterize how relative factor supply and factor demand determine allocation and prices.[27] Under this notion, we have the following proposition:

Proposition 2 *Consider two R&D centers in country i, with innovation efficiencies $z_2^R > z_1^R$ and a common production efficiency z^P. Then $[log(y(z^P, z_2^R)) - log(y(z^P, z_1^R))]$, the log difference of R&D output between these two R&D centers, increases when the inventor distribution in country i becomes more talent abundant according to Definition 1, if one of the following are satisfied: 1) z^P and z^R are independent; 2) $c_d^M = 0$ for all d.*

Proof See appendix.

---

[26]Because of offshore R&D decisions, g_i^R is not necessarily continuous. At the finite discontinuous points of g_i^R, the matching function might not be differentiable. In this case, Equation 7 is not defined on the discontinuous points of g_i^R. While T_i is still well defined and continuous, the kinks in T_i' make it challenging to solve the matching function numerically. In the quantitative section, I describe a computational algorithm suited for this context.

[27]According to this definition, when θ_1 and θ_2 are in the support of both distributions, $\frac{\tilde{h}(\theta_2)}{\tilde{h}(\theta_1)} > \frac{h(\theta_2)}{h(\theta_1)}$, meaning h' has a higher relative share of the higher-skill type.

The intuition for this result is that, under the additional conditions stated above, increases in talent abundance improve the quality of researchers for all firms. This benefits efficient firms disproportionately more, because of the complementarity between talent and efficiency.[28] Since this proposition works through improving of match quality for firms, it also applies to a change in the firm innovation efficiency distribution that results in improvements in match quality for all firms between (z_1^R, z_2^R). An example of such a change is a decrease in the "efficiency abundance" of the firm distribution in the spirit of Definition 1. The talent abundance in the proposition should thus be broadly interpreted as a *relative* measure—the talent abundance of the inventor distribution, relative to the efficiency abundance of the firm distribution.

Although Proposition 2 is stated in the context of domestic firms, it applies to all active R&D centers in a host country. We can test the model by comparing the innovation output of R&D centers affiliated with companies with different innovation efficiencies. If Proposition 2 is correct, then this difference will be larger in host countries with higher relative talent abundance. In the next section, I test this implication directly. Later I show that the complementarity channel underlying this prediction is quantitatively relevant in determining the pattern of offshore R&D across host countries with different talent distributions.

3.5 Offshore R&D

Now we can characterize firms' decisions to open offshore R&D centers. I make the following assumption 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.*

Define $\pi_i^R(z^R)$ as the expected profit (over the possible z^P draws) for an R&D center in country i, with innovation efficiency z^R:

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

Firms compare the expected profit from building an offshore R&D center to the fixed cost of setting up the center, $c_i^R w_i^P$. By definition (Equations 5 and 2), $\pi_i^R(z^P, z^R)$ increases in z^P. We can also show that $\pi_i^R(z^P, z^R)$ increases with z^R.[29] Assumption 2 then implies that $\pi_i^R(z^R)$ increases strictly in z^R, so the decision to offshore R&D follows a threshold rule: there exists a cutoff $\hat{z}_{oi}^R$, so that firms from country o will perform offshore R&D in country i if and only if its innovation efficiency is above $\hat{z}_{oi}^R$. This cutoff is given by the following zero profit condition:

$$\pi_i^R(\hat{z}_{oi}^R \phi_{oi}^R) = c_i^R w_i^P. \tag{9}$$

---

[28] Grossman et al. (2015), Sampson (2014), and Costinot and Vogel (2010) obtain similar results on the effects of trade on income inequality under the log-supermodularity assumption. Compared to these papers, additional technical assumption is needed to ensure that equilibrium changes in return to R&D, $\pi_i(z^P)$, due to the distribution change do not decrease the quality of match for any firms.

[29] From Equation 5, $\frac{\partial \log(\pi_i^R(z^P, z^R))}{\partial z^R}\big|_{\theta = T_i(z^R)} = \frac{1}{1-\gamma} \frac{\partial \log(f(z^R, \theta))}{\partial z^R} > 0.$

3.6 R&D Center Efficiency Distribution

Firms' offshore R&D decisions determine g_i^R, the distribution of innovation management efficiency, and hence the distribution of production management efficiency, in each country. Given $\hat{z}_{oi}^R$, we can now derive R&D centers' production and innovation efficiency distributions. Let $G_i^R(z^R)$ be the CDF for innovation management efficiency of the R&D centers active in country i, and let $G_o^E(\tilde{z}^R)$ be the CDF of the distribution of innovation efficiency for firms from country o. Then we have the following equation:

$$R_i G_i^R(z^R) = \sum_{o=1}^{N} \mathbb{I}_{\frac{z^R}{\phi_{oi}^R} > \hat{z}_{oi}^R} E_o G_o^E \left(\frac{z^R}{\phi_{oi}^R}\right)$$

Differentiating this equation with respect to z^R, we obtain the density function:

$$g_i^R(z^R) = \frac{1}{R_i} \sum_{o=1}^{N} \mathbb{I}_{\frac{z^R}{\phi_{oi}^R} > \hat{z}_{oi}^R} E_o g_o^E \left(\frac{z^R}{\phi_{oi}^R}\right) \frac{1}{\phi_{oi}^R}. \tag{10}$$

The PDF for R&D centers with (z^P, z^R) is $g_i(z^P, z^R) = g^P(z^P|z^R) g_i^R(z^R)$.

3.7 Aggregation

Knowing $g_i(z^P, z^R)$, I derive the total measure of varieties that are invented in a country, denoted M_i, and the distribution of these varieties over different production efficiencies. Letting $m_i(z^P)$ be the measure of varieties innovated in country i by R&D centers with a production efficiency of z^P, then we have:

$$m_i(z^P) = R_i \int_{\mathbb{Z}^R} y_i(z^P, z^R) g_i(z^P, z^R) dz^R$$

$$M_i = \int_{\mathbb{Z}^P} m_i(z^P) dz^P, \tag{11}$$

where $y_i(z^P, z^R)$ is given by Equation 4. The price index in country d is then given by the following equation:

$$P_d^{1-\sigma} = \sum_i \int_{z^P > \hat{z}_{id}^p} m_i(z^P) \left[\int \min_l \{p_{ild}(\eta)^{1-\sigma}\} d\tilde{F}(\eta)\right] dz^P$$

$$= \Gamma\left(\frac{\delta + 1 - \sigma}{\delta}\right) \left(\frac{\sigma}{\sigma - 1}\right)^{1-\sigma} \sum_i \Psi_{id}^{\frac{\sigma-1}{\delta}} \int_{z^P > \hat{z}_{id}^p} m_i(z^P) z^{P\sigma-1} dz^P \tag{12}$$

To express the aggregate objects in the model, let X_{id} be the total sales in country d of the products developed in country i. We have the following:

$$X_{id} = P_d^{\sigma-1} X_d \int_{\hat{z}_{id}^P}^{\infty} m_i(z^P) [\int \min_l \{p_{ild}(\boldsymbol{\eta})^{1-\sigma}\} d\tilde{F}(\boldsymbol{\eta})] dz^P$$
$$= (\frac{\sigma}{\sigma-1})^{1-\sigma} \Gamma(\frac{\delta+1-\sigma}{\delta}) P_d^{\sigma-1} X_d \Psi_{id}^{\frac{\sigma-1}{\delta}} \int_{\hat{z}_{id}^P}^{\infty} m_i(z^P)(z^P)^{\sigma-1} dz^P. \tag{13}$$

These sales can be fulfilled through production in any country. Letting X_{ild} denote the value of production in country l, then we have $\sum_l X_{ild} = X_{id}$. I further define Y_l to be the total production of the varieties in country l, so $\sum_{i,d} X_{ild} = Y_l$. The Frechet assumption on idiosyncratic productivity draws also implies that, for each R&D center located in country i, the share of products it sells in country d that are fulfilled through production in country l is:

$$\psi_{ild} = \frac{\Lambda_l(\frac{w_l^P \tau_{ld}}{\phi_{il}^P})^{-\delta}}{\Psi_{id}},$$

with $\Psi_{id} = \sum_l \Lambda_l(\frac{w_l^P \tau_{ld}}{\phi_{il}^P})^{-\delta}$. Because this probability is the same for all R&D centers from country i, it also applies to the aggregate sales:

$$X_{ild} = \psi_{ild} X_{id}. \tag{14}$$

Production workers are used to produce output, and to pay fixed R&D and marketing costs. The production labor market clearing condition is:

$$w_d^P L_d^P = \underbrace{\frac{\sigma-1}{\sigma} Y_d}_{\text{Production}} + \underbrace{\sum_o E_o c_d^R w_d^P (1 - G_o^R(\hat{z}_{od}^R))}_{\text{Fixed R\&D center setup costs}} + \underbrace{c_d^M w_d^P \sum_i \int_{\hat{z}_{id}^P}^{\infty} m_i(z^P) dz^P}_{\text{Fixed marketing costs}}. \tag{15}$$

Recall that the density of firms from country o with innovation efficiency $\tilde{z}^R$ is $g_o^E(\tilde{z}^R)$. We can integrate π_i^R over $g_o^E(\tilde{z}^R)$ to compute the total profits made by country i R&D centers affiliated to firms from country o, denoted Π_{oi}:

$$\Pi_{oi} = E_o \int_{\hat{z}_{oi}^R}^{\bar{z}_i^R} \pi_i^R(\tilde{z}^R \phi_{oi}^R) g_o^E(\tilde{z}^R) d\tilde{z}^R,$$

This profit is after deducting R&D, marketing, and production costs, but before deducting fixed costs for building R&D centers.

Let I_i be the total R&D expenditures in country i, defined as total compensation to researchers in country i. Let I_{oi} be the expenditures in I_i that are incurred by affiliates of firms from country o. Equations 3 and 4 imply that:

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

The income of country d comes from three sources: wages of production labor, compensation

to researchers, and the net profit made by domestic firms from the country. Current account balance requires that total consumption of each country equals total income:

$$X_d = \underbrace{w_d^P L_d^P}_{\text{Production Labor}} + \sum_i \underbrace{[\Pi_{di} - E_d c_i^R w_i^P (1 - G_d^R(\hat{z}_{di}^R))]}_{\text{Net Profit}} + \underbrace{I_d}_{\text{Researcher Compensation}}. \tag{16}$$

Definition 2 *The competitive equilibrium is defined as a set of allocations and prices, such that:*

1. *Firms' market entry decisions satisfy Equation 1.*

2. *The matching function, T_i, and wage schedule for researchers, w_i satisfy Equations 6, 7, and 8.*

3. *Firms' offshore R&D decisions satisfy Equation 9.*

4. *The distribution of R&D center innovation efficiency in each country satisfies Equation 10.*

5. *The distribution of productivity efficiency for varieties satisfies Equation 11.*

6. *The price index in each country satisfies Equation 12.*

7. *The wage for production labor satisfies Equation 15.*

8. *The total expenditure in each country satisfies Equation 16.*

3.8 The Gains from Openness

In this subsection I focus on a special case to derive an expression for the welfare gains from openness, defined as the percentage change in real income ($\frac{X_d}{P_d}$), as a country moves from complete isolation to the degree of openness observed in the data. This expression makes it clear that offshore R&D is a new channel for countries to benefit from globalization. It also relates the size of this benefit to observable information and model parameters. Specifically, I make the following assumption:

Assumption 3 *1) $f(z^R, \theta) = z^R \theta^\beta$;*
2) Production efficiency, z^P, is independent of z^R, and follows a Pareto distribution: $G_d^P(x) = 1 - (\frac{x}{z_d^P})^{-\kappa_P}$;
3) There is no fixed marketing cost: $\forall d, c_d^M = 0$;
4) Firm innovation efficiency, $\tilde{z}^R$, follows a Pareto distribution: $G_d^E(x) = 1 - (\frac{x}{\tilde{z}_d^R})^{-\kappa_R}$.

The first part of the assumption maintains that $f(z^R, \theta)$ takes a multiplicative form.[30] Under this assumption, $\frac{\partial^2 \log f(z^R, \theta)}{\partial z \partial \theta} = 0$, so $f(z^R, \theta)$ no longer satisfies the strict log-supermodularity requirement in Assumption 1. Since a CES function with elasticity of substitution smaller than 1 satisfies strict log-supermodularity, the multiplicative case represents the limiting case as the elasticity approaches 1. This simplification will allow us to solve for the equilibrium wage schedule and firm-level decisions analytically.

---

[30]The assumption that the power of z^R is 1 is without loss of generality, because the units of z^R can always be scaled so that it enters $f(z^R, \theta)$ with a power of 1.

In the general model, because firms endogenously choose how many varieties to develop, aggregation is difficult. The first three components of Assumption 3, however, imply the Pareto distribution of production efficiency for varieties, which admits analytical aggregation. The fourth component in turn allows us to derive the total fixed costs of R&D in each country. With these simplifications, we have the following:

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

$$GO_d = \Big(\frac{X_{ddd}}{\sum_l X_{dld}}\Big)^{-\frac{1}{\delta}} \Big(\frac{\sum_l X_{dld}}{X_d}\Big)^{-\frac{1}{\sigma-1}} \Big(\frac{I_{dd}}{I_d}\Big)^{-\frac{1-\gamma}{\sigma-1}} \Big(\frac{\frac{\sigma-1}{\sigma}}{\frac{\sigma-1}{\sigma}\frac{Y_d}{X_d} + \frac{(1-\gamma)\kappa_R-1}{\gamma\kappa_R}\frac{I_d}{X_d}\big(1-\frac{I_{dd}}{I_d}\big)}\Big) - 1. \tag{17}$$

Proof See appendix.

This expression highlights various forces through which a country benefits from economic integration. The first term, $\frac{X_{ddd}}{\sum_l X_{dld}}$, captures the benefits from offshore production for consumption. The second term, $\frac{\sum_l X_{dld}}{X_d}$, captures the benefits from foreign innovation for consumption. These two terms are direct effects of offshore production and trade in the model. The third term, $\frac{I_{dd}}{I_d}$, captures the importance of foreign firms in domestic R&D. Intuitively, the smaller is this ratio, the more a country relies on foreign affiliates for R&D, and the more significant are the welfare gains from offshore R&D. The last term in the equation captures the effects of profit flows on welfare through their impacts on total expenditures. This indirect effect tends to bring positive welfare impacts, for countries that specialize in R&D (smaller $\frac{Y_d}{X_d}$), and countries that rely more on domestic firms in R&D (smaller $\frac{I_d - I_{dd}}{X_d}$).[31]

In the appendix, I compare this formula to that from Arkolakis et al. (2014). The first two and the fourth terms in Equation 17 also appear in their formula, with minor adjustments to reflect the modeling differences. The third term, $\frac{I_{dd}}{I_d}$, does not show up in their formula. To have an idea of how large this term is, consider the median country in the quantitative section, with about 30% of its R&D done by foreign affiliates. The value of $\big(\frac{I_{dd}}{I_d}\big)^{-\frac{1-\gamma}{\sigma-1}}$ is around 1.055, when $\gamma = 0.4$ and $\sigma = 5$. All else equal, this term generates a 5% real income change. So offshore R&D indeed represents a quantitatively important channel through which countries benefit from global integration.

4 Empirical Evidence

The model generates several predictions that relate offshore innovation to firm innovation efficiency and host country characteristics. Specifically, firm heterogeneity implies that more efficient firms will offshore their R&D to a larger number of host countries, and perform more R&D in each of them. At host country level, the market-access motive implies that larger countries are more attractive as a host for offshore R&D centers. The talent-acquisition motive has two implications. First, host countries with higher relative inventor talent abundance attract more offshore R&D.

---

[31]Finite aggregate fixed R&D costs require $(1-\gamma)\kappa_R - 1 > 0$.

Second, as Proposition 2 indicates, this effect is especially strong for more efficient home country firms, because of the complementarity between researcher talent and firm innovation efficiency in R&D. I will test these predictions in this section.

The complementarity in innovation also implies that more innovative firms work with more talented researchers (Lemma 1). In the appendix, I test this implication using firm- and inventor-level data from the USPTO. Using past innovation as a proxy for inventor talent, and various measures of firms' R&D efficiency, I show that, among a sample of job-switching inventors, the more talented ones tend to switch to more efficient firms, consistent with assortative matching. I now proceed to discuss my tests of the model's implications for the location and output of research centers.

4.1 Specification

I use mainly the following specification to test model predictions:

$$\log(y_{foi}) = \delta_f + \gamma_i + \beta_1 \gamma_i q_f + \beta_2 x_{oi} + \epsilon_{foi}, \tag{18}$$

where f, o and i are index for parent company, home country, and host country, respectively. The dependent variable, y_{foi}, is a measure of innovation output by company f's affiliated R&D center in country i (multiple affiliates in the same countries are aggregated into one). The first independent variable, δ_f, is the firm fixed effect, which controls for characteristics that are common to all R&D centers affiliated to the same firm. I exclude the firm fixed effect in some specifications to examine the effect of firm innovation efficiency. γ_i is a vector of host country characteristics that might affect offshore R&D and patenting, including size, relative talent abundance, per-capita income, intellectual property right protection (IPR), and general human capital. When these characteristics are not of primary interest, I use host country fixed effects instead. q_f is firm innovation efficiency. The interaction term $\gamma_i q_f$ captures how host country characteristics affect firms with different efficiencies. Of prime interest among these is the interaction between host country relative talent abundance and firm efficiency. x_{oi} is a vector of variables that vary across host-home pairs, including various measures of distance. When the interest is not in host country or bilateral characteristics, I use country-pair fixed effects to capture this term. ϵ_{foi} is the error term.

4.2 Data Description

I use patent data from the USPTO to construct three key measures in the specification: offshore R&D center innovation output, firm innovation efficiency, and host country relative talent abundance.

Firm and inventor classification To construct these measures, I need to be able to identify individual inventors and firms. This is challenging because patent data is self-reported, so there are no individual or firm identifiers. Moreover, typos and misspellings are frequent, and the same company might have different abbreviations. I follow the patent literature in addressing these

issues. For the firm side, I use the 2006 update of the disambiguated data set introduced by Hall et al. (2001), which covers the patents granted from 1976 to 2006. By combining automatic cleaning procedures—which take care of common abbreviations in company names—with manual checks, Hall et al. (2001) generates a unique identifier for each patent owner. For inventors, I use the unique inventor identifiers provided in Li et al. (2014), which uses a supervised learning approach to automatically generate inventor identifiers.

Offshore R&D output measure When applying for a patent at the USPTO, the applicant, usually the owner, reports address information for both the inventor and the owner of the proposed patent. I classify a patent as invented in a country-i offshore R&D center, affiliated to firm f from country o, if its inventor is in country i and its owner in country o.[32] Counting the total number of such patents by each firm in each host country, I obtain the benchmark measure for y_{foi}.

Firm innovation efficiency measure I use the total number of patents invented by firm f in its home country o as a proxy for its innovation efficiency. I focus on home country for this measure, and drop observations from the home country of each firm from the regression, so that the results are not driven by the mechanical correlation between home innovation and the measure of innovation efficiency. To reduce measurement error, in benchmark regressions, I classify a firm according to whether its innovation efficiency is above the median in its home country. Later I will show results with different cutoffs.

Host country relative talent abundance All research firms in a country compete for talent. It is the abundance of talented inventors *relative to* the abundance of efficient firms that matters for the type of inventor a foreign offshore R&D center is able to recruit. Following Definition 1, I construct the measure for inventor talent abundance as the share of inventors in a country that are in the top 1% most productive inventors in the world; I construct the measure for firm efficiency abundance analogously. I then use the log of the ratio between the two as the benchmark measure for relative abundance. Taking the ratio also nets out some of the differential selection across countries into patenting in the U.S.[33]

I use a relative quality measure, not relative quantity measures (e.g., the number of inventors relative to the number of firms), because the model predicts that a change in the relative quantity will have no impact on the matching function or the wage schedule.[34] In robustness, I also include this relative quantity. The choice to use the top 1% of inventors and firms in constructing this measure is motivated by the importance of exceptional inventors and firms in aggregate innovation. In robustness exercises, I use different cutoffs for computing the top shares, and other measures of quality in constructing the ratio.

---

[32]The information is reported at the time of patent application, so transfers of patents between owners from different countries are unlikely to be important.

[33]For example, patenting might be easier in some countries, so marginal firms and inventors self select into patenting, resulting lower measured average inventor and firm quality.

[34]By inspecting Equation 7, we can see that $\frac{L_i^R}{R_i}$ only enters as a ratio, and by taking derivatives of both sides with respect to z^R, we can see that this ratio does not affect T_i''; hence it will not affect the matching function or the shape of the wage schedule. The ratio $\frac{L_i^R}{R_i}$ will thus only affect the research team size and output of all R&D centers in country i proportionally.

Discussion on the use of patent data to measure R&D The advantages of using patent data for these measures are obvious: in addition to having a wide country coverage, patents are also highly correlated with firm-level R&D. Figure B.1 and Table B.1 in the appendix show that a patent-based offshore R&D measure correlates reasonably well across countries with a measure based on R&D expenditures. However, the drawbacks of patent data are also well known (Pavitt, 1988). First, the benefits of patents differ across countries, so that firms might have different incentives to apply for patents in the U.S. These differences might stem from market size, intellectual property right protection, or other country characteristics, such as connections to the U.S. Second, different industries have different reliance on patents for the protection of their intellectual properties. Third, patents have heterogeneous values, so patent counts are a noisy measure of firms' innovation output.

