Innovation and Production in the Global Economy*

Costas Arkolakis	Natalia Ramondo	Andrés Rodríguez-Clare	Stephen Yeaple
Yale and NBER	UC San Diego	UC Berkeley and NBER	Penn State and NBER

December 2014

Abstract

The decline in the costs of multinational production (MP) has led some countries to specialize in innovation and others to specialize in production. To study the aggregate and distributional implications of this phenomenon, we develop a quantifiable general equilibrium model of trade and MP. Specialization is endogenously determined as a result of comparative advantage and home market effects (HME) that arise from the interaction between increasing returns to innovation and geographical frictions. The model yields simple structural expressions for bilateral trade and MP that we use to calibrate it across a set of OECD countries. Comparative statics exercises reveal that the reduction in the cost of MP or the integration of China into the world economy may hurt countries that are driven to specialize in production due to HMEs, although these losses tend to be very small. Contrary to popular fears, we find that production workers gain even in countries that further specialize in innovation.

*We thank Pol Antras, Ariel Burstein, Arnaud Costinot, Jonathan Eaton, Gene Grossman, Samuel Kortum, James Markusen, Felix Tintelnot, and Jonathan Vogel, as well as seminar participants at several seminars and conferences for insightful comments. We also thank Treb Allen, Jakub Kominiarczuk, and Masayuki Sawada for excellent research assistance. Rodriguez-Clare and Yeaple would like to thank the Human Capital Foundation, Rodriguez-Clare the Center for Equitable Growth, and Arkolakis the NSF for support. All remaining errors are our own. The statistical analysis of firm-level data on U.S. multinational corporations reported in this study was conducted at the International Investment Division, U.S. Bureau of Economic Analysis, under arrangements that maintained legal confidentiality requirements. Views expressed are those of the authors and do not necessarily reflect those of the Bureau of Economic Analysis.

1 Introduction

One consequence of globalization, and in particular the rise of multinational production (MP), is that goods are increasingly being produced far from where ideas are created. International specialization in innovation and production is clearly evident in the aggregate data. Figure 1 shows that the most innovative OECD countries, as measured by R&D expenditures in manufacturing relative to local value-added, are home to multinationals whose foreign affiliate sales exceed the sales of foreign multinational affiliates in their country. With increasing globalization, this pattern has become more pronounced over time. Figure 2 shows that R&D expenditures relative to manufacturing value-added in the United States has grown from 8.7% in 1999 to 12.7% in 2009. Over the same period, American firms have increased the share of their total global employment that is located in their foreign affiliates from 22% to 31%. This is thanks in large part to an increasing presence in China, which now accounts for one in eight employees of the foreign affiliates of U.S. firms.

Being a recent phenomenon, specialization in production or innovation raises a host of concerns. Countries that specialize in production worry that low innovation will damage their growth prospects, while countries that specialize in innovation worry that this will reduce the availability of good middle-income jobs. Viewed from the perspective of the standard two-sector trade theory, concerns about specialization in production sound misguided, since specialization would reflect comparative advantage and bring about efficiency gains. But there are two reasons why it is legitimate to worry about specialization in production as opposed to innovation. First, the expansion of production could trigger a deterioration of a country's terms of trade. Second, innovation is an increasing-returns-to-scale activity, so standard reasoning based on comparative advantage is insufficient. In particular, the combination of fixed costs of innovation and frictions to the movement of ideas across borders leads to home-market effects (HMEs) in innovation, and, as shown by Venables (1987), specialization induced by HMEs could lead to non-standard welfare effects.

In this paper we develop the first quantifiable, multicountry general-equilibrium model of trade and MP that captures these forces in a rich geographic setting. We use the model to quantify the welfare implications of shocks driving increased specialization in innovation and production, including a generalized reduction in the cost of transferring technologies across borders, the integration of China into the global economy, and the impact of selective integration between countries.