I add additional controls to address these concerns. Specifically, for the first concern, I either control directly for host country size, IPR protection, and other country characteristics, or simply include host country fixed effects. For the second concern, I use firm fixed effects to absorb firms' characteristics, including their industry. Moreover, I construct measures at the patent-category level so that host country specialization does not drive the results. Finally, to address the third concern, I also use citation as an alternative measure of innovation output.

Sample period The patent data spans 1976-2006. Since both the dependent variable, offshore R&D output, and the key independent variable, the relative talent abundance, are constructed based on patenting data, measurement error will lead to correlation between the two measures. To avoid this problem, I split the sample into two periods, 1976-1996 and 1997-2006. I use only information from the first period to measure the number and quality of innovating firms and individuals. I then use the 1997-2006 data to measure R&D output for each parent company and its foreign subsidiaries. Further, my regressions include only observations from *new* offshore R&D centers—those that enter in the second period—in order to prevent any R&D centers used as regression observations from affecting host talent quality measures. This sample split also prevents reverse causality, i.e., the entry of innovative and efficient foreign firms attracting more talented individuals to become inventors.

Additional data Additional variables used for the regressions are from the following sources: GDP, population, per-capita income, and a human capital index come from the Penn World Table 8.0; bilateral distance information is from the CEPII geodistance database (Mayer and Zignago, 2011); intellectual property protection information is taken from Park (2008). All these variables are averaged over 1997-2006 for consistency.

### 4.3 Main Results

Figure 4 provides evidence on the effect of firm innovation efficiency on offshore R&D through the extensive margin over the period 1997-2006, focusing on firms headquartered in the U.S. Each dot represents a firm. The horizontal axis is the number of patents granted to the firm and invented in the U.S. The vertical axis is the number of countries in which the firm performs R&D. The

figure indicates that firms with higher innovation efficiency tend to perform offshore R&D in more countries. Among the firms that enter the largest number of countries, IT and chemical companies are the most common.

Figure 4: Firm Efficiency and Offshore R&D Entry
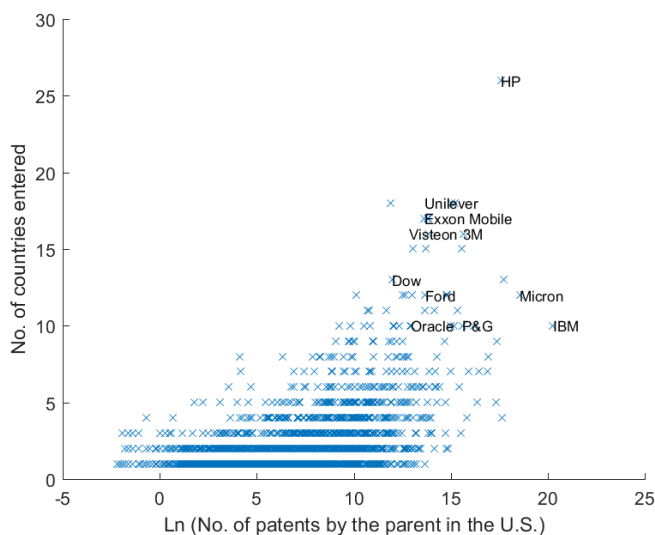
Notes: Each dot represents a firm headquartered in the U.S. The horizontal axis is the log of the number of patents the firm invented in the U.S. The vertical axis is the number of host countries it entered for offshore R&D, defined as $\sum_i \mathbb{I}_{y_{foi}>0}$.

I now estimate Equation 18 to test additional model predictions. Table 1 presents the baseline results. The first column includes the indicator for firm innovation efficiency, a vector of host country characteristics, bilateral distance measures, and home country fixed effects. Consistent with the first implication of the talent-acquisition motive, host country relative talent abundance has a positive and statistically significant impact on innovation output. The estimate has an elasticity interpretation: a 1% increase in host relative talent abundance increases firm-level offshore R&D by around 0.1%. Consistent with the market-access motive, host GDP also has a positive effect with a similar point estimate. Firms with above median innovation efficiency generate 63% higher R&D output, on average, so innovation efficiency increases offshore R&D through not only the extensive margin, but also the intensive margin. The estimate for host country intellectual property right protection is small and statistically insignificant, reassuring us that differential selection into patenting due to intellectual property protection differences are not driving the results. Host per-capita income does not have a significant effect. Distance measures are mostly insignificant, except for the common language indicator.

The second column adds the interaction term between the host country relative talent abundance and firm innovation efficiency. This interaction term is positive and statistically significant, with a point estimate of 0.15. While most other coefficients do not change, the coefficient for host relative talent abundance is no longer significant: consistent with the prediction from Proposition 2, the impacts of host talent quality are mainly concentrated in the top half of firms as ranked by innovation efficiency.

25

Results so far are supportive of a market-access motive and a talent-acquisition motive. Since the market-access motive is closely related to the extensive existing literature on the effects of market size on innovation (Acemoglu and Linn, 2004) and the location choice of multinational firms (Head and Mayer, 2004), I now focus on the talent-acquisition motive by further examining the interaction term. In the third column, I add host country and parent firm fixed effects to further absorb unobserved heterogeneity. The point estimate of the interaction term rises to 0.177, meaning that a 1% increase in relative talent abundance in the host country increases the R&D output by 0.17% more for R&D centers with above-median efficiency. After adding these better controls for host country and firm heterogeneity, bilateral geographic distance becomes significant, with an elasticity of -0.119. Common language, on the other hand, is no longer significant. The fourth column adds country pair fixed effects to capture differential economic connections between countries. The point estimate of the interaction term barely changes.

4.4 Robustness and Heterogeneous Effects

As discussed earlier, firms' incentives to perform R&D and to patent their R&D output are potentially affected by country characteristics. There may be plausible alternative theories that generate heterogeneous effects of these other characteristics for higher productivity firms. I now examine whether such alternatives can explain the baseline findings and generally find they cannot.

First, relative talent abundance might pick up an income effect. High-income consumers prefer high-quality products, which might be more R&D intensive than low-quality products. If efficient firms have comparative advantage in doing R&D, they might perform more R&D in high income host countries. I capture this by including the interaction between host country per-capita income and firm efficiency.

Second, the returns to both R&D and patenting are higher in large countries. Firms with higher efficiency might benefit disproportionately more because they tend to be more efficient in production. This concern motivates me to include the interaction between host country GDP and firm efficiency. Following the same reasoning, the effect of stronger patent protection enforcement might also benefit efficient firms more, encouraging them to patent more. Therefore I further include the interaction between firm efficiency and the IPR protection index.

Finally, the complementarity between talent and firm efficiency might happen in the production stage. Firms with better management can make better use of skilled workers in production, which reduces production costs and increases the return to R&D. I incorporate the interaction between the host country human capital index and firm efficiency to address this concern.

The first column of Table 2 reports the regression with these additional terms. The interaction between host talent and home efficiency is still significant, although it shrinks by about 40%. The variable that explains this drop is the interaction between host GDP and firm efficiency. Other interaction terms do not have strong effects.

I use a quality-based relative talent abundance measure in my baseline regressions because according to the model, the relative *quantity* of inventors and research firms will affect firms with

Table 1: Determinants of Offshore R&D: Baseline Results

Dependent Variable:	Log (Offshore patents invented in a host country)			
	(1)	(2)	(3)	(4)
Host relative talent abundance	0.091**	0.020		
	(0.039)	(0.041)		
I (Parent R&D >median)	0.625***	0.994***		
	(0.076)	(0.077)		
Host inventor relative abundance * I (Parent R&D >median)		0.147***	0.177***	0.171***
		(0.026)	(0.040)	(0.040)
Host GDP	0.083**	0.088**		
	(0.035)	(0.035)		
Host per-capita income	-0.004	-0.001		
	(0.076)	(0.075)		
Host IPR protection	0.032	0.026		
	(0.117)	(0.118)		
Distance	-0.007	0.003	-0.119**	
	(0.026)	(0.028)	(0.055)	
Common border	-0.009	0.003	0.203	
	(0.051)	(0.052)	(0.131)	
Common language	0.152***	0.151**	0.095	
	(0.056)	(0.056)	(0.094)	
Colonial tie	-0.075	-0.074	0.027	
	(0.049)	(0.051)	(0.092)	
Home country FE	X	X	X	
Host country FE			X	
Home-Host FE				X
Firm FE			X	X
Observations	14803	14803	7914	7716
R^2	0.050	0.053	0.454	0.490

Notes: The level of observation is host country-parent company. The dependent variable is the log of the total number of patents invented by an affiliate of a parent company in a host country over 1997-2006. For the I(Parent R&D >median) indicator, Parent R&D is measured by the total number of patents invented by the parent in its home country during the same period, and median is computed for all patenting firms in the home country of the parent company. *Host relative talent abundance* is defined as the log difference between the share of inventors in a host country that fall into the global top 1%, and the share of firms in that country that falls into the global top 1%. This measure is constructed using only patenting information for 1976-1996 to avoid mechanical correlation. *Host IPR* is the intellectual property right protection index from Park (2008), averaged over 1997-2006. Other host country characteristics are from the Penn World Table 8.0, averaged over 1997-2006.
Standard errors (two way clustered at the host-country and parent-company levels) are in parentheses.
* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

different efficiencies proportionally. To make sure the empirical finding is not the result of improperly measuring the relative talent abundance, in the second column, I add the interaction between firm efficiency and the ratio of the *number* of inventors and the number of R&D active firms. Reassuringly, this quantity ratio does not have a statistically significant impact itself and does not substantially change the interaction term between host relative talent quality and firm efficiency.

For a fuller picture of how host talent affects firms with different efficiencies, in the third column, I add indicators for firms with R&D efficiency above the 25th, 75th, and 90th percentiles of the R&D efficiency distribution in their home countries, as well as the interaction of these indicators with the full set of controls in the second column. The effect of a better host talent distribution is substantially larger for firms in the upper tail of the distribution. A 1% increase in host relative talent abundance leads to a 0.45% larger increase in R&D output for firms in the top 10% of the firm efficiency distribution than for firms in the bottom 25% of the distribution.

So far, all regressions patents pooled over all categories to construct measures for both the dependent and independent variables. Aggregation reduces measurement errors, but given that industries do not equally rely on patents for IPR, using aggregate patenting data might confound sectoral composition with country-level relative talent quality. For robustness, I also construct all variables within each individual patent category, classified by Hall et al. (2001).[35] For each firm, I keep only the category in which it patents most. Columns four and five perform regressions using category-level data. The fourth column controls for host-category fixed effects, and the last column controls for bilateral pair-category fixed effects. Both columns confirm that, the effect of host country talent is significantly larger for more efficient firms. Although the sample size is substantially smaller as more fixed effects are added, the coefficients are quantitatively similar to those in the third column, so the differential sectoral specializations of countries are unlikely to be the explanation for the benchmark results.

In Table B.2 in the appendix, I report robustness checks using alternative measures of host relative talent abundance, R&D center innovation output, and firm innovation efficiency. First, to address concerns about differential selection into patenting, and about the heterogeneous value of patents, I use citation counts, instead of patent counts, to measure R&D center innovation output and firm efficiency. Second, the benchmark measure for host talent abundance is somewhat ad-hoc, so I also use the following alternatives: 1) define "top" inventors and "top" firms as being the top 10% as opposed to the top 1%; 2) use the ratio between per-inventor patent counts and per-firm patent counts; 3) use the ratio between the standard deviation of inventor patent counts and the standard deviation of firm patent counts.[36] In all regressions, I include the control variables used in the second column of Table 2. The point estimates in these regressions are all similar to those in the benchmark estimate.

In summary, this section presents evidence that is consistent with the main predictions of the

---

[35]There are in total six categories: chemical (excluding drugs), computers and communications, drugs and medical, electrical and electronics, mechanical, and others.

[36]A higher standard deviation captures a higher share of inventors with a large quantity of patents.

Table 2: Determinants of Offshore R&D: Alternative Explanations and Heterogeneous Effects

Dependent Variable:	Log (Offshore patents invented in a host country)				
	Aggregated Across All Categories			Only Firms' Main Category	
	(1)	(2)	(3)	(4)	(5)
Host relative talent abundance * I (Parent R&D >25%)			0.131	0.107	0.187
			(0.084)	(0.142)	(0.158)
Host relative talent abundance * I (Parent R&D >median)	0.105*	0.093*	0.117	0.134	0.217
	(0.054)	(0.051)	(0.091)	(0.159)	(0.175)
Host relative talent abundance * I (Parent R&D >75%)			0.208**	0.126	0.217
			(0.095)	(0.148)	(0.185)
Host relative talent abundance * I (Parent R&D >90%)			0.465***	0.354*	0.579**
			(0.146)	(0.173)	(0.217)
Host per-capita income * I (Parent R&D >median)	-0.093	-0.063			
	(0.148)	(0.146)			
Host GDP * I (Parent R&D >median)	0.119***	0.101**			
	(0.039)	(0.043)			
Host human capital * I (Parent R&D >median)	-0.102	-0.161			
	(0.214)	(0.224)			
Host IPR * I (Parent R&D >median)	0.167	0.155			
	(0.196)	(0.197)			
Host inventor relative *quantity* * I (Parent R&D >median)		0.125			
		(0.146)			
Home-Host FE	X	X	X	X	X
Firm FE	X	X	X	X	X
Host-Category FE				X	
Host-Home-Category FE					X
Host Characteristics *Full set of firm efficiency indicators			X	X	X
Observations	7687	7687	7687	4381	3907
R^2	0.490	0.490	0.496	0.562	0.605

Notes: See Table 1 for descriptions of variables and sample period. The first three columns use information aggregated over all patent categories. The last two columns use variables similarly constructed at the patent category level. For each firm, only the category in which it generates the most patents is used in regressions. Although not reported, columns 3-5 also include the interaction of host country characteristics in column 2 with the full set of firm R&D efficiency indicators.

Standard error (two way clustered at the host-country and parent-company levels) are in parentheses.

* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

model in terms of how host country relative talent abundance affects firms with different R&D efficiencies. Additional exercises that consider alternative explanations and measurements do not substantially weaken the evidence.

5 Parameterization

I now perform a quantitative analysis of the model outlined in Section 3. I focus on a sample with 25 countries and a statistical aggregation of another 22 countries.[37] I parameterize the model to be consistent with the data in its predictions on the international interactions between countries and the size distribution of firms within the U.S. This section describes the parameterization procedures, starting with the functional form assumptions.

5.1 Additional Assumptions

In the quantification, I embed an occupational choice into the model. Throughout the rest of the paper, I assume that each country is endowed with L_i number of workers, with talent distribution $H_i(\theta)$. Workers sort into production labor or research. Each production worker has one unit of production labor, and each researcher has θ units of talent in research. Adding occupational choice generates endogenous responses in the supply of inventors in counterfactuals experiments.

The function $f(z^R, \theta)$ determines the complementarity between the innovation management efficiency of firms and the talent of researchers. I assume that f is a CES function with elasticity of substitution $\alpha < 1$:

$$f(z^R, \theta) = (z^{R\frac{\alpha-1}{\alpha}} + \theta^{\frac{\alpha-1}{\alpha}})^{\frac{\alpha}{\alpha-1}}.$$

This specification satisfies the log-supermodularity assumption. As α approaches 1, the complementarity between researcher talent and firm efficiency weakens.

The distributions of worker talent and firm innovation efficiency are parameterized to be truncated Pareto distributions, given below:

$$H_i(\theta) = \frac{(\underline{\theta}_i^{-\kappa_i^\theta} - \theta^{-\kappa_i^\theta})}{(\underline{\theta}_i^{-\kappa_i^\theta} - \overline{\theta}_i^{-\kappa_i^\theta})}, \ G_i^E(z^R) = \frac{(\underline{z}_i^{R-\kappa_i^R} - z_i^{R-\kappa_i^R})}{(\underline{z}_i^{R-\kappa_i^R} - \overline{z}_i^{R-\kappa_i^R})}.$$

In these expressions, the letters with upper and lower bars indicate the upper and lower bounds for their respective distributions. κ_i^R and κ_i^θ are the truncated-Pareto counterparts to the shape parameter in the Pareto distribution.

To capture the correlation between innovation and production efficiency at the firm level, I assume that there are two distributions, indicated by H and L (for high and low, respectively), from which firms draw their productivity z^P. The probability of drawing from the high distribution

---

[37]The list of countries in this statistical aggregation is provided in the appendix.

depends on a firm's innovation efficiency in the following fashion:

$$\text{Prob}(z^P \in H | z^R) = \frac{\exp(A + B \times z^R)}{1 + \exp(A + B \times z^R)},$$ (19)

where A and B are parameters to be estimated. A positive value for B means that more innovative firms tend to be more productive as well. H and L are both Pareto distribution with the same shape parameter κ^P:

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

I assume that $\underline{z}_L^P < \underline{z}_H^P$, so the H distribution first-order stochastically dominates the L distribution.

5.2 Parameters Assigned Directly

I set the number of workers in a country, L, to total employment from the Penn World Tables. To focus on differences in the firm efficiency distributions and to abstract from the differences in the number of firms, I set E, the measure of firms, to be proportional to L. This proportion is chosen so that the average employment per firm in the model equals the average employment per firm in the U.S.

I directly assign values to a few parameters in the model. Parameter σ, the elasticity of substitution between varieties, determines the markup charged by firms. I set this parameter to be 5, following recent studies in international trade (see Simonovska and Waugh, 2014, for example). This value also implies that 20% of sales are variable profits. In the U.S., R&D expenditures account for about 8% of manufacturing sales. The model counterpart of R&D expenditures is researcher compensation. I set γ, the share of researcher compensation in variable profit, to 0.4, so that researcher compensation accounts for about 8% of sales in the model.

Equation 14 implies $\log(X_{ild}) = \alpha_{id} + \beta_{il} - \delta \log(\tau_{ld})$, where α_{id} and β_{ld} are pair fixed effects. δ therefore determines the elasticity of X_{ilds} with respect to the cost of shipping from l to d. Based on Arkolakis et al. (2014), which estimates this specification using the affiliate production and sales data of U.S. multinationals ($i = U.S.$), I set δ to 10.9.

Calibrating firm efficiency and worker talent distributions for each country requires comparable data across countries. I use the World Management Survey by Bloom et al. (2012) and the cognitive test score data by Hanushek and Woessmann (2012) to calibrate these distributions.

The World Management Survey provides firm-level management scores for each country in the sample. 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 then averaged over these sub scores. The talent management score intends to capture whether firms follow good managerial practice for retaining and incentivizing its talent, so it is closely related to whether a firm is able to make full use of its research talent. I use it to calibrate innovation management distributions. I obtain three distribution statistics of $\tilde{z}^R$ for each country: mean, standard deviation, and dispersion.

Table 3: Parameters Calibrated Externally

Symbol	Descriptions	Value	Source
σ	Elasticity of substitution between varieties	5	Simonovska and Waugh (2014)
γ	Return to research team size in R&D	0.4	Manufacturing R&D share
δ	Dispersion in offshore production efficiency draws	10.9	Arkolakis et al. (2014)
A	Probability of having a high production efficiency	-6.3	Estimated
B	Dependence of z^P on z^R	0.17	Estimated

I use firm-level talent and production management scores to estimate A and B, the coefficients linking a firm's innovation and production efficiencies. Specifically, I classify a firm as being from a high productivity distribution if its production management score falls into the top 1% of the distribution in the world (the top 4% 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 model given by Equation 19. This procedure determines $A = -6.3$, $B = 0.167$.[38]

For the talent distribution, I obtain average cognitive score and the share of students reaching "basic" and "top" performance from the test score database, which are defined based on a common absolute level across countries. To pin down the relative scale of management efficiency and talent, I take the U.S. as the benchmark. Specifically, I set $H_{US}(\theta)$ and $G^R_{US}(\tilde{z}^R)$ to be the same, and use the three statistics on the talent management score to pin down all three parameters in $G^R_{US}(\tilde{z}^R)$.[39] I then determine the distributions for other countries, by relating their distribution statistics to those of the U.S.[40]

Table 3 summarizes the information on the parameters determined directly. I choose additional parameters jointly in equilibrium, a process I describe below.

5.3 Parameters Determined in Equilibrium

Overview The remaining parameters to be determined include international frictions, $\{\tau_{ld}\}$, $\{\phi^R_{oi}\}$, $\{\phi^P_{il}\}$, $\{c^M_d\}$, and $\{c^R_i\}$; country-specific productivity, $\{\Lambda_l\}$; production efficiency distri-

---

[38]The choice of the top 1% cutoff is motivated by the importance of the most productive firms in international business and in production in general. A high cutoff allows me to better capture the distribution of the very top firms. The implicit assumption underlying this calibration strategy is that firms drawing their production efficiency from the L distribution constitute the bottom 99% in the production efficiency distribution, whereas firms drawing from the H distribution constitute the top 1% of production efficiency. This assumption does not hold exactly because under the Pareto assumption, $G_L(z^P)$ will always overlap with $G_H(z^P)$. Given the choice of the cutoff (1%), however, the calibrated $\underline{Z}^P_H$ will be large enough so the overlap is negligible.

[39]The talent management score is approximately normal in the data. Since it is well known that the firm size distribution has a fat tail, I take the exponential of original scores and use that to match firms' innovation efficiency distribution. The statistics I use to pin down each country's distribution are based on these exponents of scores. A few countries in the quantitative analysis are not covered by the World Management Survey. I impute their statistics based on country characteristics. The calibration appendix reports the procedures used in the imputation process.

[40]For firm innovation distributions, the three moments can be perfectly matched by the three parameters in the truncated Pareto distribution. For the talent distribution, however, the truncated Pareto distribution cannot perfectly match all three moments. I therefore use only the average score and the top student share to pin down the upper bound and the shape parameter, while setting the lower bound to be the same across countries. This simplification, however, does not leave out important information, as the correlation between the share of students reaching basic performance and the average score is 0.92.

bution parameters, $\underline{z}_L^P$, $\underline{z}_H^P$, and κ_P; and complementarity between management and talent, α. Although in equilibrium these parameters are jointly identified, for certain parameters some moments are more informative than others. I describe below how each parameter is determined.

The iceberg components of international frictions, $\{\tau_{ld}\}$, $\{\phi_{oi}^R\}$, and $\{\phi_{il}^P\}$, determine the aggregate flows of international integration. I use them to match bilateral trade shares, offshore R&D shares, and offshore production shares. The data sources for these bilateral relationship include: the multinational production data sets introduced in Ramondo et al. (2015); bilateral trade including domestic absorption from the World Input-Output Database; and bilateral offshore R&D information based on patenting statistics at USPTO from the OECD patent database. To reduce measurement errors, I average the bilateral patenting and trade data over the period of 1998-2007.[41]

The fixed components of international costs, $\{c_d^M\}$ and $\{c_i^R\}$, determine the extensive margin of firms' global operation. Due to lack of this information for a large sample of countries, I assume that these fixed costs are the same for all country pairs, and choose them to match the share of exporters (0.35) and the share of foreign affiliates among research active firms (0.037) in the U.S. manufacturing sector, respectively. I calibrate $\{\Lambda_l\}$, the labor productivity in production, by matching the real per-capita income of each country.

I normalize $\underline{z}_L^P$ to 1, and determine α, $\underline{z}_H^P$, and κ_P jointly. α affects both the pattern of matching between firms and researchers, and the firm size distribution. Strong complementarity (small α) puts efficient firms at an advantage in working with talent, which affects the shape of the matching function and the concentration of researchers. Figure 5a plots the model matching function under various α. The matching functions corresponding to smaller α tend to be more convex, with a larger share of researchers working for the top firms.[42] I measure the overall convexity of the matching function using the ratio between the average slope of the matching function for the top 50% z^R firms, and the average slope for the bottom 50% firms. This convexity conveys information about the value of α, and will be used as a calibration target. I discuss below how I construct the model and empirical measures for this convexity.

By determining the distribution of talent across firms, α also affects the number of products a firm develops, and hence the firm size distribution. Numerically, it is mostly informative about the size of firms in the top 1%. In addition to α, κ_P and $\underline{z}_H^P$ are also important about the firm size distribution: κ_P directly affects the Pareto shape of the firm size distribution at the very top, while $\underline{z}_H^P$ effectively determines the scale of the top 4% firms relative to the bottom 96%, as about 4% of U.S. firms draw from the H distribution.

Specifics about the matching function I estimate parameters of matching function based on the test of positive assortative matching between inventors and firms, presented in the appendix. Specifically, I measure firm innovation efficiency using the per-inventor innovation output, and

---

[41]The multinational production database is averaged over 1996-2001 in the original source.

[42]From Equation 7, other things equal, the slope of the matching function reflects the size of the research teams. The steeper the curve, the larger is the research team. A more convex matching function thus means a more unequal distribution of research team size, similar to the Lorenz curve.

inventor talent using past innovation. Focusing on a sample of job switchers, I then estimate nonparametrically how the talent of an inventor is related to the innovation efficiency of the new firm, controlling for inventor and firm characteristics as well as time and patent category fixed effects. A positive correlation indicates positive assortative matching. The solid line in Figure 5b presents the estimates, along with a 2 s.e. band. The overall convexity measure of this empirical matching function is 1.71.

We cannot directly compare the model convexity measure to its data counterpart. In the data, matches are noisy, so the range of the estimated matching function is not $[1, 100]$, whereas in the model, this is always the case. To make the two comparable, I take the stand that in choosing the optimal types of researchers, firms make mistakes. They cannot differentiate workers whose talent satisfies firms' first order condition (Equation 6) within a certain "error margin". I fix the wage schedule at the benchmark equilibrium, and then choose the size of this margin so that the estimated matching function using simulated data has the same range as the empirical matching function. I then compute the convexity measure based on this simulated matching function.[43]

Figure 5b plots the simulated noisy matching function when α is 0.7, which will be the benchmark calibration, and two different values. The benchmark value offers the best fit for the overall concavity, determined by the value of the matching function at the 50th firm percentile. A smaller α could fit the overall shape reasonably well, but misses the top range. A larger α, on the other hand, is a poor fit overall.

5.4 Computational Algorithm

A detailed account of the computational algorithm is provided in the appendix. This section briefly describes the nested procedure I use. In the outer loop, I choose $\underline{z}_H^P$, κ^P, c^M, c^R, and α to match the targets described above. In the middle loop, I iterate over $\{\tau\}$, $\{\phi^R\}$, $\{\phi^P\}$, and $\{T\}$ to match all bilateral shares and per-capita real income of countries. The inner loop solves the model given exogenous parameters.

This computation algorithm requires solving the researcher market equilibrium for all countries at different parameter values. With offshore R&D decisions in the model, the distribution of R&D center innovation efficiency, $g_i^R(z^R)$, is an endogenous outcome. The cutoff rule in offshore R&D decisions implies that $g_i^R(z^R)$ could have multiple discontinuities. As a result, the matching function, $T_i(z^R)$, is not necessarily differentiable. In this case, general boundary value problem solvers routinely fail or takes a long time to find the solution. In the appendix, I develop a computational algorithm that is well-suited for this exercise.

Figure 5: The Model and Empirical Matching Function

(a) Model matching function
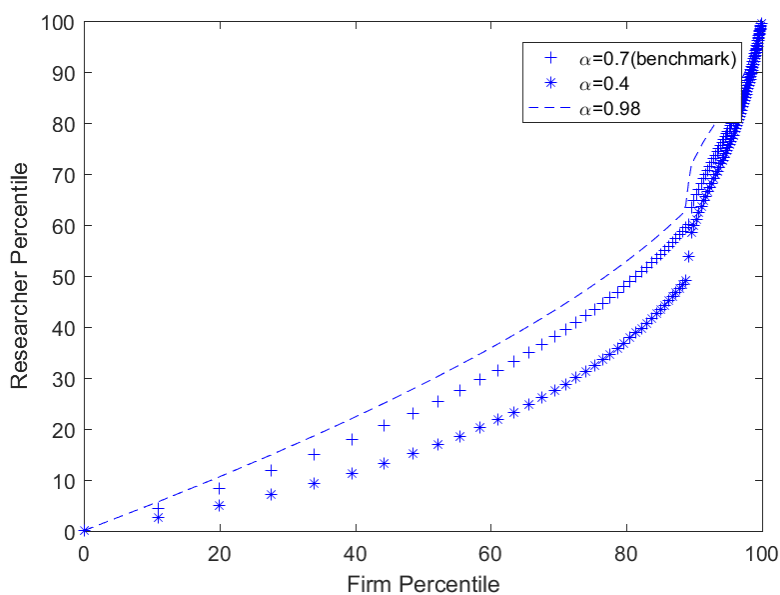
(b) Empirical and "noisy" model matching function
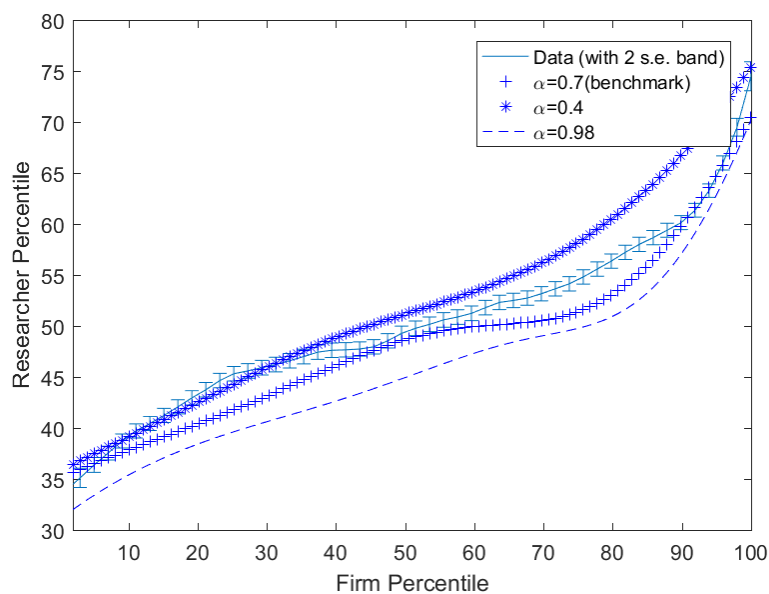
Notes: In both panels, the horizontal axis plots percentiles of firm innovation efficiency, and the vertical axis plots percentiles of researcher talent. The upper panel is the model matching function under different α. The lower panel shows the empirical matching function estimated by the author using the USPTO data (solid line), and simulated "noisy" model matching functions under different α.

Table 4: Fit of the Targeted Moments

Parameter	Value	Moments	Model	Data
c^M	0.0693	Share of exporters	0.35	0.35
c^R	2.6	Share of foreign affiliates	0.042	0.037
	$z_H^P = 1.2$	Fraction of firms with emp.<100	0.99	0.99
$\underline{z}_H^P, \alpha, \kappa^P$	$\alpha = 0.7$	Fraction of firms with emp.<20	0.95	0.95
	$\kappa^P = 8.16$	Matching function slope between 0%-50%/ slope between 50-100%	1.58	1.71
		Share of emp. in firms with>500 emp.	0.41	0.47
		Power law coefficient of firm size dist.	1.04	1.05

5.5 Model Fit

The calibration process determines that $c^M = 0.069$, $c^R = 2.6$, $z_H^P = 1.2$, $\kappa^P = 8.16$, and $\alpha = 0.7$. Table 4 reports the value of parameters and the model moments that help pin down these parameters. Overall, the model is able to fit data along these dimensions well.

The value of α suggests strong complementarity between innovation efficiency and researcher talent. Since the complementarity and the resulting talent-acquisition motive are an important channel in the model, in the following, I first discuss the role of α in determining the model predictions and explaining the patterns in the data. I then present additional implications of the model under the benchmark calibration and compare them to the data whenever possible.

Figure 6: Complementarity and Offshore R&D
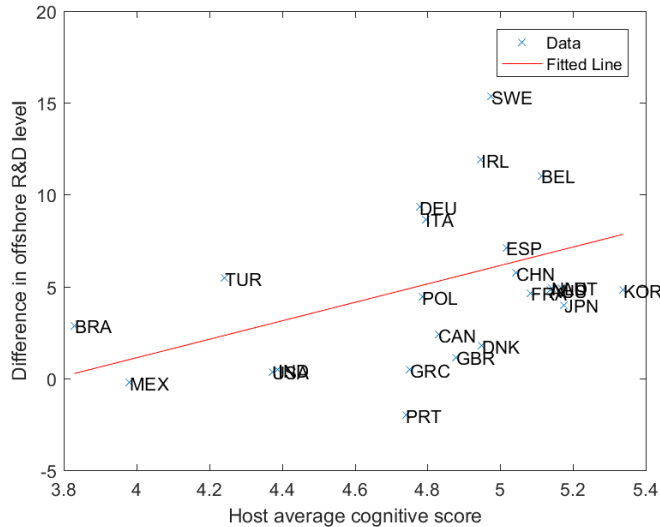
Notes: The vertical axis plots the percentage point difference between the benchmark parameterization and an alternative parameterization with $\alpha = 0.98$ in the share of R&D expenditures by foreign affiliates. The horizontal axis is host average talent. Host average innovation efficiency is netted out from both axis.

---

[43] For different α, the size of the "error margin" needed to match the range of the empirical matching function varies. But as long as $\alpha < 1$, the simulated matching function can always match the range of its empirical counterpart. When α approaches 1, firms become increasingly indifferent between different researchers. A small amount of mistakes in recruiting would then result in a flat matching function.

The importance of complementarity To understand the role of complementarity in shaping offshore R&D between countries, I solve a counterfactual experiment with $\alpha = 0.98$, keeping other parameters at the benchmark. This parameter value implies much weaker complementarity than the benchmark calibration. The vertical axis in Figure 6 shows the percentage point difference between the benchmark and the counterfactual equilibrium in the share of domestic R&D done by foreign affiliates. The horizontal axis is host average talent quality. The figure indicates that higher complementarity increases offshore R&D, particularly in host countries with high talent, so complementarity is an important force for the pattern of offshore R&D.

The calibration suggests that a relatively strong complementarity ($\alpha = 0.7$) fits the pattern of matches and moments of the firm size distribution well. Does it also explain the pattern of offshore R&D better than under weak complementarity (as α approaches 1)? Although not reported here, I find that the empirical results reported in Section 4 using firm-level patent data also hold at the aggregate level. Specifically, offshore R&D between a country pair increases with home innovation efficiency, host talent, and their interaction term.[44] I evaluate the model's ability to generate these features under $\alpha = 0.7$ and $\alpha = 0.98$.

Because the calibration exactly matches bilateral offshore R&D and the distribution of talent and innovation efficiency in the cross section, I evaluate the model in changes. I simulate a counterfactual equilibrium in which countries receive random shocks to their distributions of talent or efficiency.[45] I then use the simulated data to perform a difference-in-difference regression of changes in offshore R&D on changes in host talent distribution and home innovation efficiency, in which each pair of country is an observation.

Table 5: Complementarity and the Patterns of Offshore R&D: Simulated Data

| | Benchmark Calibration | | | Weak Complementarity | | |
	(1)	(2)	(3)	(4)	(5)	(6)
Log(average home mgt. efficiency)	0.157***		-0.650*	0.010***		-0.018
	(0.044)		(0.375)	(0.003)		(0.025)
Log(average host talent)		0.071*	-0.625**		0.006**	-0.020
		(0.038)	(0.298)		(0.003)	(0.019)
Interaction			0.353**			0.012
			(0.164)			(0.011)
Bilateral FE	X	X	X	X	X	X
Observations	1352	1352	1352	1352	1352	1352
Within R^2	0.019	0.005	0.025	0.019	0.007	0.021

Standard errors in parentheses

* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Columns 1-3 of Table 5 report the results under the benchmark specification. Bilateral pair fixed effects are included in all three columns, so the model is identified from changes. The first two columns show that host country talent and home country innovation efficiency both have

---

[44]Results from country-level regressions are available upon request.

[45]Specifically, I reduce the upper bound of the talent distribution by a random fraction for one third of the countries, reduce the upper bound of innovation efficiency by a random fraction for one third of the countries, and then keep the remaining one third of countries intact.

significant positive impacts. The third column adds an interaction term. The interaction is positive and significant, while the non-interactive terms turn negative. So the effects are concentrated in the pairs of countries that experience improvements in both host talent and home efficiency, consistent with the empirical findings. Columns 4-6 of Table 5 report the same specifications under the case where $\alpha = 0.98$. In this case, host talent and home efficiency both have significant marginal impacts on offshore R&D. The within R square terms in the first two columns are also similar to those under the benchmark specification. In contrast to the empirical finding, however, the interaction term is not significant.

Together, these results suggest that under the benchmark calibration, the model is able to generate a relationship between offshore R&D and the distributions of endowments similar qualitatively to that observed in the data, while a model without complementarity cannot.

Table 6: Additional Untargeted Moments

	Model	Data
Management Score and Firm Size		
Management score difference between large and small firms	1.18	1.32
The Management Efficiency of Foreign Affiliates		
Foreign affiliate advantage	1.33	1.16
Coefficient of variation across countries	0.094	0.075
Correlation with domestic average score	-0.67	-0.84

The management score difference between large and small firms The calibration procedure for κ_P and z_H^P takes the stand that management efficiency differences are the fundamental cause of performance differences among firms.[46] To validate this assumption, I examine the model's performance in matching the mapping from management score to firm size. This is a valid test because the calibration uses the information on management scores and the firm size distribution separately, but puts no restrictions on how a one-point increase in the management score at different percentiles of the firm size distribution translates into increases in firm size. For this comparison, I use the total management score, defined as the sum of the innovation and production score, for a consistent comparison with the empirical evidence.

The first panel of Table 6 reports the difference in total management score between firms with 10000 employees and firms with 10 employees for the model and data.[47] In the model, the difference in total management score between an average firm with 10000 employees and an average firm with 10 employees is 1.18 times the standard deviation of the management score, which is close to the empirical counterpart of 1.32.

Figure 7 plots nonparametrically the relationship between management score and firm size from the model and the data. The estimated curve from the data, in the left panel, displays some

---

[46]The calibration essentially takes the management score distribution from data, and chooses κ_P and z_H^P so that the variation in firm size is close to that in the data.

[47]The empirical counterpart of this number is from Bloom et al. (2014), which estimates this relationship nonparametrically, focusing on medium-sized U.S. manufacturing firms with 10-10000 employees. Because the two surveys have different scales for scoring, I normalize the increase by the standard deviation of total management score.

Figure 7: Management Score Difference Between Large and Small Firms

(a) Data
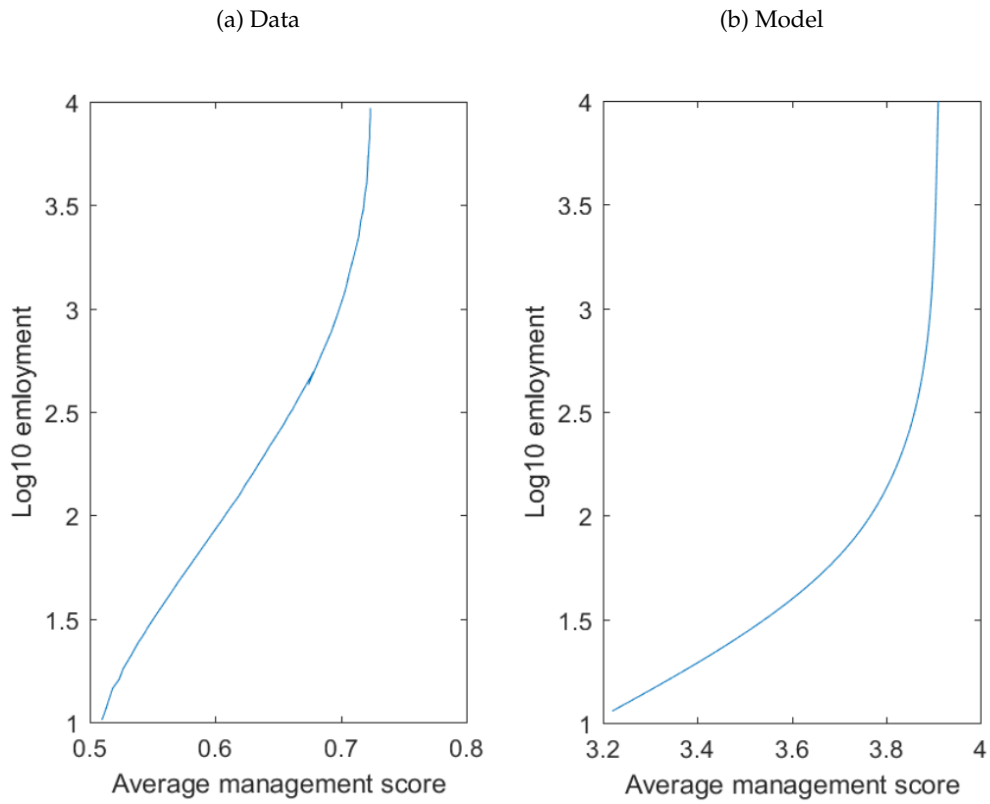
(b) Model

Notes: The left panel shows the model relationship between management score and firm employment in the data estimated in Bloom et al. (2014); the right panel shows the model counterpart. Both are based on the sub-sample of firms with employment between 10 and 10000. The range of variation in x-axis is 1.18 times the standard deviation of the management score in the data, and 1.32 in the model.

convexity: initially, firm size increases relatively slowly with management score; at the top range, however, a small increase in management score results in a larger percentage increase in firm size. Such a relationship can always be captured by the model by choosing how management score scales into productivity, which is partially determined by z_H^P. The question is whether the scale chosen to match other moments is able to generate this relationship. The right panel is the model relationship between management score and employment. Consistent with the data, the model also generates some convexity.

The multinational managerial advantage One important assumption of the model is that affiliates' innovation efficiency depends on that of their parents, rather than that of host country domestic firms. This assumption, together with the self-selection mechanism, implies that foreign affiliates tend to be more management efficient than domestic firms, and that the managerial advantage of foreign affiliates is larger in countries with worse domestic innovation management efficiency.

I validate these implications quantitatively by calculating the foreign affiliate managerial advantage for each country. The measure I use is the ratio between average foreign affiliate innovation efficiency and average domestic firm innovation efficiency, . I then compare the statistics of this measure among the sample countries to their data counterpart, constructed using the database introduced in Bloom et al. (2012).