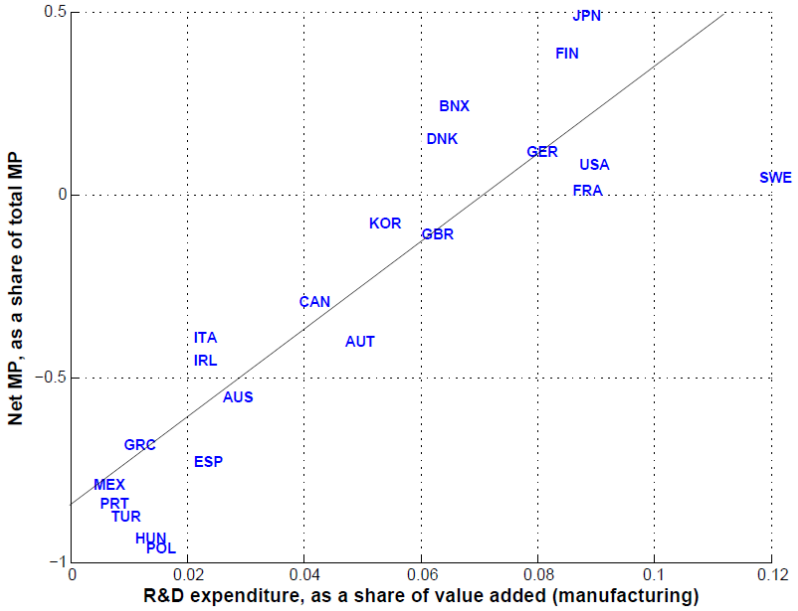


Figure 1: R&D and Net MP

Note: R&D expenditure in manufacturing value-added is from OECD STAN for 1999. Net MP is defined as outward affiliate sales - inward affiliate sales divided by their sum.

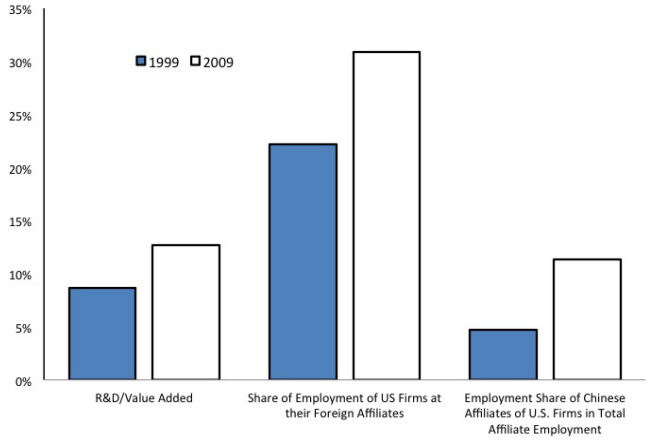


Figure 2: R&D and Employment of U.S. manufacturing firms and multinationals

Note: Sources OECD STAN, US Bureau of Economic Analysis. The employment share of US firms at their foreign affiliates is defined as total employment of US majority-owned, manufacturing affiliates abroad divided by total US manufacturing employment plus US majority-owned, manufacturing affiliates abroad minus the employment of the affiliates of foreign-owned manufacturing affiliates operating in the US.

Following Melitz (2003), we model innovation as the creation of heterogeneous firms that sell differentiated goods in monopolistically competitive markets separated by fixed and variable trade costs. We depart from the Melitz model by assuming that firms can locate production outside of their home market and assume that firm productivity levels across locations are drawn from a multivariate distribution. Firms face a tradeoff in choosing where to produce for any particular market: they could locate production close to their customers to avoid trade costs or they could locate where production costs are lower. By allowing firms to produce outside of their home country, MP allows some countries to specialize in innovation and others to specialize in production, with profits flowing from producing to innovating countries to compensate for the cost of innovation.[1] Loosely speaking, innovative countries export ideas and import goods.