The bottom panel of Table 6 reports the statistics of the foreign affiliate managerial advantage for the sample countries. Both the model and the data indicates a larger innovation management score for foreign affiliates compared to domestic firms, although the difference is larger in the model (33%) than in the data (16%). The variability of the foreign affiliate advantage measure across countries, captured by its coefficient of variation, is 0.094 in the model, and 0.075 in the data. The correlation between this measure and the host country average domestic innovation score is -0.67 in the model, and -0.84 in the data. So quantitatively, the model fits the cross-country pattern of foreign affiliate innovation advantage well. In the appendix, I also plot the model foreign affiliate managerial advantage against its data counterpart for individual countries that are common to both samples. Overall the model is a reasonable fit.

Share of non-production income in GDP The model predicts countries will specialize differently in R&D or production. Figure 8 plots the share of income from non-production labor in the model against its counterpart in the data, the share of R&D in GDP. There is a strong correlation between the model and the data across countries, even though the model best captures the manufacturing industry while the data is from the aggregate economy.[48]

International Frictions Finally, I check if the calibrated bilateral frictions are reasonable by comparing their correlations with geographic distance. The correlations between the logs of τ, ϕ^R, and ϕ^P and the log of distance are 0.2, -0.22, and -0.42, respectively. The signs of these correlations are consistent with larger international frictions for longer distances (ϕ^P and ϕ^R are the inverse of costs). The difference between offshore production and offshore R&D in distance

---

[48]The model prediction better matches the ratio between R&D expenditures and manufacturing value added.

Figure 8: Share of R&D in Income: Model versus Data
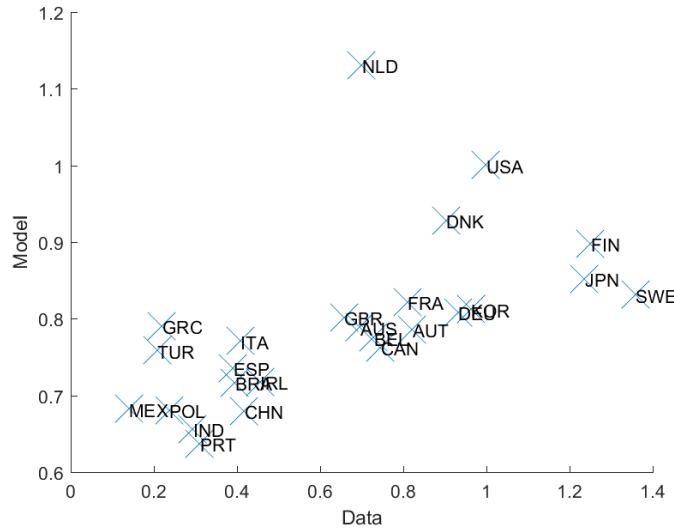
Notes: This figure plots the share of income from non-production labor in the model against the share of R&D in GDP in the data across countries. The measure for the U.S. is normalized to have 1 in both the model and the data.

elasticity also supports that these two activities are different in nature.

6 Counterfactual Experiments

In this section I perform counterfactual experiments using the parameterized model to shed light on the determinants and impacts of offshore R&D.

6.1 What Determines Offshore R&D

I first examine the quantitative importance of the talent-acquisition and market-access motives for offshore R&D. This is a relevant exercise, because policy makers around the world are looking to attract R&D intensive FDI. Domestic research talent and access to foreign countries through trade and offshore production are cited as important determinants of the attractiveness of a country as a host for offshore R&D centers (Guimón, 2009). I perform a set of experiments in which I either change the distribution of talent or management endowments, or the market access of a country. To isolate the effects from changes in other countries, when computing these counterfactual equilibria, I change parameters for one single country at a time, keeping model parameters at the benchmark for all other countries.

The role of endowment distributions The first set of experiments aim to quantify the importance of cross-country differences in the distributions of firm efficiency and researcher talent in determining offshore R&D. Specifically, I increase innovation efficiency of each host country, and decrease the talent of their workforce, to see how these two factors affect the equilibrium offshore R&D. I choose the U.S. innovation efficiency distribution and the Brazilian talent distribution as

Table 7: The Determinants of Offshore R&D

| Country | Benchmark | "Talent Acquisition" | | | "Market Access" | | | All |
| | | Efficiency | Talent | Both | Consumers | Producers | Both | |
	1	2	3	4	5	6	7	8
Developed								
AUS	26.83	0.05	7.61	0.00	30.30	17.83	0.27	0.00
AUT	50.21	5.87	25.81	0.00	51.46	39.88	0.00	0.00
BEL	57.12	12.25	14.98	2.48	68.15	31.62	0.00	0.00
CAN	33.52	13.38	12.19	0.13	38.88	22.73	0.00	0.00
DEU	23.85	4.22	7.53	1.03	36.32	11.51	1.65	0.00
DNK	33.55	0.93	15.98	0.13	41.22	18.88	0.00	0.00
ESP	42.92	0.28	30.96	0.00	45.64	41.03	29.53	0.00
FIN	17.93	0.00	0.46	0.00	26.86	1.62	0.00	0.00
FRA	33.74	1.00	17.66	0.31	40.32	23.61	2.89	0.00
GBR	45.65	17.83	29.04	7.13	54.25	33.23	5.63	0.08
GRC	58.00	10.46	53.50	5.72	56.90	57.02	49.54	1.37
IRL	55.20	30.23	29.72	0.01	55.21	51.22	0.00	0.00
ITA	29.20	0.33	22.11	0.21	32.05	26.31	19.19	0.10
JPN	5.05	0.00	2.03	0.00	9.00	2.39	2.35	0.00
KOR	4.81	0.00	0.00	0.00	8.80	3.34	1.36	0.00
NLD	34.66	1.13	2.83	0.07	54.67	0.81	0.00	0.00
POL	60.79	31.85	49.68	21.07	60.72	60.52	43.06	0.25
PRT	50.21	0.14	41.17	0.00	60.27	46.95	0.00	0.00
SWE	26.97	0.65	3.22	0.00	31.92	7.07	0.00	0.00
USA	7.93	7.93	2.97	2.99	13.89	1.92	1.18	0.45
Average	34.91	6.93	18.47	2.06	40.84	24.97	7.83	0.11
Emerging								
BRA	37.14	5.37	37.14	5.37	40.15	36.49	34.23	4.68
CHN	52.13	22.66	46.33	19.40	53.46	52.12	51.28	18.73
IND	57.88	35.06	57.08	34.66	58.67	57.74	56.96	33.99
MEX	49.26	24.32	49.49	24.40	54.50	48.07	36.78	14.80
TUR	51.79	12.04	46.24	9.38	52.43	51.60	49.40	8.37
Average	49.64	19.89	47.25	18.63	51.84	49.20	45.73	16.11
Average	37.85	9.52	24.23	5.38	43.04	29.82	15.41	3.31

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 percentage points. The first column shows the results from the benchmark calibration. The second column changes the firm innovation efficiency distribution for each country to that of the U.S. The third column changes the worker talent distribution for each country to that of Brazil. The fourth column combines the changes in the second and third columns. The fifth column increases exporting costs to infinity. The sixth column increases countries' outward offshore production costs to infinity. The seventh column combines changes in the fifth and sixth columns. The last column combines changes in the fourth and seventh columns.

the benchmark, because these two countries respectively have the highest management efficiency, and the lowest average talent.

The second column in Table 7 reports the share of R&D done by foreign affiliates for each host country when it is given the U.S. management efficiency distribution. With an improvement in domestic management efficiency, domestic firms are more competitive in both labor and product markets. Domestic wages increase and prices decrease, reducing foreign firms' incentive to enter. Indeed, compared to the share of foreign R&D in the benchmark equilibrium in the first column, the shares in these counterfactual equilibria are much lower. The average share of R&D done by foreign affiliates across all countries is 9.52% , or a quarter of the benchmark value. The decreases are smaller for the large developing countries in the sample—China and India. This likely occurs for two reasons: first, these countries have relatively large domestic market so product market is an important consideration, and second, these countries are further away from major innovation countries, so offshore production to these countries is more costly.

I then change each country's talent distribution to that of Brazil, while keeping its firm efficiency distributions at the benchmark. Intuitively, when domestic talent distribution improves, R&D outputs of both domestic and foreign-affiliated R&D centers increase. The increase in the latter is larger for two reasons. First, foreign affiliates are on average more productive, so they benefit more from the improvement in researcher quality. Second, the increase in R&D output allows more foreign firms to overcome the fixed costs and enter. Column 3 of Table 7 reports the share of R&D by foreign affiliates in each country. On average, foreign affiliates account for about 25% of domestic R&D, which is a decrease of around one-third from the benchmark value. The size of the decrease, again, varies considerably across countries. Perhaps because developed countries had higher talent distribution to begin with, they experience larger drops in inward offshore R&D.

Finally, I combine the two experiments by changing the distributions of both management efficiency and talent. As can be seen from the fourth column of Table 7, the global average share of R&D done by foreign affiliates is around 5%, about one-fifth of the benchmark value. Overall, cross-country differences in the distributions of talent and firm efficiency can account for most of the observed offshore R&D for developed countries, and a smaller but still significant share for large developing countries.

The role of foreign access I now examine the impact of the host country's access to foreign countries on offshore R&D. In the model, foreign access consists of two channels: access to foreign consumers through exporting, and access to foreign producers through offshore production. I consider their separate and joint impacts.

In the first experiment, I increase each host country l's iceberg export cost, $\tau_{ld}, l \neq d$ to infinity. This shuts down host countries' direct access to foreign consumers, but R&D centers there can still indirectly access foreign consumers through offshore production. The shares of R&D by foreign affiliates in these counterfactual equilibria are reported in column 5 of Table 7, which shows small but universal increases in offshore R&D shares across countries.

This result might seem surprising at first glance, given the partial equilibrium intuition below:

eliminating the access to foreign consumers through direct exporting reduces the return to doing R&D in a host country. This effect is especially strong for more productive firms, because they export more. So fewer foreign firms enter, and their share in total R&D decreases. In a model with both trade and offshore production, however, this direct channel is muted—without exporting, firms can still serve foreign consumers by offshoring their production to other countries. Moreover, due to the lower demand for labor from production, wages for both inventors and production workers decrease, which makes the country more attractive as a host for R&D centers. An increase in export costs thus has a similar effect to a decrease in a host country's production efficiency, which strengthens its comparative advantage in innovation, driving it to specialize in R&D activities.

In the second experiment, I increase the costs of offshore production in each country to infinity (by setting $\phi_{il}^R, i \neq l$ to zero), so it is impossible for R&D centers to perform offshore production in other countries. The 6th column of Table 7 shows that, compared to the benchmark equilibrium, most countries experience a decrease in offshore R&D. The average share of R&D by foreign affiliates decreases by about 8 percentage points from the benchmark economy, to around 30%. Because firms located in emerging economies in this sample do not perform outward offshoring activities to begin with, the decrease in offshore R&D resulting from this change tends to be more significant for developed economies than for emerging economies.

The general equilibrium effect works in the same direction as the partial equilibrium effect in this case. 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 for production workers and inventors, making the country less attractive as a host for R&D centers. An increase in offshore production costs is therefore similar to a reduction in R&D innovation efficiency of a country, which strengthens its comparative advantage in production.

Column 7 of Table 7 reports the experiment when both exporting and offshore production are shut down. Compared to column 5, the share of offshore R&D is much smaller for developed countries, because when offshore production is not an option, countries can no longer specialize in innovation. For developing countries, the differences between columns 5 and 7 are small, mainly because they do not perform much outward offshoring production in the benchmark equilibrium.

Finally, I combine the two sets of experiments reported in this section, by changing the two distributions and also eliminating host access to foreign consumers and producers. The average of foreign R&D shares, reported in the last column of Table 7, is 3.3%, or one-tenth of the benchmark value. The only countries that attract a significant share of offshore R&D are large emerging economies, such as Brazil, China, and India. The large markets of these countries, and their relative closeness to export and offshore production from other countries, are the reasons for foreign firms to perform R&D in those countries.

In summary, the experiments in this section show that the two main forces incorporated in the model have significant impacts on firms' offshore R&D decisions. Differences in the management and worker quality distributions together explain about six-seventh of the equilibrium offshore

Figure 9: The Welfare Gains from Offshore R&D
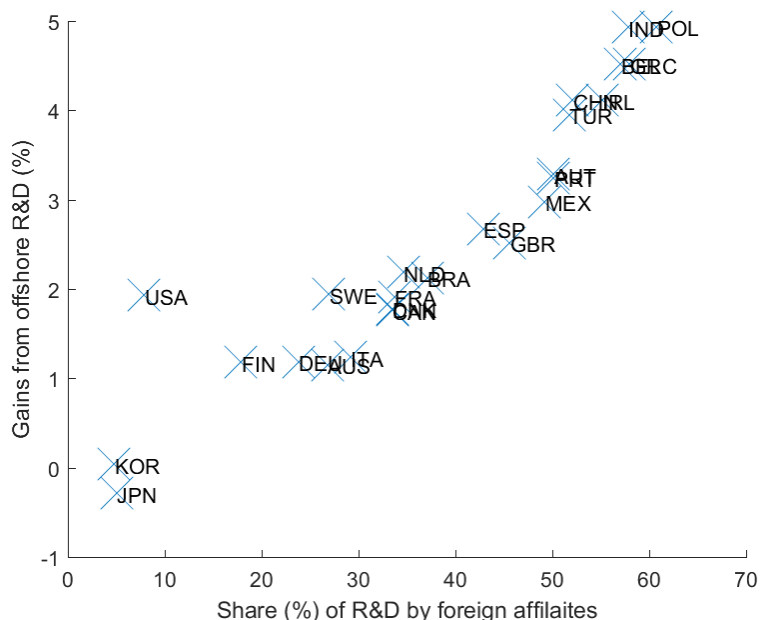
Notes: The vertical axis shows the welfare gains from offshore R&D (%), and the horizontal axis shows the share of R&D in a host country performed by foreign affiliates in the model economy.

R&D. Host access to foreign customers reduces offshore R&D in the country, while its access to foreign producers increases it. Combined, international differences in the distributions of talent and firm efficiency, and access to foreign markets and producers, explain more than 92% of the average level of offshore R&D in the benchmark equilibrium. The small remaining offshore R&D activities are concentrated in emerging economies with large domestic markets.

6.2   The Gains from Offshore R&D

I now turn to the normative aspect of offshore R&D. As a starting point, I examine the welfare gains from various forms of economic integrations by eliminating each channel from the model separately. I define the gains from offshore R&D as the increase in real income as a country moves from an equilibrium where offshore R&D is not allowed to the baseline equilibrium. I define gains from trade and gains from offshore production analogously. I compare these gains from individual channels to the overall gains from openness.

The first column in Table 8 presents the welfare gains from offshore R&D. The unweighted average welfare gain is 2.6%. This average, however, masks a great deal of country heterogeneity. Some countries, such as China, India, and Greece, benefit by around 4% or higher. Meanwhile countries like Japan and Korea barely receive any benefits or even lose, due to the general equilibrium effect from international competition. Figure 9 plots the gains from offshore R&D against the share of foreign affiliates in domestic R&D for each country. Countries with a higher share of R&D done by foreigners tend to benefit more from offshore R&D.

45

Table 8: The Welfare Gains from International Economic Integration

ISO	Offshore R&D 1	Trade 2	Offshore Prod. 3	Openness 4	Trade & Offshore Prod. 5
Developed					
AUS	1.14	5.64	2.09	19.47	17.24
AUT	3.28	10.67	3.12	52.43	43.98
BEL	4.51	15.94	4.52	74.15	61.53
CAN	1.76	9.43	3.53	43.56	38.74
DEU	1.18	7.08	3.17	24.87	22.88
DNK	1.77	17.23	2.32	43.02	40.40
ESP	2.67	3.95	0.32	17.61	12.69
FIN	1.17	8.12	3.00	21.70	20.86
FRA	1.91	6.90	2.34	22.51	19.62
GBR	2.51	8.47	4.59	34.68	28.82
GRC	4.51	7.67	0.03	20.54	13.32
IRL	4.11	12.55	2.17	60.98	48.39
ITA	1.23	3.33	0.42	11.35	9.22
JPN	-0.29	3.16	1.31	4.75	4.69
KOR	0.03	3.83	0.15	6.60	6.45
NLD	2.18	19.47	9.43	64.29	62.48
POL	4.93	5.01	0.61	28.20	17.45
PRT	3.24	5.23	2.48	43.77	33.17
SWE	1.93	8.79	3.52	34.33	31.72
USA	1.93	8.92	3.97	16.08	15.90
Average	2.29	8.57	2.65	32.24	27.48
Emerging					
BRA	2.12	-0.46	0.09	6.91	3.27
CHN	4.11	-0.92	-0.05	7.96	2.40
IND	4.93	-0.53	0.04	9.48	3.70
MEX	2.97	3.61	0.73	20.73	13.60
TUR	3.95	3.27	-0.18	12.06	6.14
Average	3.61	1	0.13	11.43	6.41
Average	2.55	7.06	2.15	28.08	23.27

The second and third columns report the gains from trade and the gains from offshore production, respectively. The average is around 7.0% for trade, and 2.1% for offshore production. Again, the welfare gains take a wide range of values. As expected, smaller economies and countries that are closer to major markets, such as Belgium, Netherlands, and Ireland, gain more from both trade and offshore production. Larger and more remote economies, such as India, gain less. Some countries even receive modest losses from trade and offshore production.

In the fourth column are the overall gains from openness. They range from 16% for the U.S. to 74% for Belgium, with an average of 35%. The gains from openness are almost always larger than the sum of the gains from the three forms of economic integration, which means these three forms of integration are substitutes—the benefit from additional openness is smaller once a country is already open in other dimensions. The substitution between trade and offshore production is intuitive—since these two are alternative ways of serving goods from where they are invented to where they are consumed, when one channel is present, the marginal benefits from the other channel are lower. The last column of Table 8 reports the combined gains from trade and offshore production, computed from a counterfactual scenario where both trade and offshore production are eliminated. Indeed, the values in column 5 are universally larger than the sum of columns 2 and 3. This result is consistent with the finding in Arkolakis et al. (2014), in a setting without offshore R&D.

The interaction pattern between offshore R&D and the combined effect of trade and offshore production is more nuanced—while in most countries, the sum of columns 1 and 5 is still smaller than column 4, the difference is small. In countries like the U.S., the sum of gains from offshore R&D and the gains from trade and offshore production is actually larger than the gains from openness.

This difference is again related to the interaction among various forces through country specialization. First, there is a demand-for-R&D channel. The option to export and to produce offshore raises the return to innovation. Because of the fixed marketing cost, this benefits more efficient firms particularly, who are also the ones most likely to perform offshore R&D. This demand side channel therefore tends to increase the gains from offshore R&D when trade and offshore production are present. However, there is also a labor-supply channel. Since innovation and production compete for workers, the general equilibrium effect discussed in the previous section sets in. When one sector expands in a country, wages increase, making the country benefit less from new opportunities in the other sector.

Offshore R&D tends to increase the R&D efficiency in countries that are relatively scarce in high efficiency firms, which weakens the comparative advantage of these countries in production. Because those are also the countries that tend to specialize in production, offshore R&D reduces their gains from trade and offshore production by weakening their comparative advantage. The substitution between offshore R&D, trade and offshore production is stronger for countries with strong comparative advantage in production. We can use the ratio between the sum of gains from offshore R&D and the combined gains from trade and offshore production over the gains from

Figure 10: The Substitution Between International Activities
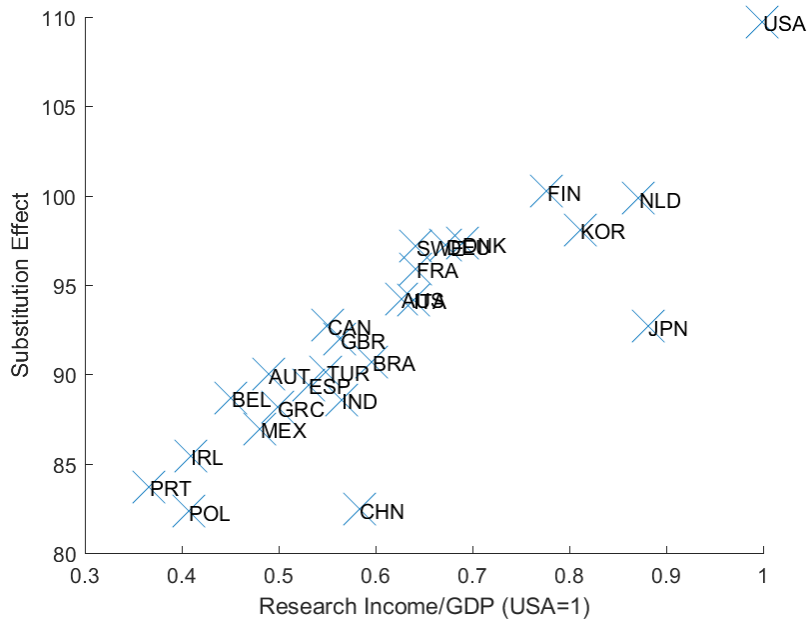
Notes: The vertical axis shows the sum of gains from offshore R&D and the combined gains from trade and offshore production, divided by the overall gains from openness. The horizontal axis shows the share researcher compensation in total income.

openness as a measure of the strength of this substitution. A lower ratio means smaller marginal gains from further integrating the economy once it is already integrated through other ways, and therefore represents stronger substitution. I use the share of income generated by R&D labor in the calibrated equilibrium as a measure for comparative advantage in innovation. Figure 10 displays the relationship between these two measures. As conjectured, the substitution is more important, for countries with comparative advantage in production.