The model provides a natural framework to explore the aggregate and distributional implications of openness to trade and MP. We find that countries that innovate more as a result of openness tend to realize larger gains from openness than implied by current models (e.g. Ramondo & Rodríguez-Clare (2013)), while countries that are most at risk from adverse welfare effects are those that experience a contraction of innovation. To analyze the impact of openness on income distributions within countries we consider variants of the model with limited labor mobility between innovation and production sectors and show that openness to trade and MP not only alters the distribution of income in intuitive ways but can even lead to a loss of welfare for some workers even as the country's aggregate real consumption increases. This result resonates with the popular fear that the real wage of production workers in innovative countries such as the United States may fall as multinationals move production abroad.

Our quantitative analysis starts by deriving and testing a novel implication of our model, namely that trade flows restricted to the parents and affiliates of firms from a given country are more sensitive to trade costs than overall trade flows. Using high-quality data from the Bureau of Economic Analysis on the sales of U.S. firms and their foreign affiliates, we estimate restricted and standard gravity equations and find that the trade elasticities are consistent with this prediction. These two elasticities will also serve as key targets in our calibration.

The model is calibrated using trade, MP, and production data for 26 countries. We identify the full set of trade and MP frictions between countries and a vector of parameters that

[1] In the absence of MP, the share of labor devoted to innovation would be the same in all countries. This is consistent with the version of the Melitz model presented in Arkolakis, Demidova, Klenow & Rodríguez-Clare (2008), where entry is endogenous, but not affected by trade costs. An equivalent result is derived by Eaton & Kortum (2001) in a setting with Bertrand competition.

govern comparative advantage by fitting aggregate bilateral trade and MP data under the assumptions that all trade and MP frictions are symmetric between pairs of countries. Intuitively, specialization in innovation or production that cannot be explained by geography are due to comparative advantage.[2]

We use the calibrated model to perform two counterfactual exercises. First, we consider a 5% reduction in all MP costs from their calibrated levels. This reduction results in greater specialization across countries in innovation and production and real incomes rise on average by about 2.2%. Some of the countries that experience a decline in innovation also suffer welfare losses, but these are very small. Contrary to popular fears, we find that production workers gain everywhere, and it is innovation workers who experience losses in countries that face a contraction in their innovation sector. Second, we explore the implications of the integration of China to the world economy. The result is that developed countries follow what we could refer to as the "Apple model": they specialize in innovation while China becomes a key manufacturing center for the whole world. All countries experience welfare gains, and again production workers gain everywhere, although they benefit by less than innovation workers in the countries that specialize in innovation (e.g., the United States).

The mechanisms at work in our model have antecedents in the classic work on trade and MP (see Markusen (2002)). This literature highlights four key ideas: (i) MP allows innovation (entry) to be geographically separated from production; (ii) countries differ in their relative costs of innovation and production, which leads to specialization in one of these two activities; (iii) the non-rivalry of technology within the firm allows multi-plant production; and (iv) trade costs encourage, while MP costs discourage, multi-plant production. The incorporation of these features into a general-equilibrium trade model dates back to Helpman (1984) and Markusen (1984).[3] By modeling firm-level productivity in different countries as coming from a multivariate distribution and by replacing plant-level fixed costs with marketing fixed costs, we gain the ability to construct a tractable, quantifiable, and multi-country general equilibrium model that incorporates the most important mechanisms found in this

[2]Before proceeding, one conceptual issue is worth noting. In this paper we focus on MP as the vehicle through which international specialization takes place, but there are alternative arrangements, such as the licensing of technology and other contractual relationships such as outsourcing that do not involve ownership. Our model is consistent with these mechanisms, but because there is little data on arm's length offshoring we can only measure the offshoring done within multinationals.

[3]Examples of work that most closely resembles our own are Markusen & Venables (1998) and Markusen & Venables (2000) in which the authors analyze the interaction between comparative advantage in production and innovation, trade costs, and plant and corporate fixed costs in a two-country, Heckscher-Ohlin-like setting. Grossman & Helpman (1991) extend this framework to an endogenous growth setting in which the more efficient use of the world's resources made possible by MP may affect the long-run growth rate in rich and poor countries. Non-homothetic preferences together with home market effects determine specialization and foreign investment patterns in high-quality or low quality goods in Fajgelbaum, Grossman & Helpman (2014).

earlier work.