How important is accounting for offshore R&D in understanding the gains from openness? Figure 11 plots the relationship between host income and the ratio between the gains from openness in the benchmark model, shown in Table 8, and the gains from openness in a restricted-version of the model without offshore R&D. With offshore R&D as an additional channel for gains from openness, the ratio is generally larger than 1, indicating higher gains from openness in the benchmark model. The average of this ratio among the model countries is 1.2. This amplification, however differs significantly across countries. For emerging countries in the sample, such as China, India, Brazil, and Turkey, the gains from openness are more than 100% higher in the benchmark model with offshore R&D. This amplification is much lower for developed countries. For example, for the U.S., the inclusion of offshore R&D only increases the gains from openness by 15%. The wide range of the ratio also underscores the importance of incorporating offshore R&D—overlooking this channel will not only understate the gains from openness, but also bias the comparison of the gains from openness across countries.

Why do developing countries benefit more from offshore R&D? Further examination of coun-

Figure 11: Relative Importance of Offshore R&D
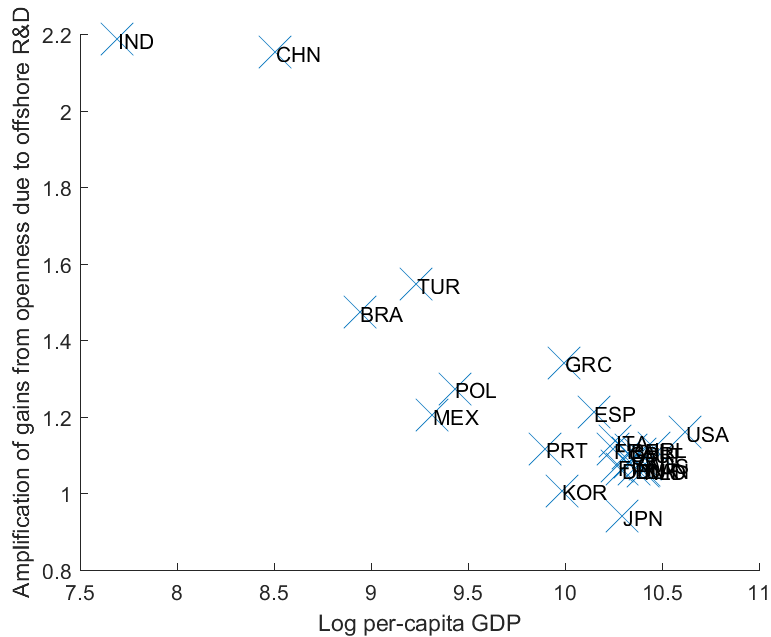
Notes: The vertical axis shows the ratio between gains from openness in the benchmark model, and the gains from openness in a model without offshore R&D. The horizontal axis shows host income.

tries' participation in various forms of integration suggests that during the sample period, developing countries participated more intensively in offshore R&D than in trade and offshore production. By fitting this pattern, the model implies that the frictions impeding offshore R&D increase more slowly with distance than the frictions impeding trade and offshore production. As a result, developing countries which are far away from major home countries of innovating firms—U.S., West Europe, and Japan—participate more intensively in offshore R&D, and less intensively in offshore production.

To sum up, the counterfactual experiments in this section demonstrate that offshore R&D represents a quantitatively important new channel through which countries benefit from globalization. It is a weak substitute for trade and offshore production in general, although the substitution patterns depend on a country's specialization in innovation or production in the world economy. Further, by showing that offshore R&D and other forms of globalization have very different impacts across countries, the results also highlight the importance of modelling offshore R&D separately, rather than treating it as part of the offshore production process.

6.3 Further Liberalization of China and India

Existing quantitative research on multinational activities usually does not allow firms to make independent decisions on offshore R&D and production. I evaluate whether this is an important restriction by comparing the welfare implications of liberalizing offshore R&D and offshore production. Doing so is potentially important because policy makers usually have at their disposal

policies that specifically target production or innovation activities.

As an example, I focus on the case of China and India and evaluate two types of openness policies. This exercise is interesting in its own right, because both countries are becoming popular destinations for offshore production and R&D. Related to this trend, their governments are attempting to attract more foreign companies, especially R&D intensive ones, by cutting red tape and speeding up the entry approval process.

I first consider an inward offshore R&D liberalization that makes it easier for foreign firms to open R&D centers in India and China. More specifically, I reduce the fixed costs of R&D in these countries by 20%. This reduction in cost can be interpreted as a tax credit for the upfront investment in R&D, subsidized land, or speedy approval of entry. The magnitude of the reduction is well within the range of policies commonly used.[49] The first column of Table 9 reports the results. As we can see, China and India benefit by 0.6 and 0.9 percent in welfare from such a policy, while other countries are not significantly affected.

The second experiment is a liberalization in inward offshore production, which increases ϕ_{oi}^{P} by 10% for $i =$ India, China, $o \neq i$. Because these two types of liberalizations do not necessarily share the same fiscal costs or administrative burdens, I do not compare the levels of the welfare gains, but instead focus on the distributions of the welfare gains across countries. The second column of Table 9 shows that India and China still benefit from this liberalization. But differently from the first experiment, major developed countries also benefit significantly. The difference between these two experiments is due to the interaction between offshore R&D and countries' specialization in the world economy. Because offshore R&D into China and India reduces these two countries' comparative advantage in production, it pushes developed countries to be less specialized in innovation. As a result, they do not benefit much in the first experiment. Inward offshore production liberalization, on the other hand, allows China and India to be more specialized in production, and developed countries to be more specialized in innovation, thus benefiting everyone. These two experiments demonstrate that openness to offshore R&D and offshore production could have different welfare implications for other countries. It is thus very important to separate offshore R&D and offshore production in the model, to better evaluate specific policies.

Because of the interactions among the three forms of global integration, incorporating offshore R&D also affects our understanding of the effects of other types of policies. I focus on China and India as an example to illustrate this point. Specifically, I consider the same liberalization for China and India in inward offshore production 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 9. Compared to the second column, the welfare gains are significantly smaller for China and India, but larger for developed countries. The reason for the difference is that, when there is no offshore R&D, openness to offshore production only crowds out R&D by domestic firms, so Chinese and Indian firm owners bear all the reduction in profit from increased

---

[49]For example, in 2012, to attract a $30 billion investment in a chip factory from Samsung Electronics Co. Ltd., the Chinese city Xi'An offered an package of favorable policies, including free land, infrastructure, and tax credits. The land alone was valued at $4 billion, more than 10% of the initial investment cost.

Table 9: Further Inward FDI Liberalization in China and India

	Policy 1	Policy 2	Policy 3
CHN	0.65	0.20	0.13
IND	0.89	0.19	0.12
DEU	0.00	0.17	0.21
JPN	0.00	0.15	0.24
USA	-0.01	0.09	0.26

Notes: All numbers are in percentage point terms. Policy 1 is a unilateral reduction of 20% in *fixed* inward offshore R&D costs for China and India from the benchmark equilibrium. Policy 2 is a unilateral reduction of 10% in inward offshore production costs for China and India. Policy 3 simulates the same shock as in Policy 2 in a restricted version of the model without offshore R&D.

inward offshore production. This reduces the aggregate welfare gains in these two countries, but increases the welfare gains to developed countries. The difference in welfare impacts suggests that even if one's goal is solely to understand the effect of liberalizing offshore production, it is important to incorporate offshore R&D into the model.

7 Conclusion

Talented researchers and efficient firms are both necessary inputs to invention of new products, but they are distributed unevenly across countries. By carrying out their R&D activities offshore, firms mobilize their management technology across borders, which might generate important aggregate gains. This paper develops a unified model of firms' global R&D and production decisions, featuring talent-acquisition and market-access motives for offshore R&D. These motives generate testable implications on the pattern of offshore R&D by firms with different efficiencies and in host countries with different characteristics. Empirical evidence based on USPTO patenting data by firms performing R&D in different countries confirms these predictions. Counterfactuals using the calibrated model further suggest these two motives can account for over 90% of the observed offshore R&D.

Quantitatively, the welfare gains from offshore R&D are on average 2.5% of real income. Incorporating this channel amplifies the welfare gains from openness by a factor of 1.2, with more amplification for developing countries than for developed countries. Further experiments show that a country's openness to offshore R&D and offshore production have very different spillover effects to other countries. Moreover, because of the interaction among various forms of international integration, whether offshore R&D decisions are allowed makes a difference when evaluating

the effects of liberalizing offshore production. All these results point to the importance of incorporating offshore R&D for a better understanding of globalization.

As a first step towards quantitatively evaluating offshore R&D, this paper abstracts from many potential channels through which it can affect aggregate welfare, such as the spillover between foreign and domestic R&D centers, and the effects of offshore R&D on growth. Moreover, although offshore R&D is one way for efficient firms and talented researchers from different countries to work together, an alternative is for researchers to migrate, which this paper abstracts from. All these are exciting opportunities for future research.

References

Abowd, John M, Francis Kramarz, and David N Margolis, "High Wage Workers and High Wage Firms," *Econometrica*, 1999, *67* (2), 251–333.

Acemoglu, Daron and Joshua Linn, "Market Size in Innovation: Theory and Evidence from the Pharmaceutical Industry," *The Quarterly journal of economics*, 2004, *119* (3), 1049–1090.

Akcigit, Ufuk, Salomé Baslandze, and Stefanie Stantcheva, "Taxation and the International Mobility of Inventors," *NBER Working Paper*, 2015.

Alfaro, Laura, Sebnem Kalemli-Ozcan, and Vadym Volosovych, "Why Doesn't Capital Flow from Rich to Poor Countries? An Empirical Investigation," *The Review of Economics and Statistics*, 2008, *90* (2), 347–368.

Alviarez, Vanessa, "Multinational Production and Comparative Advantage," 2016.

Ambos, Björn, "Foreign Direct Investment in Industrial Research and Development: A Study of German MNCs," *Research Policy*, 2005, *34* (4), 395–410.

Antràs, Pol and Stephen Yeaple, "Multinational Firms and the Structure of International Trade," *Handbook of International Economics*, 2014, *4*, 55.

__ , **Luis Garicano, and Esteban Rossi-Hansberg**, "Offshoring in a Knowledge Economy," *Quarterly Journal of Economics*, 2006, *121* (1).

Arkolakis, Costas, Natalia Ramondo, Andrés Rodríguez-Clare, and Stephen Yeaple, "Innovation and Production in the Global Economy," *National Bureau of Economic Research WP*, 2014.

Bilir, L Kamran and Eduardo Morales, "The Impact of Innovation in the Multinational Firm," 2016.

Bloom, Nicholas, Benn Eifert, Aprajit Mahajan, David McKenzie, and John Roberts, "Does Management Matter? Evidence from India," *The Quarterly Journal of Economics*, 2013, *1* (51), 51.

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

__ , **Erik Brynjolfsson, Lucia Foster, Ron S Jarmin, Megha Patnaik, Itay Saporta Eksten, and John Van Reenen**, "IT and Management in America," *CEPR Discussion Paper No. DP9886*, 2014.

Branstetter, Li, Guangwei Li, and Francisco Veloso, "The Globalization of R&D: China, India and the Rise of International Co-invention," *Unpublished Manuscript*, 2013.

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.

Costinot, Arnaud and Andres Rodriguez-Clare, "Trade Theory with Numbers: Quantifying the Consequences of Globalization," *Handbook of International Economics*, 2014, *4*, 197.

_ **and Jonathan Vogel**, "Matching and Inequality in the World Economy," *Journal of Political Economy*, 2010, *118* (4).

_ **, _ et al.**, "Beyond Ricardo: Assignment Models in International Trade," *Annual Review of Economics*, 2015, *7* (1), 31–62.

Gersbach, Hans and Armin Schmutzler, "Foreign Direct Investment and R&D-Offshoring," *Oxford Economic Papers*, 2011, *63* (1), 134–157.

Grossman, Gene M and Elhanan Helpman, "Growth, Trade, and Inequality," *National Bureau of Economic Research Working Paper*, 2014.

_ **and Giovanni Maggi**, "Diversity and Trade," *American Economic Review*, 2000, pp. 1255–1275.

_ **, Elhanan Helpman, Philipp Kircher et al.**, "Matching, Sorting, and the Distributional Effects of International Trade," *Journal of Political Economy*, 2015.

Guadalupe, Maria, Olga Kuzmina, and Catherine Thomas, "Innovation and Foreign Ownership," *The American Economic Review*, 2012, *102* (7), 3594–3627.

Guimón, José, "Government Strategies to Attract R&D-Intensive FDI," *The Journal of Technology Transfer*, 2009, *34* (4), 364–379.

Haas, Martine R and Prithwiraj Raj Choudhury, "Collaborate to Innovate? Teams and Patent Generation in a Global R&D Center," 2015.

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

_ **, Adam B Jaffe, and Manuel Trajtenberg**, "The NBER Patent Citation Data File: Lessons, Insights and Methodological Tools," *National Bureau of Economic Research*, 2001.

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, "FDI as an Outcome of the Market for Corporate Control: Theory and Evidence," *Journal of International Economics*, 2008, *74* (1), 2–20.

_ **and Thierry Mayer**, "Market Potential and the Location of Japanese investment in the European Union," *Review of Economics and Statistics*, 2004, *86* (4), 959–972.

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.

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, p. rdv008.

Hsieh, Chang-Tai and Peter J Klenow, "Misallocation and Manufacturing TFP in China and India," *The Quarterly journal of economics*, 2009, *124* (4), 1403–1448.

Irarrazabal, Alfonso, Andreas Moxnes, and Luca David Opromolla, "The Margins of Multina-

tional Production and the Role of Intrafirm Trade," *Journal of Political Economy*, 2013, *121* (1), 74–126.

Ito, Banri and Ryuhei Wakasugi, "What Factors Determine the Mode of Overseas R&D by Multinationals? Empirical Evidence," *Research Policy*, 2007, *36* (8), 1275–1287.

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

Kerr, Sari Pekkala and William R Kerr, "Global Collaborative Patents," *NBER Working Paper*, 2014.

_ , _ , **Caglar Ozden, and Christopher Robert Parsons**, "Global Talent Flows," *World Bank Policy Research Working Paper*, 2016, (7852).

Li, Guan-Cheng, Ronald Lai, Alexander DiAmour, David M Doolin, Ye Sun, Vetle I Torvik, Z Yu Amy, and Lee Fleming, "Disambiguation and Coauthorship Networks of the US Patent Inventor Database (1975–2010)," *Research Policy*, 2014, *43* (6), 941–955.

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

Mayer, Thierry and Soledad Zignago, "Notes on CEPII's Distances Measures: The GeoDist Database," *CEPII Working Paper*, 2011.

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

Nocke, Volker and Stephen Yeaple, "Cross-Border Mergers and Acquisitions vs. Greenfield Foreign Direct Investment: The Role of Firm Heterogeneity," *Journal of International Economics*, 2007, *72* (2), 336–365.

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

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

Pavitt, Keith, "Uses and Abuses of Patent Statistics," in A. F. J. van Raan, ed., *Handbook of Quantitative Studies of Science and Technology*, Sole Distributors for the U.S.A. And Canada, Elsevier Science Pub. Co., 1988, pp. 509–536.

Ramondo, Natalia, "A Quantitative Approach to Multinational Production," *Journal of International Economics*, 2014, *93* (1), 108–122.

_ **and Andre Rodrıguez-Clare**, "Trade, Multinational Production, and the Gains from Openness," *Journal of Political Economy*, 2013, *121* (2).

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

Roys, Nicolas and Ananth Seshadri, "Economic Development and the Organization of Production," *Working Paper*, 2014.

Sampson, Thomas, "Selection into Trade and Wage Inequality," *American Economic Journal: Microeconomics*, 2014, *6* (3), 157–202.

Schank, Thorsten, Claus Schnabel, and Joachim Wagner, "Do Exporters Really Pay Higher Wa-

ges? First Evidence from German Linked Employer–Employee Data," *Journal of international Economics*, 2007, *72* (1), 52–74.

Shimizutani, Satoshi and Yasuyuki Todo, "What Determines Overseas R&D Activities? The Case of Japanese Multinational Firms," *Research Policy*, 2008, *37* (3), 530–544.

Simonovska, Ina and Michael E Waugh, "The Elasticity of Trade: Estimates and Evidence," *Journal of international Economics*, 2014, *92* (1), 34–50.

Terviö, Marko, "The Difference that CEOs Make: An Assignment Model Approach," *The American Economic Review*, 2008, pp. 642–668.

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," *Quarterly Journal of Economics*, forthcoming.

Wei, Shang-Jin, "How Taxing is Corruption on International Investors?," *Review of economics and statistics*, 2000, *82* (1), 1–11.

Yeaple, Stephen Ross, "The Role of Skill Endowments in the Structure of US Outward Foreign Direct Investment," *Review of Economics and Statistics*, 2003, *85* (3), 726–734.

_ , "A Simple Model of Firm Heterogeneity, International Trade, and Wages," *Journal of International Economics*, 2005, *65* (1), 1–20.

Appendix A   Theory

A.1   Lemma 1

Proof   Country index i is omitted in this proof. Consider two R&D centers characterized by management scores (z_1^P, z_1^R), and (z_2^P, z_2^R), with $z_2^R > z_1^R$. Let $T : \mathbb{Z}^{\mathbb{P}} \times \mathbb{Z}^{\mathbb{R}} \to \Theta$ be the mapping from the type of an R&D center to the type(s) of researchers it recruits. Let $\theta_1 = T(z_1^P, z_1^R)$, $\theta_2 = T(z_2^R, z_2^P)$, so the first R&D center recruits θ_1 and the second θ_2.

I prove by contradiction that $\theta_1 <= \theta_2$. Suppose $\theta_1 > \theta_2$, given that (z_1^P, z_1^R) hires θ_1, it must be the case that it at least weakly prefer researchers with talent θ_1 to researchers with ability θ_2. From Equation 5, this implies

$$\pi(z_1^P)^{\frac{1}{1-\gamma}} w(\theta_2)^{-\frac{\gamma}{1-\gamma}} f(z_1^R, \theta_2)^{\frac{1}{1-\gamma}} \leq \pi_i(z_1^P)^{\frac{1}{1-\gamma}} w(\theta_1)^{-\frac{\gamma}{1-\gamma}} f(z_1^R, \theta_1)^{\frac{1}{1-\gamma}}$$

$$[\frac{w(\theta_1)}{w(\theta_2)}]^{\frac{\gamma}{1-\gamma}} \leq [\frac{f(z_1^R, \theta_1)}{f(z_1^R, \theta_2)}]^{\frac{1}{1-\gamma}}.$$

Similarly, for (z_2^P, z_2^R):

$$[\frac{w(\theta_2)}{w(\theta_1)}]^{\frac{\gamma}{1-\gamma}} \leq [\frac{f(z_2^R, \theta_2)}{f(z_2^R, \theta_1)}]^{\frac{1}{1-\gamma}}. \tag{A.1}$$

We therefore have:

$$[\frac{w(\theta_1)}{w(\theta_2)}]^{\frac{\gamma}{1-\gamma}} \geq [\frac{f(z_2^R,\theta_1)}{f(z_2^R,\theta_2)}]^{\frac{1}{1-\gamma}} > [\frac{f(z_1^R,\theta_1)}{f(z_1^R,\theta_2)}]^{\frac{1}{1-\gamma}} \geq [\frac{w(\theta_1)}{w(\theta_2)}]^{\frac{\gamma}{1-\gamma}},$$

where the first inequality is from Equation A.1, and the second from Assumption 1.

The above contradiction suggests that $\theta_1 \leq \theta_2$, and that $T(z^P, z^R)$ is weakly increasing in z^R. Now suppose $\theta_1 = \theta_2$, given the weak monotonicity of T, for R&D centers with $z^R \in (z_1^R, z_2^R)$, regardless of their production efficiency, will also recruit θ_1. Therefore in equilibrium, the demand for researchers with ability θ_1 will have a mass point, which contradicts with the assumption that talent distribution in each country has no mass point. Therefore the equilibrium matching function, $T(z^P, z^R)$, will be strictly increasing in z^R.

Now consider (z_1^P, z_1^R) and (z_2^P, z_1^R). If these two R&D centers hire different types of researchers, θ_1, and θ_2, then from monotonicity, all researchers with ability between θ_1 and θ_2 would be recruited by R&D centers with innovation efficiency z_1^R. The demand for researchers by R&D centers with this efficiency will be a positive mass, which contradicts that the distribution of efficiency for R&D centers have no mass point. Therefore $T_i(z^P, z^R)$ is independent of z^P.

A.2 Lemma 2

Proof   Country index i is omitted. To show that $w(\theta)$ is differentiable, we consider an R&D center with innovation efficiency z_1^R, which is matched to $T(z_1^R)$. Consider z^R and $\theta = T(z_1^R)$. By the definition of $T(z_1^R)$, R&D centers with innovation efficiency z_1^R prefers researchers with ability θ instead of those with $\theta + d\theta$. Following Equation A.1, this implies:

$$[\frac{w(\theta + d\theta)}{w(\theta)}]^{\gamma} \geq [\frac{f(z_1^R, \theta + d\theta)}{f(z_1^R, \theta)}]$$

Consider $z_2^R = T^{-1}(\theta + d\theta)$, then similarly, we also have:

$$[\frac{w(\theta + d\theta)}{w(\theta)}]^{\gamma} \leq [\frac{f(z_2^R, \theta + d\theta)}{f(z_2^R, \theta)}]$$

From these two equations, we have:

$$\frac{f(z_1^R, \theta + d\theta) - f(z_1^R, \theta)}{f(z_1^R, \theta)} \leq \frac{w(\theta + d\theta)^{\gamma} - w(\theta)^{\gamma}}{w(\theta)^{\gamma}} \leq \frac{f(z_2^R, \theta + d\theta) - f(z_2^R, \theta)}{f(z_2^R, \theta)}$$

Dividing all three terms in the above inequality by $d\theta$, and letting $d\theta \to 0$, the first term approaches $\frac{f_2(z_1^R,\theta)}{f(z_1^R,\theta)}$, and the third term approaches $\frac{f_2(z_2^R,\theta)}{f(z_2^R,\theta)}$, which in term equals $\frac{f_2(z_1^R,\theta)}{f(z_1^R,\theta)}$, given the continuity of $f(z^R, \theta)$ and $T(z^R)$. Therefore we have

$$\lim_{d\theta \to 0} \frac{w(\theta + d\theta)^{\gamma} - w(\theta)^{\gamma}}{w(\theta)^{\gamma} d\theta} = \frac{f_2(z_1^R, \theta)}{f(z_1^R, \theta)}$$

Therefore $w(\theta)^\gamma$ is differentiable, with derivative being $w(\theta)^\gamma \frac{f_2(z_1^R,\theta)}{f(z_1^R,\theta)}$. This implies $w(\theta)$ is also differentiable, and its derive satisfies the following equation:

$$\frac{w'(\theta)}{w(\theta)} = \frac{f_2(z^R,\theta)}{\gamma f(z^R,\theta)}.$$

A.3 Proposition 2

Proof Country index i is omitted. Consider two talent distributions $H(\theta)$ and $\tilde{H}(\theta)$, with $\tilde{H}(\theta)$ more talent abundant than $H(\theta)$ according to Definition 1, and $h(\tilde{\theta})$ and $h(\theta)$ being the corresponding PDFs. I use tilde to denote variable under $\tilde{H}(\theta)$. I first show that $\tilde{T}(z^R) \geq T(z^R)$, i.e., firms are matched with more talented researchers under $\tilde{H}(\theta)$ than under $H(\theta)$.

I prove by contradiction. From the definition of talent abundance, $\tilde{T}(\underline{z}^R) = \underline{\tilde{\theta}} > T(\underline{z}^R) = \underline{\theta}$, and $\tilde{T}(\overline{z}^R) = \overline{\tilde{\theta}} > T(\overline{z}^R) = \overline{\theta}$. Suppose for $z^R \in (\underline{z}^R, \overline{z}^R)$, $\tilde{T}(z^R) < T(z^R)$, then there must be $z_1^R < z^R$ and $z_2^R > z^R$, so that $\tilde{T}$ crosses T from above at z_1^R, and crosses it again from below at z_2^R. In $z^R \in (z_1^R, z_2^R)$, $\tilde{T}(z^R) < T(z^R)$.

For this to be possible, it must be the case that $\frac{\tilde{T}'(z_2^R)}{\tilde{T}'(z_1^R)} > \frac{T'(z_2^R)}{T'(z_1^R)}$. Using Equation 7, and that $T(z_1^R) = \tilde{T}(z_1^R)$, $T(z_2^R) = \tilde{T}(z_2^R)$, this implies:

$$[\frac{\tilde{w}(\tilde{T}(z_1^R))}{\tilde{w}(\tilde{T}_i(z_2^R))}]^{\frac{1}{1-\gamma}} \frac{\int_{\mathbf{Z}^P} \tilde{\pi}(z^P)^{\frac{1}{1-\gamma}} g^P(z^P|z_2^R) dz^P}{\int_{\mathbf{Z}^P} \tilde{\pi}(z^P)^{\frac{1}{1-\gamma}} g^P(z^P|z_1^R) dz^P} \frac{\tilde{h}(\tilde{T}(z_1^R))}{\tilde{h}(\tilde{T}(z_2^R))} > [\frac{w(T(z_1^R))}{w(T(z_2^R))}]^{\frac{1}{1-\gamma}} \frac{\int_{\mathbf{Z}^P} \pi(z^P)^{\frac{1}{1-\gamma}} g^P(z^P|z_2^R) dz^P}{\int_{\mathbf{Z}^P} \pi(z^P)^{\frac{1}{1-\gamma}} g^P(z^P|z_1^R) dz^P} \frac{h(T(z_1^R))}{h(T(z_2^R))}.$$

Note that $\pi(z^P)$ depend on the talent distribution because the latter determines general equilibrium outcomes, such as X and P. However, if one of the two additional conditions stated in Proposition 2 is satisfied,

$$\frac{\int_{\mathbf{Z}^P} \tilde{\pi}(z^P)^{\frac{1}{1-\gamma}} g^P(z^P|z_2^R) dz^P}{\int_{\mathbf{Z}^P} \tilde{\pi}(z^P)^{\frac{1}{1-\gamma}} g^P(z^P|z_1^R) dz^P} = \frac{\int_{\mathbf{Z}^P} \pi(z^P)^{\frac{1}{1-\gamma}} g^P(z^P|z_2^R) dz^P}{\int_{\mathbf{Z}^P} \pi(z^P)^{\frac{1}{1-\gamma}} g^P(z^P|z_1^R) dz^P}.$$

Then the above inequality further simplifies to:

$$[\frac{\tilde{w}(\tilde{T}(z_1^R))}{\tilde{w}(\tilde{T}(z_2^R))}]^{\frac{1}{1-\gamma}} \frac{\tilde{h}(\tilde{T}(z_1^R))}{\tilde{h}(\tilde{T}(z_2^R))} > [\frac{w(T(z_1^R))}{w(T(z_2^R))}]^{\frac{1}{1-\gamma}} \frac{h(T(z_1^R))}{h(T(z_2^R))}.$$

From the definition of talent abundance, $\frac{\tilde{h}(\tilde{T}(z_1^R))}{\tilde{h}(\tilde{T}(z_2^R))} < \frac{h(T(z_1^R))}{h(T(z_2^R))}$. From Equation 6 and log-supermodularity, $[\frac{\tilde{w}(\tilde{T}(z_1^R))}{\tilde{w}(\tilde{T}(z_2^R))}]^{\frac{1}{1-\gamma}} < [\frac{w(T(z_1^R))}{w(T(z_2^R))}]^{\frac{1}{1-\gamma}}$, so the above inequality cannot hold. Thus we have proved that $\tilde{T} \geq T, \forall z^R \in (\underline{z}^R, \overline{z}^R)$.

Let $y(z^P, z^R)$ and $\tilde{y}(z^P, z^R)$ denote the number of varieties an R&D center with efficiency (z^P, z^R) develops when the talent distribution is $H(\theta)$ and $\tilde{H}(\theta)$, respectively. Now consider the output

difference between R&D centers with $z_1^R < z_2^R$. From Equation 4, we have:

$$\log(\frac{y(z_P, z_1^R)}{y(z^P, z_2^R)}) = \log(y(z^P, z_1^R)) - \log(y(z^P, z_2^R))$$

$$= \int_{z_1^R}^{z_2^R} \frac{\partial \log(y(z^P, z^R))}{\partial z^R} dz^R$$

$$= \frac{1}{1-\gamma} \int_{z_1^R}^{z_2^R} \frac{f_1(z^R, T(z^R))}{f(z^R, T(z^R))} dz^R$$

$$\leq \frac{1}{1-\gamma} \int_{z_1^R}^{z_2^R} \frac{f_1(z^R, \tilde{T}(z^R))}{f(z^R, \tilde{T}(z^R))} dz^R$$

$$= \log(\frac{\tilde{y}(z_P, z_1^R)}{\tilde{y}(z^P, z_2^R)}),$$

where the inequality uses the definition of log-supermodularity and the above conclusion that $\tilde{T}(z^R) \geq T(z^R)$.

A.4 Proposition 3

Proof   To derive the gains from openness under Assumption 3, I proceed in three steps. The first step is to derive expression for production workers' real wage, $\frac{w_i^P}{P_i}$, in terms of measurable flows and total number of domestically invented varieties. The second step is to derive the relationship between production wage and total expenditure, $\frac{X_i}{w_i^P}$ in order to obtain $\frac{X_i}{P_i}$, the real income of a country. In the final step, I use $\frac{X_i}{P_i}$ to derive the gains from openness.

Step1: real wage for production worker I first derive real wage for production workers, $\frac{w_i^P}{P_i}$. The key step is to derive the total measure of varieties in each country. Under the assumption that $f(z^R, \theta) = z^R \theta^\beta$, Equation 6 becomes $\frac{w_i'(\theta)}{w_i(\theta)} = \frac{\beta}{\gamma} \frac{1}{\theta}$. Therefore the researcher wage schedule can be solved directly:

$$w_i(\theta) = \underline{w}_i \theta^{\frac{\beta}{\gamma}}$$

Under the assumption that the fixed marketing cost is 0, the per-variety variable profit given by Equation 2 becomes $\pi_i(z^P) = z^{P^{\sigma-1}} \sum_d \frac{1}{\sigma} (\frac{\sigma}{\sigma-1})^{1-\sigma} \Gamma(\frac{\delta+1-\sigma}{\delta}) P_d^{\sigma-1} X_d \Psi_{id}^{\frac{\sigma-1}{\delta}}$. The total innovation output by an R&D center with (z^P, z^R) is therefore:

$$y_i(z^P, z^R) = (\frac{\gamma}{\underline{w}_i})^{\frac{\gamma}{1-\gamma}} \pi_i(z^P)^{\frac{\gamma}{1-\gamma}} z^{R \frac{1}{1-\gamma}}$$

$$= (\frac{\gamma}{\underline{w}_i})^{\frac{\gamma}{1-\gamma}} [\sum_d \frac{1}{\sigma} (\frac{\sigma}{\sigma-1})^{1-\sigma} \Gamma(\frac{\delta+1-\sigma}{\delta}) P_d^{\sigma-1} X_d \Psi_{id}^{\frac{\sigma-1}{\delta}}]^{\frac{\gamma}{1-\gamma}} z^{P \frac{\gamma(\sigma-1)}{1-\gamma}} z^{R \frac{1}{1-\gamma}}$$

The measure of varieties invented in country i that can be produced with z^P is

$$m_i(z^P) = R_i \int_{\mathbb{Z}_i^R} y_i(z^P, z^R) g(z^P | z^R) g_i(z^R) dz^R$$

$$= R_i g(z^P) \int_{\mathbb{Z}_i^R} y_i(z^P, z^R) g_i(z^R) dz^R$$

$$= K_i z^{P \frac{\gamma(\sigma-1)}{1-\gamma} - \kappa_P - 1},$$

$$\text{Where } K_i = \underline{z}_i^{P \kappa_P} \kappa_P (\frac{\gamma}{\underline{w}_i})^{\frac{\gamma}{1-\gamma}} [\sum_d \frac{1}{\sigma}(\frac{\sigma}{\sigma-1})^{1-\sigma} \Gamma(\frac{\delta+1-\sigma}{\delta}) P_d^{\sigma-1} X_d \Psi_{id}^{\frac{\sigma-1}{\delta}}]^{\frac{\gamma}{1-\gamma}} R_i \int_{\mathbb{Z}_i^R} z^{R \frac{1}{1-\gamma}} g_i(z^R) dz^R$$

$$\tag{A.2}$$

The total measure of varieties developed in country i, M_i is

$$M_i = \int_{\underline{z}_i^P}^{\infty} m_i(z^P) dz^P = \frac{K_i}{\kappa^P - \frac{\gamma(\sigma-1)}{1-\gamma}} \underline{z}_i^{P \frac{\gamma(\sigma-1)}{1-\gamma} - \kappa_P} \tag{A.3}$$

From Equation A.2, the productivity distribution for varieties developed in country i follows a Pareto distribution, with minimum $\underline{z}_i^P$ and dispersion parameter $\frac{\gamma(\sigma-1)}{1-\gamma} - \kappa_P$. Given the measure of new varieties, M_i, the offshore production and trade block of the model corresponds to the model in Arkolakis et al. (2014) with exogenous entry. Following their notation, I define $\lambda_{id}^E = \frac{\sum_l X_{ild}}{X_d}$, $\lambda_{ld}^T = \frac{\sum_i X_{ild}}{X_d}$. Then λ_{id}^E denotes the share of consumption expenditure in country d that are spent on goods invented in country i, and λ_{ld}^T denotes the share of consumption in country d that are imported from country l.

Given that the total measure of varieties developed in country i is M_i, and that their productivity distribution is Pareto, the ideal price index in country d is given by:

$$P_d^{1-\sigma} = \Gamma(\frac{\delta+1-\sigma}{\delta})(\frac{\sigma}{\sigma-1})^{1-\sigma} \sum_i \Psi_{id}^{\frac{\sigma-1}{\delta}} M_i \frac{\kappa_P - \frac{\gamma(\sigma-1)}{1-\gamma}}{\underline{z}_i^{P \frac{\gamma(\sigma-1)}{1-\gamma} - \kappa_P}} \int_{\underline{z}_i^P}^{\infty} z^{P(\sigma-1) + \frac{\gamma(\sigma-1)}{1-\gamma} - \kappa_P - 1} dz^P$$

$$= \Gamma(\frac{\delta+1-\sigma}{\delta})(\frac{\sigma}{\sigma-1})^{1-\sigma} \sum_i \Psi_{id}^{\frac{\sigma-1}{\delta}} M_i \frac{\kappa_P - \frac{\gamma(\sigma-1)}{1-\gamma}}{\kappa_P - \frac{\sigma-1}{1-\gamma}} \underline{z}_i^{P(\sigma-1)}.$$

By definition of X_{id} from Equation 13, using the above expression for price, we have

$$\lambda_{id}^E = \frac{X_{id}}{X_d}$$

$$= \frac{\Psi_{id}^{\frac{\sigma-1}{\delta}} M_i \underline{z}_i^{P(\sigma-1)}}{\sum_i \Psi_{id}^{\frac{\sigma-1}{\delta}} M_i \underline{z}_i^{P(\sigma-1)}}$$

$$= \Gamma(\frac{\delta+1-\sigma}{\delta})(\frac{\sigma}{\sigma-1})^{1-\sigma} \frac{\kappa_P - \frac{\gamma(\sigma-1)}{1-\gamma}}{\kappa_P - \frac{\sigma-1}{1-\gamma}} \frac{\Psi_{id}^{\frac{\sigma-1}{\delta}} M_i \underline{z}_i^{P(\sigma-1)}}{P_d^{1-\sigma}}.$$

Therefore,

$$P_d^{1-\sigma} = \frac{\Psi_{dd}^{\frac{\sigma-1}{\delta}} M_d \underline{z}_d^{P(\sigma-1)}}{\lambda_{dd}^E} \Gamma(\frac{\delta+1-\sigma}{\delta})(\frac{\sigma}{\sigma-1})^{1-\sigma} \frac{\kappa_P - \frac{\gamma(\sigma-1)}{1-\gamma}}{\kappa_P - \frac{\sigma-1}{1-\gamma}}. \tag{A.4}$$

Note that $\lambda_{dd}^T = \sum_i \psi_{idd} \lambda_{id}^E$, where $\psi_{idd} = \frac{T_d(\frac{w_d^P \tau_{dd}}{\phi_{id}^P})^{-\delta}}{\Psi_{id}}$, so production wage satisfies:

$$w_d^{P\delta} = \frac{1}{\lambda_{dd}^T}(\sum_i \frac{T_d \phi_{id}^{P\delta} \lambda_{id}^E}{\Psi_{id}}) \tag{A.5}$$

To express $\sum_i \frac{T_d \phi_{id}^{P\delta} \lambda_{id}^E}{\Psi_{id}}$ in flow units, consider:

$$\frac{X_{idd}}{X_d} = \lambda_{id}^E \frac{T_d(\frac{w_d^P \tau_{dd}}{\phi_{id}^P})^{-\delta}}{\Psi_{id}}$$

$$\Leftrightarrow \frac{X_{idd}}{X_d} \Psi_{dd} w_d^{P\delta} = \lambda_{id}^E T_d (\frac{\tau_{dd}}{\phi_{id}^P})^{-\delta} \frac{\Psi_{dd}}{\Psi_{id}}$$

$$\Leftrightarrow \frac{\sum_i X_{idd}}{X_d} \Psi_{dd} w_d^{P\delta} = \sum_i \lambda_{id}^E T_d \phi_{id}^{P\delta} \frac{\Psi_{dd}}{\Psi_{id}}$$

$$\Leftrightarrow \frac{\sum_i X_{idd}}{X_d} \frac{T_d}{\psi_{ddd}} = \sum_i \lambda_{id}^E T_d \phi_{id}^{P\delta} \frac{\Psi_{dd}}{\Psi_{id}}$$

$$\Leftrightarrow T_d \frac{\sum_i X_{idd}}{X_d} \frac{\sum_l X_{dld}}{X_{ddd}} = \sum_i \lambda_{id}^E T_d \phi_{id}^{P\delta} \frac{\Psi_{dd}}{\Psi_{id}}$$

The real wage for production workers, $\frac{w_d^P}{P_d}$, is

$$\begin{aligned}
\frac{w_d^P}{P_d} &= z_d^P \Gamma(\frac{\delta + 1 - \sigma}{\delta})^{\frac{1}{\sigma-1}} (\frac{\sigma}{\sigma-1})(\frac{\kappa_P - \frac{\gamma(\sigma-1)}{1-\gamma}}{\kappa_P - \frac{\sigma-1}{1-\gamma}})^{\frac{1}{\sigma-1}} M_d^{\frac{1}{\sigma-1}} \lambda_{dd}^{T^{-\frac{1}{\delta}}} \lambda_{dd}^{E^{-\frac{1}{\sigma-1}}} (\sum_i \lambda_{id}^E T_d \phi_{id}^{P\delta} \frac{\Psi_{dd}}{\Psi_{id}})^{\frac{1}{\delta}} \\
&= z_d^P \Gamma(\frac{\delta + 1 - \sigma}{\delta})^{\frac{1}{\sigma-1}} (\frac{\sigma}{\sigma-1})(\frac{\kappa_P - \frac{\gamma(\sigma-1)}{1-\gamma}}{\kappa_P - \frac{\sigma-1}{1-\gamma}})^{\frac{1}{\sigma-1}} M_d^{\frac{1}{\sigma-1}} \lambda_{dd}^{T^{-\frac{1}{\delta}}} \lambda_{dd}^{E^{-\frac{1}{\sigma-1}}} (T_d \frac{\sum_i X_{idd}}{X_d} \frac{\sum_l X_{dld}}{X_{ddd}})^{\frac{1}{\delta}} \\
&= z_d^P \Gamma(\frac{\delta + 1 - \sigma}{\delta})^{\frac{1}{\sigma-1}} (\frac{\sigma}{\sigma-1})(\frac{\kappa_P - \frac{\gamma(\sigma-1)}{1-\gamma}}{\kappa_P - \frac{\sigma-1}{1-\gamma}})^{\frac{1}{\sigma-1}} M_d^{\frac{1}{\sigma-1}} \lambda_{dd}^{T^{-\frac{1}{\delta}}} \lambda_{dd}^{E^{-\frac{1}{\sigma-1}}} (T_d \lambda_{dd}^E \frac{\sum_i X_{idd}}{X_{ddd}})^{\frac{1}{\delta}} \\
&= z_d^P \Gamma(\frac{\delta + 1 - \sigma}{\delta})^{\frac{1}{\sigma-1}} (\frac{\sigma}{\sigma-1})(\frac{\kappa_P - \frac{\gamma(\sigma-1)}{1-\gamma}}{\kappa_P - \frac{\sigma-1}{1-\gamma}})^{\frac{1}{\sigma-1}} T_d^{\frac{1}{\delta}} M_d^{\frac{1}{\sigma-1}} \lambda_{dd}^{T^{-\frac{1}{\delta}}} \lambda_{dd}^{E^{\frac{1}{\delta} - \frac{1}{\sigma-1}}} (\frac{X_{ddd}}{\sum_i X_{idd}})^{-\frac{1}{\delta}}
\end{aligned} \tag{A.6}$$

Step 2: Relating consumption to wage: In the second step, I derive the ratio between production wage and expenditures, $\frac{X_d}{P_d}$. I start with the market clearing condition for production workers:

$$\begin{aligned}
w_d^P L_d^P &= \frac{\sigma-1}{\sigma} Y_d + \sum_o E_o c_d^R w_d^P (1 - G_o(\hat{z}_{od}^R)) \\
&= \frac{\sigma-1}{\sigma} Y_d + \sum_o E_o c_d^R w_d^P (\frac{\hat{z}_{od}^R}{\underline{z}_o^R})^{-\kappa_R}
\end{aligned} \tag{A.7}$$

The first term on the right hand side is total demand for production workers from production, while the second term on the right hand side is demand from the overhead of R&D centers. The second line uses the fact that $G_o^E(\tilde{z}^R)$ follows Pareto distribution. The goal here is to express the

second term in the right hand side in the form of aggregate flow variables.

Consider the expected profit for a firm with z^R to perform offshore R&D in country i:

$$\pi_i^R(z^R) = \int_{\mathbb{Z}_i^P} \pi_i^R(z^P, z^R) g(z^P | z^R) dz^P$$

$$= (1-\gamma) \int_{\mathbb{Z}_i^P} y_i(z^P, z^R) \pi_i^P(z^P) g(z^P) dz^P$$

$$= (1-\gamma)\kappa_P(\frac{\gamma}{\underline{w}_i})^{\frac{\gamma}{1-\gamma}} [\sum_d \frac{1}{\sigma}(\frac{\sigma}{\sigma-1})^{1-\sigma}\Gamma(\frac{\delta+1-\sigma}{\delta})P_d^{\sigma-1}X_d\Psi_{id}^{\frac{\sigma-1}{\delta}}]^{\frac{1}{1-\gamma}} z^{R\frac{1}{1-\gamma}} \underline{z}_i^{P\kappa_P} \int_{\underline{z}_i^P}^{\infty} z^{P\frac{\sigma-1}{1-\gamma}-\kappa_P-1} dz^P$$

$$= K_i' z^{R\frac{1}{1-\gamma}},$$

where $K_i' = \frac{\kappa_P}{\kappa_P - \frac{\sigma-1}{1-\gamma}}(1-\gamma)(\frac{\gamma}{\underline{w}_i})^{\frac{\gamma}{1-\gamma}} [\sum_d \frac{1}{\sigma}(\frac{\sigma}{\sigma-1})^{1-\sigma}\Gamma(\frac{\delta+1-\sigma}{\delta})P_d^{\sigma-1}X_d\Psi_{id}^{\frac{\sigma-1}{\delta}}]^{\frac{1}{1-\gamma}} \underline{z}_i^{P\frac{\sigma-1}{1-\gamma}}$. Therefore the cutoff innovation efficiency level for a firm from o to open R&D center in country i is given by:

$$(\hat{z}_{oi}^R)^{\frac{1}{1-\gamma}} = \frac{c_i^R w_i^P}{K_i' \phi_{oi}^{R\frac{1}{1-\gamma}}}$$

The total fixed costs paid by firms from country o, doing R&D in country i, is

$$E_o c_d^R w_d^P (\frac{\hat{z}_{od}^R}{\underline{z}_o^R})^{-\kappa_R} = E_o (\underline{z}_o^R)^{\kappa_R} (\hat{z}_{oi}^R)^{\frac{1}{1-\gamma}-\kappa_R} (\phi_{oi}^R)^{\frac{1}{1-\gamma}} K_i'$$

Now consider total R&D expenditures incurred by offshore R&D centers of country o firms:

$$I_{oi} = E_o \frac{\gamma}{1-\gamma} \int_{\hat{z}_{oi}^R}^{\infty} \int_{\underline{z}_i^P}^{\infty} \pi_i^R(z^P, z^R \phi_{oi}^R) g_i(z^P) dz^P g_o(z^R) dz^R$$

$$= E_o \frac{\gamma}{1-\gamma} K_i' \phi_{oi}^{R\frac{1}{1-\gamma}} \frac{\kappa_R}{\kappa_R - \frac{1}{1-\gamma}} (\underline{z}_o^R)^{\kappa_R} (\hat{z}_{oi}^R)^{\frac{1}{1-\gamma}-\kappa_R},$$

so the overhead cost is a fixed share, $\frac{(1-\gamma)\kappa_R - 1}{\gamma \kappa_R}$, of total R&D expenditure by foreign firms. Noting that this ratio holds true for offshore R&D center from all other countries, except for the home country because they do not incur additional fixed costs, the labor market clearing condition, A.7, becomes:

$$w_d^P L_d^P = \frac{\sigma-1}{\sigma} Y_d + \frac{(1-\gamma)\kappa_R - 1}{\gamma \kappa_R}(1 - \frac{I_{dd}}{I_d}) I_d,$$

and we have:

$$\frac{X_d}{w_d^P} = \frac{L_d^P}{\frac{\sigma-1}{\sigma}\frac{Y_d}{X_d} + \frac{(1-\gamma)\kappa_R - 1}{\gamma \kappa_R}\frac{I_d}{X_d}(1 - \frac{I_{dd}}{I_d})}.$$

We can combine this equation with Equation A.6 to obtain the expression for gains from openness:

$$GO_d = (\frac{M_d}{M_d'})^{\frac{1}{\sigma-1}} (\frac{X_{ddd}}{\sum_i X_{idd}})^{-\frac{1}{\delta}} \lambda_{dd}^{T-\frac{1}{\delta}} \lambda_{dd}^{E\frac{1}{\delta}-\frac{1}{\sigma-1}} \frac{\frac{\sigma-1}{\sigma}}{\frac{\sigma-1}{\sigma}\frac{Y_d}{X_d} + \frac{(1-\gamma)\kappa_R - 1}{\gamma \kappa_R}\frac{I_d}{X_d}(1 - \frac{I_{dd}}{I_d})} - 1, \quad (A.8)$$

where M_d is the measure of varieties innovated in country d in the benchmark equilibrium, while M_d' is the measure of output invented under the counterfactual autarky equilibrium.

Step 3: deriving relative change in measure of varieties: The final step is to express $\frac{M_d}{M_d'}$ in

terms of observable flows. To do this, we first derive $\underline{w}_d$, the wage for the bottom researcher in country d. Notice that under the multiplicative assumption, wage schedule is $w_i(\theta) = \underline{w}_i\theta^{\frac{\beta}{\gamma}}$, and the optimal demand for researcher satisfies $l_i(z^P, z^R) = [\frac{\gamma}{\underline{w}_i}\pi_i^P(z^P)z^R]^{\frac{1}{1-\gamma}}\theta^{-\frac{\beta}{\gamma}}$.

In this case, the wage schedule can be interpreted as each unit of researcher efficiency, defined as $\theta^{\frac{\beta}{\gamma}}$, is paid a unit of wage, $\underline{w}_i$. Therefore the payment to a researcher whose ability is $2\theta^{\frac{\beta}{\theta}}$ is simply twice the payment to a researcher with ability $\theta^{\frac{\beta}{\theta}}$. The labor demand equation can be manipulated into:

$$\frac{l_i(z^P, z^R)\theta^{\frac{\beta}{\gamma}}}{z^{R\frac{1}{1-\gamma}}} = (\frac{\gamma}{\underline{w}_i})^{\frac{1}{1-\gamma}}\pi_i(z^P)^{\frac{1}{1-\gamma}},$$

which states that for an R&D center (z^P, z^R), each unit of innovation management efficiency, $z^{R\frac{1}{1-\gamma}}$, is matched with $(\frac{\gamma}{\underline{w}_i})^{\frac{1}{1-\gamma}}\pi_i(z^P)^{\frac{1}{1-\gamma}}$ unit of researcher efficiency talent. The labor market clearing condition for researcher efficiency unit is then:

$$L_i^R \int_{\Theta_i} \theta^{\frac{\beta}{\gamma}}h_i(\theta)d\theta = R_i(\frac{\gamma}{\underline{w}_i})^{\frac{1}{1-\gamma}}[\int_{\mathbb{Z}_i^P}\pi_i(z^P)^{\frac{1}{1-\gamma}}g_i(z^P)dz^P][\int_{\mathbb{Z}_i^R}z^{R\frac{1}{1-\gamma}}g_i(z^R)dz^R]$$

Therefore

$$(\frac{\gamma}{\underline{w}_i})^{\frac{1}{1-\gamma}} = \frac{L_i^R \int_{\Theta_i} \theta^{\frac{\beta}{\gamma}}h_i(\theta)d\theta}{R_i\frac{\kappa_P}{\kappa_P-\frac{\sigma-1}{1-\gamma}}z_i^{P\frac{\sigma-1}{1-\gamma}}[\sum_d \frac{1}{\sigma}(\frac{\sigma}{\sigma-1})^{1-\sigma}\Gamma(\frac{\delta+1-\sigma}{\delta})P_d^{\sigma-1}X_d\Psi_{id}^{\frac{\sigma-1}{\delta}}]^{\frac{1}{1-\gamma}}\int_{\mathbb{Z}_i^R}z^{R\frac{1}{1-\gamma}}g_i(z^R)dz^R}.$$

Substituting this into Equation A.3, we obtain the expression for the measure of varieties developed:

$$M_i = \frac{\kappa_P-\frac{\sigma-1}{1-\gamma}}{\kappa_P-\frac{\gamma(\sigma-1)}{1-\gamma}}(L_i^R\int_{\Theta_i}\theta^{\frac{\beta}{\gamma}}h_i(\theta)d\theta)^{\gamma}(R_i\int_{\mathbb{Z}_i^R}z^{R\frac{1}{1-\gamma}}g_i(z^R)dz^R)^{1-\gamma}$$

As this expression makes clear, under the multiplicative assumption of $f(z^R, \theta)$, the aggregate innovation output is a Cobb-Douglas function of total stock of innovation efficiency stock, and researcher talent stock in an economy. This expression also implies that in the absence of immigration, the ratio between the measure of varieties in the benchmark equilibrium and in the autarky equilibrium is $\frac{M_i}{M_i'} = (\frac{R_i\int_{\mathbb{Z}_i^R}z^{R\frac{1}{1-\gamma}}g_i(z^R)dz^R}{R_i'\int_{(\mathbb{Z}_i')^R}z^{R\frac{1}{1-\gamma}}g_i'(z^R)dz^R})^{1-\gamma}$, where the denominator is the stock of innovation efficiency units in autarky equilibrium when offshore R&D is not possible.

Recall that each unit of $z^{R\frac{1}{1-\gamma}}$ is matched with $\frac{\gamma}{\underline{w}_i}^{\frac{1}{1-\gamma}}$ units of researcher efficiency units. Since the draw of z^P is independent of z^R, the share of researcher efficiency units recruited by foreign R&D centers in the open economy is proportional to the share of innovation efficiency units of these R&D centers in the country, that is, $\frac{R_i\int_{\mathbb{Z}_i^R}z^{R\frac{1}{1-\gamma}}g_i(z^R)dz^R}{R_i'\int_{(\mathbb{Z}_i')^R}z^{R\frac{1}{1-\gamma}}g_i'(z^R)dz^R} = \frac{I_i}{I_{ii}}$.

The expression for the welfare gains from openness therefore is:

$$
\begin{aligned}
GO_d &= \left(\frac{X_{ddd}}{\sum_i X_{idd}}\right)^{-\frac{1}{\delta}} \lambda_{dd}^{T\,-\frac{1}{\delta}} \lambda_{dd}^{E\,\frac{1}{\delta}-\frac{1}{\sigma-1}} \left(\frac{I_{dd}}{I_d}\right)^{-\frac{1-\gamma}{\sigma-1}} \frac{\frac{\sigma-1}{\sigma}}{\frac{\sigma-1}{\sigma}\frac{Y_d}{X_d} + \frac{(1-\gamma)\kappa_R-1}{\gamma\kappa_R}\frac{I_d}{X_d}\left(1-\frac{I_{dd}}{I_d}\right)} - 1 \\
&= \left(\frac{X_{ddd}}{\sum_i X_{idd}}\right)^{-\frac{1}{\delta}} \left(\frac{\sum_i X_{idd}}{X_d}\right)^{-\frac{1}{\delta}} \left(\frac{\sum_l X_{dld}}{X_d}\right)^{\frac{1}{\delta}-\frac{1}{\sigma-1}} \left(\frac{I_{dd}}{I_d}\right)^{-\frac{1-\gamma}{\sigma-1}} \frac{\frac{\sigma-1}{\sigma}}{\frac{\sigma-1}{\sigma}\frac{Y_d}{X_d} + \frac{(1-\gamma)\kappa_R-1}{\gamma\kappa_R}\frac{I_d}{X_d}\left(1-\frac{I_{dd}}{I_d}\right)} - 1,
\end{aligned}
\tag{A.9}
$$

which is equivalent to Equation 17 in Section 2.

A.5 The Gains from Openness in the Literature

I compare the gains-from-openness formula in this model, given by Equation A.9, to the formula in Ramondo et al. (2015) and Arkolakis et al. (2014), both of which feature trade and offshore production, but not offshore R&D. Their formulas are given by the following:

$$
GO_d = \overbrace{\left[\left(\frac{X_{ddd}}{X_d}\right)^{-\frac{1}{\delta}} \left(\frac{\sum_l X_{dld}}{X_d}\right)^{\frac{1}{\delta}-\frac{1}{\kappa_P}}\right]}^{\text{Arkolakis et al. (2014)}} \left[\left(\frac{X_d}{Y_d}\right)^{\frac{\kappa_P+1-\sigma}{\sigma-1}\frac{1}{\kappa_P}+1}\right] - 1
\tag{A.10}
$$

$$
\underbrace{\phantom{\left[\left(\frac{X_{ddd}}{X_d}\right)^{-\frac{1}{\delta}} \left(\frac{\sum_l X_{dld}}{X_d}\right)^{\frac{1}{\delta}-\frac{1}{\kappa_P}}\right]}}_{\text{Direct Effect (Ramondo et al., 2015)}}
$$

Like Equation A.9, this equation consists of an direct and an indirect effect. There are three main differences between these two equations. First and most important, Equation A.9 features an extra term, $\frac{I_{dd}}{I_d}$, the gains
from having foreign affiliates doing R&D domestically. The second difference is that, the power on the direct effect is different across these two equations. Specifically, κ_P, the dispersion parameter for production efficiency distribution, does not appear in Equation 17. This is because by assuming out the fixed marketing costs, the extensive margin of exporting vanishes, and the elasticity of substitution between varieties alone determines trade elasticity. Third, while in Equation A.10, the strength of the indirect effect only depends on $\frac{Y_d}{X_d}$, in the present paper, it depends on $\frac{I_d}{X_d}$ and $\frac{I_{dd}}{I_d}$, too.

Appendix B   Data and Empirics

In this section I provide the background of the data used in the empirical and quantification section, and additional robustness. I then present empirical evidence on positive assortative matching between firms and inventors.

B.1   Patent Data

The main source of data I use to measure offshore R&D is patenting information from the USPTO. I use firm-level data for empirical exercises in Section 4. I use aggregate bilateral statistics constructed by the OECD from the USPTO database for calibration and quantitative exercises. The notion of offshore R&D employed in the construction of the OECD data is the same as how I define offshore R&D, but the OECD took extra efforts to ensure, to the extent possible, that patents filed under the name of affiliates in host countries are rightly classified as invented in by foreign affiliates, rather than domestic firms. For example, a patent filed by Apple China should be classified as invented by a foreign affiliate, rather than a domestic Chinese company.

One important drawback of using patent data is that, it might be biased due to differential selection into patenting across countries. For example, if only firms selling to the U.S. patent at USPTO, then the measure of offshore R&D will be biased towards these firms. I present two pieces of evidence to show this selection is not important in the context of this paper.

Figure B.1: The Comparison of Two Measure
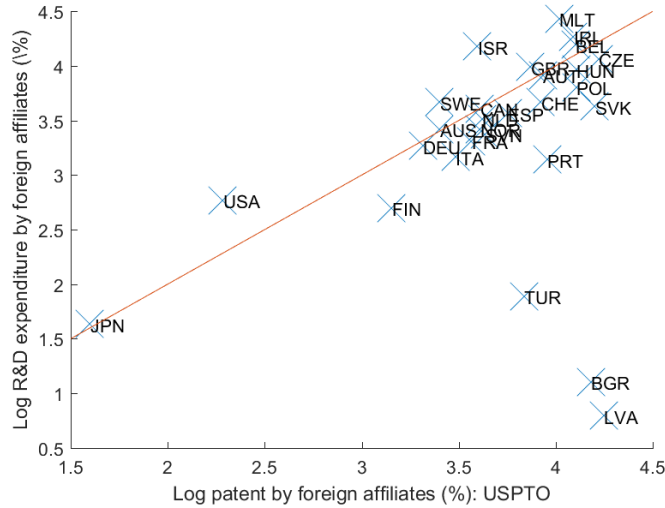
Notes: The figure plots measure of offshore R&D based on R&D expenditure against the measure based on patenting from the OECD harmonized USPTO data, averaged over 1998-2007. Redline indicates perfect correlation.

Table B.1: Correlation Between Various Measures of Offshore R&D

	Full Sample			
	R&D	USPTO	EPO	PCT
R&D	1.00			
USPTO	0.37	1.00		
EPO	0.42	0.89	1.00	
PCT	0.57	0.88	0.93	1.00

	Excluding Three Outliers			
	R&D	USPTO	EPO	PCT
R&D	1.00			
USPTO	0.72	1.00		
EPO	0.67	0.90	1.00	
PCT	0.76	0.93	0.93	1.00

Notes: The upper panel presents correlations between four measures of the share of R&D activities done by foreign firms in a host country. The four measures are based on R&D expenditures, and patenting data from three patent offices: USPTO, EPO, and PCT. The lower panel present the correlation excluding three outlier countries: Latvia, Bulgaria, and Turkey.

First, in Figure B.1, I compare the patent-based measure of offshore R&D to the expenditure-based measure, to show that such selection is unlikely to be important in the context of offshore R&D. As the figure indicates, other than three outliers, Latvia, Turkey, and Bulgaria, the two measures line up closely.

Second, if differential selection in patenting in the U.S. due to product market consideration is important, one should expect the data based on the European Patent Office (EPO) and the Patent Cooperation Treaty (PCT) to give different results. The upper panel in Table B.1 presents the correlation matrix of the four measures. As the table indicates, the three patent-based measures are close to each other, and they are all different from the expenditure-based measure. However, the discrepancies are mainly driven by the three outliers, Latvia, Turkey, and Bulgaria. Once the three outliers are excluded, as the lower panel in Table B.1 shows, all patent-based measures are strongly correlated with the expenditure-based measure.

B.2 Additional Robustness with Alternative Measures

In this subsection I discuss additional robustness tests of results in Section 4, using alternative measures of host relative talent abundance, R&D center innovation output, and firm innovation efficiency.

There are two sets of robustness, which I report in Table B.2. In columns 1-4, I aggregate data across all patent categories to construct measure for firm R&D output and host relative talent abundance. In columns 5-8, I use category-level data to construct R&D output as well as talent abundance (in regression, only the main category of each firm is kept). In all these regressions, I include full control variables from the second column in Table 2.

Within each of these two sets, I vary how I measure key variable to see if the results are sensitive. In columns 1-3 and 5-7, I use the same measure for the dependent variable, but vary how I construct independent variables. Specifically, I first vary the cutoff in defining "top" inventors and "top" firms from the top 1% in the baseline analysis, to top 10% in columns 1 and 5. In columns 2 and 6, I use the ratio between the average number of patents by inventors and the average number of patents by firms as the measure for the relative abundance in talented researchers. Since the dispersion in inventors' output is primarily driven by the output of the most talented inventors in a country, in columns 3 and 7, I use the ratio between the standard deviation of inventor output and the standard deviation of firm innovation output as a proxy for the relatively abundance of talented researchers. Finally, in columns 4 and 8, I use citation counts, rather than patent counts, to measure both the outcome variables and firms' R&D efficiency. All these alternative measures yield similar results.

Table B.2: Determinants of Offshore R&D: Alternative Measures

Regression Level:	Aggregated Across Categories				Only Firms' Main Category			
Dependent Variable:	Benchmark (patent)			Citation	Benchmark (patent)			Citation-based
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Host relative talent abundance 2* I (Parent R&D >median)	0.146*				0.259*			
	(0.083)				(0.145)			
Host relative talent abundance 3* I (Parent R&D >median)		0.110*				0.254*		
		(0.056)				(0.130)		
Host relative talent abundance 4* I (Parent R&D >median)			0.140**				0.225*	
			(0.066)				(0.127)	
Host relative talent abundance* I (Parent R&D2 >median)				0.106**				0.173*
				(0.048)				(0.093)
Home-Host FE	X	X	X	X				
Host-Home-Category FE					X	X	X	X
Firm FE	X	X	X	X	X	X	X	X
Full Controls	X	X	X	X	X	X	X	X
Observations	8083	8149	8149	7687	4882	5265	5238	3907
R^2	0.486	0.486	0.486	0.501	0.591	0.588	0.587	0.600

Notes: All regressions include the full set of controls in the second column of Table 2. The definition of common variables are the same as Tables 2. *Host relative talent abundance 2* is defined as the ratio between the share of top inventors and the share of top firms, in which "top" is defined as among global top 10%. *Host relative talent abundance 3* is defined as the ratio between the average number of patents by inventors, and the average number of patents by firms. *Host relative talent abundance 4* is defined as the ratio between the standard deviation of the number of patents by inventors, and the standard deviation of the number of patents by firms. *Parent R&D 2* is the total citation of the patents invented by a parent company in its home country.

standard errors (two way clustered at the host-country and parent-company levels) are in parentheses.

* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

B.3 Evidence on Positive Assortative Matching

The log-supermodularity assumption implies strict positive assortative matching between firms and researchers. While there is a large literature that studies assortative matching in the labor market in general (see, for example, Abowd et al., 1999 and the references thereto), and a more recent literature that focuses on the markets for CEOs and managers (Terviö, 2008, among others), to the best of my knowledge, there has been no prior research that specifically investigates the match between inventors and firms. This subsection describes the data and the main findings from this test.

B.3.1 Measures and Specification

I rely also on the USPTO patent-level data for this test. As in the empirical exercises reported in Section 4, I construct a panel of inventors and firms using the inventor identifier from Li et al. (2014) and the firm identifier from the NBER patent database project (Hall et al., 2001). This data set has a structure that resembles that of a matched employer-employee data set, except that here a match only shows up in a given year, when a patent is filed.

Using lagged innovation as a proxy for inventor ability, and various lagged measures of firm innovation efficiency, I investigate whether more talented inventors are more likely to switch to high-efficiency firms. The idea is that, if there are greater values for high-talent inventors to work with high-efficiency firms, such matches should show up more in the data than other kinds of matches. Since patents are invented jointly by inventors and firms, correlating these two measures would pick up their mechanical correlation. To avoid this problem, I focus on a sample of inventors that switch firms and examine, among them, whether the more innovative ones are more likely to move to more productive firms.

B.3.2 Results

The main findings are reported in Table B.3. There are two panels in Table B.3, each corresponding to a set of regressions with the same outcome variable. Each specification in the table regresses a measure of firm innovation efficiency on a measure of inventor talent, on a sample of inventors that have just moved to a new firm. The independent variable, same across all panels, is my preferred measure of inventor quality, which is the lagged value of log total forward citations to the patents filed by the inventor to date, adjusted by the number of inventors on each of these patents. The lagged value refers to the previous observation of the inventor in the database when he/she does not work for the present employer. This might be a few years back, however, if an inventor's last patent is from the distant past. The dependent variable in panel A is the lagged value of the log of total number of forward citations to the patents a firm has been granted. I use lag value here to ensure that the inventor under investigation is not also included in the outcome variable, leading to a mechanical correlation.

The first column adds no control variables. In the second column, I add the years since first patenting for firms and inventors to capture the life cycle effects, as it is plausible that inventors with different ages prefers firms at different stages of growth, for reasons not necessarily related to firms' innovation efficiency. In the third column, I add year fixed effects as well as category fixed effects. After controlling for these fixed effects, the point estimate shrinks somewhat, but is still statistically significant. Column 3 is my preferred specification. The point estimate indicates that, an inventor with talent that is 1% higher will be matched to a firm with 0.2% higher innovation efficiency.

Columns 4 through 6 push further by adding fixed effects for previous employers, current

Table B.3: The Match Between Firms and Inventors

	(1)	(2)	(3)	(4)	(5)	(6)
Panel A			Outcome: ln (Total citation to patents of the firm) $_{t-1}$			
Inventor Quality: Measure 1$_{t-1}$	0.192***	0.290***	0.228***	0.025***	0.095***	0.025***
	(0.006)	(0.006)	(0.006)	(0.004)	(0.007)	(0.005)
Observations	54733	54733	54733	54733	54733	54733
Firm/Inventor Controls		X	X	X	X	X
Year/Category FE			X	X	X	X
New Employer FE				X		X
Previous Employer FE					X	X
R^2	0.017	0.240	0.315	0.898	0.664	0.927

	(1)	(2)	(3)	(4)	(5)	(6)
Panel B			Outcome: Firm Productivity (total citation per inventor) $_{t-1}$			
Inventor Quality: Measure 1$_{t-1}$	2.460***	3.598***	3.244***	0.301***	1.480***	0.283***
	(0.080)	(0.092)	(0.091)	(0.061)	(0.098)	(0.066)
Observations	56096	56096	56096	56096	56096	56096
Firm/Inventor Controls		X	X	X	X	X
Year/Category FE			X	X	X	X
New Employer FE				X		X
Previous Employer FE					X	X
R^2	0.037	0.059	0.101	0.673	0.458	0.801

Notes: The regressions reported in this table use a sample of inventors that have switched firms. The independent variable is the lagged value of the log of total forward citations to the patents filed by the inventor to date, adjusted for the number of inventors for each of these patents. *Firm/Inventor Controls* refers to years since first patenting for the firm and for the inventor. *Year/Category FE* refers to year fixed effects and category fixed effects. Categories here are defined by Hall et al. (2001). There are in total six categories: chemical (excluding drugs), Computers and Communications, drugs and medical, electrical and electronics, mechanical, and others.

Robust standard errors in parentheses.

* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

employers, and both, respectively. Identifying the effect from *overtime* changes for a given origin or destination employer help overcome biases that might arise from unobserved firm heterogeneity. The cost is that the attenuation effect might be stronger. As indicated by the R^2, when both current and former employer fixed effects are added in Column 6, they absorb most of the variation. The coefficient decreases by more than 80%. However, it is still statistically significant.

Total forward citations might be a noisy measure of invention efficiency for the firm. For example, firms with larger researcher teams tend to have more inventions, hence higher citations, to their patents. Although researcher team size is a theory-consistent measure for innovation efficiency, some firms might have more inventors for reasons outside the model. Panel B uses a measure similar in spirit to firms' "labor productivity", defined as the per-inventor total forward citations in a given year, to address this concern. The results are all statistically and economically significant. The preferred specification in column 3 suggests that one percent increase in inventor productivity would increase the per-inventor citation of his/her employer by 3.

Although not reported, I also perform robustness using alternative numeric measures of productivity, results based on rankings rather than numeric measures, and results from a sample of job switchers likely due to "exogenous job loss" (original employers stopping patenting). The evidence on positive assortative matching between inventors and firms is robust to these alternative choices. These results are available upon request.

Appendix C   Quantification

This section provides additional information on the quantification section.

C.1   Data

I use the OECD harmonized USPTO data to construct the bilateral offshore R&D measure as discussed in Appendix B. The first column in Table C.1 reports the R&D by foreign affiliates based on this measure.

The calibration uses the World Management Survey (Bloom et al., 2012) and an internationally comparable cognitive ability score database (Hanushek and Woessmann, 2012). I compute the mean, standard deviation, and skewness of innovation management efficiency distribution in each country, by computing the corresponding statistics of the exponent of firm-level talent management scores for each country. I take exponent so that the distribution of scores has a right tail that resembles the firm size distribution. The distribution statistics for cognitive test scores are directly from Hanushek and Woessmann (2012). These statistics include the average cognitive score for high school students in a country, the share of students that achieve "top" performance, and the share of student that achieve "basic" performance. Thresholds for "top" and "basic" performance are defined in absolute level so the shares are comparable internationally. These statistics are reported in Table C.1.

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 R^2 of these regressions are all above 0.85. In general, geographic-region dummies have biggest explanatory power. Table C.1 indicates which countries have imputed management scores.

The model economy consists of the 25 countries reported in the table, and a statistical aggregation of another 23 countries: Argentina, Belarus, Switzerland, Chile, Colombia, Costa Rica, Guatemala, Croatia, Iran, Islamic Rep, Israel, Lebanon, Malaysia, Norway, New Zealand, Saudi Arabia, Singapore, El Salvador, Thailand, Tunisia, Uruguay, Venezuela, South Africa. The main

constraint in modelling these countries explicitly is the availability of World Management Survey and the World Input-Output Database. In calibrating the distributions for this "country", I use the same imputation method as described above when World Management Survey is not available, and then use country population as weights to compute the average distribution statistics.

Table C.1: Country Characteristics

| ISO | Offshore R&D | Innovation Mgt. Dist. | | | | Talent Dist. | | |
		Mean	Std.	Skewness	Imputed	Mean	Top Share	Basic Share
AUS	27.09	6.43	3.64	1.88		5.09	11.241	93.84
AUT	50.41	6.14	3.80	2.49	Yes	5.09	9.7359	93.106
BEL	57.25	6.20	3.85	2.48	Yes	5.04	9.3775	93.129
BRA	37.29	5.26	3.33	2.35		3.64	1.087	33.846
CAN	33.73	8.40	6.09	2.01		5.04	8.3302	94.843
CHN	52.24	5.94	2.74	1.85		4.94	8.3393	93.478
DEU	23.99	8.21	5.25	2.20		4.96	8.1654	92.694
DNK	33.69	6.87	5.00	2.67	Yes	4.96	8.7538	88.78
ESP	43.23	5.29	3.46	2.21		4.83	7.9337	85.88
FIN	18.22	7.00	5.10	2.67	Yes	5.13	12.386	95.775
FRA	33.97	6.43	4.25	2.52		5.04	8.4914	92.621
GBR	45.88	7.33	4.84	2.06		4.95	8.7857	92.885
GRC	58.26	5.63	3.70	1.89		4.61	4.2429	79.772
IND	58.03	5.93	5.00	3.02		4.28	1.2742	92.188
IRL	55.46	7.14	6.73	3.86		4.99	9.4048	91.371
ITA	29.70	6.47	4.15	2.17		4.76	5.4455	87.543
JPN	5.33	7.83	5.57	1.82		5.31	16.758	96.669
KOR	5.05	6.76	4.05	1.88	Yes	5.34	17.835	96.159
MEX	49.42	6.90	4.43	1.66		4.00	0.88366	48.933
NLD	34.82	6.56	4.14	2.46	Yes	5.12	9.1594	96.536
POL	60.79	7.25	4.60	1.73		4.85	9.8585	83.764
PRT	50.24	5.38	2.99	1.94		4.56	3.1628	80.269
SWE	27.25	7.06	4.17	1.99		5.01	8.7552	93.943
TUR	51.82	5.86	2.58	2.09		4.13	3.9199	58.233
USA	8.08	10.94	8.15	2.15		4.90	7.3299	91.824

Notes:"*Offshore R&D*" refers to the share (%) of patents invented in a country but owned by firms from foreign countries, based on the USPTO data. "*Innovation Mgt. Dist.*" refers to the sample distribution statistics constructed from the World Management Survey as described in Section C.1. "*Imputed*" indicates whether the innovation management distribution statistics are imputed. "*Talent*" refers talent distribution statistics from Hanushek and Woessmann (2012), in which "*Mean*" is the mean score for a country, and "*Top Share*" and "Basic Share" are share of students achieving "top" and "basic" performance, respectively. The performance standards are common across countries.

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
$\mathbb{I}_{z^P \in G_H^P}$	11340	0.0109	0.104	0	1

Notes: This table presents the summary statistics for firm-level innovation management score and the indicator for whether a firm is in the top 1% production efficiency.

C.2 Calibration

C.2.1 Relating Production Efficiency to Innovation Efficiency

To discipline the relationship between firms' innovation and production management efficiencies, I use micro data from the World Management Survey to estimate the following equation:

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

This data base covers around 11000 firms from 34 countries. I classify a firm as being a H type, if its production management scores falls in the top 1% in the sample. Because in calibration, I assume the management score in the model, z^R, is exponent of the management score in the data, in this estimation, I transform the innovation score accordingly. Table C.2 presents summary statistics on innovation management score, defined this way, and the indicator for H type.

Table C.3 presents result from Logit estimation of Equation C.1, using the full sample. Consistent with positive correlation between innovation and management efficiency, the estimate for A is positive and statistically significant.

C.2.2 Estimating the Matching Function

In calibration, I use a nonparametric patching function to determine the value for the complementarity parameter. I estimate this matching function based on Panel B of Table B.3. The measure for firm innovation efficiency and inventor talent is the same as in the regression. I estimate a local linear regression of inventor talent on firm innovation efficiency, focusing on the job-switching inventors. I control for the fixed effects for year and patent category, as well as firm and inventor age, defined as years since firs time the firm/inventor appears in the USPTO database.

C.2.3 Model Fit: Additional Figures

In assessing the model fit, I construct a measure of foreign affiliates' managerial advantage. The measure I use is the ratio between the average innovation management score of foreign affiliates in each country, and the average score of domestic firms in that country. Figure C.1 plots this ratio to its data counterpart for each country. The figure shows that, consistent with the summary statistics in Table 4, there is a positive relationship between the model and the data, although the model over predicts the premium.

Table C.3: Estimates for A and B

Dependent Variable:	$\mathbb{I}_{z^P \in G_H^P}$
	Coeff
z^R	0.167***
	(0.009)
Constant	-6.30***
	(0.17)
Pseudo R^2	0.2545
N	11338

Notes: This table presents results from a Logit regression of the high production efficiency indicator $\mathbb{I}_{z^P \in G_H^P}$, 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 1% in the world.

Standard errors are in parentheses

* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Figure C.1: Foreign Affiliate Managerial Advantage
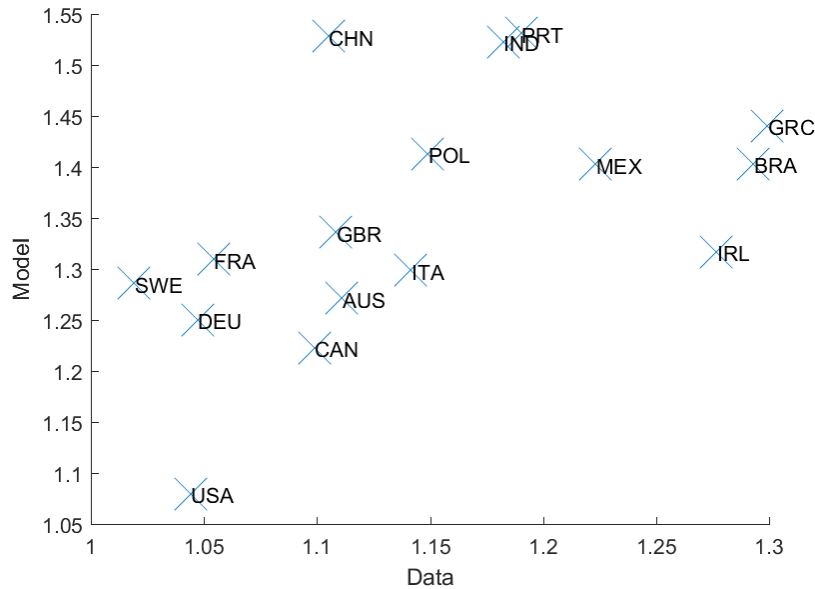
Notes: The vertical axis shows the model-based foreign affiliate innovation managerial advantage, defined as the ratio between the average innovation management score of foreign affiliates in each country and the average score of domestic firms. The horizontal axis shows the empirical counterpart of this ratio, based on the data from Bloom et al. (2012). The correlation between the model and the data is 0.53.

C.3 Computational Algorithm

To solve the model and calibrate it to match the data, I use the following calibration algorithm:

1. Choose c^M, c^R, z_H^R, α, κ^P.

2. Choose $\{\Lambda_l\}$, country-specific production efficiency.

3. Choose $\{\phi_{oi}^R\}$, $\{\phi_{il}^P\}$, $\{\tau_{ld}\}$, bilateral frictions.

4. Solve the model, compare the model-predicted bilateral shares in offshore R&D, offshore production, and trade, to the data. If they are not the same, go back to Step 3, and update international frictions accordingly. Otherwise proceed to the next step.

5. Compare the model-generated GDP with the data. If they are not the same, go back to step 2, and update country-level productivity accordingly.

6. Compute the moments on the firm size distribution, and the extensive margin of offshore R&D and trade, reported in Table 4. Compare these moments to their data counterparts. If these moments are not the same with the data, go back to Step 2. If they are the same, then the calibration procedure is finished.

I solve the model in step 4 of this calibration procedure, using the following algorithm. I start with a guess of aggregate variables $\{X_i\}$, $\{P_i\}$, $\{w_i^P\}$, cutoffs for workers to become researchers, $\{\hat{\theta}_i\}$, and cutoffs for offshore R&D, $\{\hat{z}_{oi}^R\}$. Given the aggregate variables, I solve for the cutoffs to export for each producer-consumer country pairs, $\{\hat{z}_{ld}^P\}$, and the corresponding per-variety profit, $\{\pi_i^P(z^P)\}$. I then use $\{\hat{z}_{oi}^R\}$ to solve for the measure and efficiency distribution of R&D centers in each country. With this information, I solve the researcher labor market equilibrium in each country, finding the researcher wage schedule, $w_i(\theta)$ and the matching function $T_i(z^R)$, which further allows me to solve for the number of varieties in each country, and the productivity for these varieties. The offshore and trade block of the model then determines $\{X_{ild}\}$. Based on $\{X_{ild}\}$, I distribute all revenues from sales to the production workers, researchers, and firm owners from different countries.

I then update the guess $\{X_d\}$ using the current account balance conditions, $\{w_d\}$ using the production labor market clearing condition, and $\{P_d\}$ using the model-implied price indices (Equation 12). I also update the guess for occupation choice and offshore R&D based using their respective indifference conditions. I continue this process until the updated aggregate objects and the cutoffs are the same as the input.

A crucial step in solving the model is to solve for the research market equilibrium for each country, characterized by Equations 6, 7, and 8. This is computationally difficult for two reasons. First, due to firms' offshore R&D decisions, the density of R&D center efficiency distribution is discontinuous. This discontinuity leads to kinks in the matching function. Commonly used boundary value problem solver take a long time or, under many parameter values, fail to find the solution. Second, the density of R&D center efficiency distribution depends on the offshore R&D decision of firms from all over the world. Evaluating the R&D density function therefore requires summing over all home countries (Equation 10). Similarly, we also need to evaluate $\pi_i^P(z^P)$, which depends on aggregate variables of all countries. The computational burden increases quadratically in the number of countries in the economy.

I solve the first problem by using the "shooting" method, that is, to recast the boundary value problem as a sequence of initial value problems. Specifically, given a wage for the bottom researcher, $w_i(\hat{\theta}) = \underline{w}_i$, Equations 6, 7 constitute an initial value problem. This problem can be solved

by simply forward integrating the two Equations starting from the initial value of wage. I use the Runge-Kutta Cash–Karp method in solving the initial value problems.

Let the solution to the initial problem be $T_i(z^R|\underline{w}_i)$. If $T_i(\bar{z}^R|\underline{w}_i) = \bar{\theta}_i$, then the solution to the initial problem is also the solution to the original boundary value problem. We can therefore search over the initial wage, $w_i(\hat{\theta}_i)$ and solve a sequence of initial value problems until we find the solution to the original problem. Further, as shown in Proposition 4 at the end of this section, $T_i(\bar{z}^R|\underline{w}_i)$ decreases monotonically in $\underline{w}_i$. This feature of the model makes this search efficient and robust.

To further speed up the process, I use the model feature that given all the aggregate variables, the research market equilibrium in each country is independent. I use the OpenMP protocol to parallelize the computation. In solving for each researcher market equilibrium, evaluating $g_i^R(z^R)$ and $\pi_i^P(z^P)$ requires summing over all home countries. I further parallelize this process by using SIMD.

In the following proposition, I prove the monotonicity of $T_i(\bar{z}^R|\underline{w}_i)$ in w_i:

Proposition 4 *Define $T_i(z^R|\underline{w})$ and $w_i(\theta|\underline{w})$ as the solution to the initial problem given by Equations 7, 6, and initial conditions $w_i(\hat{\theta}_i) = \underline{w}$, $T_i(\underline{z}^R) = \underline{\theta}_i$. Then the end value of the solution to this initial problem, $T_i(\bar{z}_i^R|\underline{w})$, decreases in $\underline{w}$.*

Proof Consider two wages for the bottom researcher in country i, $w^1 < w^2$. This proposition claims that $T_i(\bar{z}_i^R|w^1) \geq T_i(\bar{z}_i^R|w^2)$. I prove by contradiction.

Suppose $T_i(\bar{z}_i^R|w_1) < T_i(\bar{z}_i^R|w_2)$. Given that $w^1 < w^2$, Equation 7 implies that at $T_i'(\underline{z}^R|w_1) > T_i'(\underline{z}^R|w_2)$, that is, at least initially at $\underline{z}_i^R$, when facing a lower wage w_i, R&D centers will hire a larger number of researchers. This means that at an ϵ interval to the right of $\underline{z}_i^R$, $T_i(z^R|w_1) > T_i(z^R|w_2)$. Since $T_i(\bar{z}_i^R|w_1)$ and $T_i(\bar{z}_i^R|w_2)$ are both continuous function of z^R, for $T_i(\bar{z}_i^R|w_1) \leq T_i(\bar{z}_i^R|w_2)$ to hold, there must be at a point $\tilde{z}^R$, such that $T_i(z^R|w_2)$ crosses $T_i(z^R|w_1)$ at $\tilde{z}^R$ for the first time from below. Suppose $\tilde{\theta} = T_i(\tilde{z}^R|w_1) = T_i(\tilde{z}^R|w_2)$. From Equation 7, $w_i(\tilde{\theta}|w_1) > w_i(\tilde{\theta}|w_2)$.

From Equation 6, $w_i(\theta|\underline{w}_i) = \underline{w}_i \exp \int_{\hat{\theta}_i}^{\theta} \frac{f_2(T_i^{-1}(x),x)}{\gamma f(T_i^{-1}(x),x)} dx$. Under the log-supermodularity assumption of function f, the integrant on the right hand side increases with z^R. Because $\tilde{z}^R$ is the first point where the two matching functions intercept, for all $\theta \in (\hat{\theta}_i, \tilde{\theta})$, $T_i^{-1}(\theta|w_1) < T_i^{-1}(\theta|w_2)$. Therefore $w_i(\tilde{\theta}|w_1) < w_i(\tilde{\theta}|w_2)$, which contradicts the above result.