Our paper is closely related to a recent literature on trade and MP. Ramondo & Rodríguez-Clare (2013) also have a probabilistic representation of multi-country productivity and a large number of countries, allowing for counterfactual analysis in a rich geographic setting. The key difference is that there is no innovation in their perfectly-competitive model. Our monopolistic-competition framework is also related to a recent paper by Irarrazabal, Moxnes & Opromolla (2013), which is the quantitative application of Helpman, Melitz & Yeaple (2004). They focus on understanding the frictions that rationalize the export versus MP decisions of Norwegian firms, but abstract from "export-platform MP" (any market can be served only from a local affiliate or by exports from the firm's home country) and fix firm innovation locations and country wages by assumption. With a probabilistic structure similar to ours, Tintelnot (2014) extends Irarrazabal et al. (2013) to allow for export-platform MP in a general-equilibrium model, although again in this setting there is no innovation as firm entry is exogenous.

Our paper is also related to a literature that considers the movement of managerial or knowledge capital from one country to another, interpreted as MP, while trade takes place only as a way to transfer the returns to capital (see, for example, Burstein & Monge-Naranjo (2009), McGrattan & Prescott (2010), McGrattan (2011), and Ramondo (2012)).[4] The simplification on the trade dimension in these papers allows for a more detailed modeling of the effect of specific policies, such as taxes on profits of foreign owned firms, as well as the transition path as countries open up to MP. Because they do not allow for increasing returns and frictions to trade and MP, these papers have nothing to say about bilateral trade and MP flows or about the role of HMEs and their related welfare implications.

Finally, by distinguishing between innovation and production activities, we make contact with a body of theory that emphasizes the effect of offshoring on the set of activities done within a country and on real wages (e.g. Feenstra & Hanson (1999), Grossman & Rossi-Hansberg (2008) and Rodríguez-Clare (2010)). By considering the impact of China's integration into world markets in our counterfactuals, our paper also makes contact with an empirical literature that has documented the negative effect of Chinese manufacturing exports on the employment and wages of manufacturing workers in developed country (e.g. Autor, Dorn & Hanson (2012)).

The rest of the paper is organized as follows. Section 2 develops our analytical framework

[4]Recent papers that present stylized models of innovation versus production are Eaton & Kortum (2007) and Rodríguez-Clare (2010). In principle, these models could be adapted for quantitative analysis but this task has not been undertaken so far.

and characterizes its properties. Sections 3 and 4 provide empirical estimates and calibrate the model. Section 5 assesses its quantitative implications regarding the effect of MP on welfare and the location of innovation. Finally, Section 6 concludes.

2 The Model

We consider a world economy comprised of $i = 1, ..., N$ countries; one factor of production, labor; and a continuum of goods indexed by $\omega \in \Omega$. Labor is inelastically supplied and immobile across countries. Let L_i and w_i denote the total endowment of labor and the wage in country i, respectively. In each country i, there is a representative agent with Constant Elasticity of Substitution (CES) utility function with elasticity of substitution $\sigma > 1$. The associated price index is given by

$$P_i = \left(\int_{\omega \in \Omega} p_i(\omega)^{1-\sigma} d\omega \right)^{\frac{1}{1-\sigma}}, \tag{1}$$

where $p_i(\omega)$ is the price of good ω in country i.

Each good ω is potentially produced by a single firm under monopolistic competition. Firms can produce anywhere in the world with varying productivity levels as specified below. To the extent possible, we use index i to denote the firm's country of origin (the source of the *idea*), index l to denote the *location* of production, and index n to denote the country where the firm sells its product. Firms that export from l to country n incur a marketing fixed cost in units of labor in the destination country, $w_n F_n$, and an iceberg transportation cost $\tau_{ln} \geq 1$ with $\tau_{nn} = 1$. Firms originated in country i that produce in country l incur a productivity loss that we model as iceberg bilateral MP costs, $\gamma_{il} \geq 1$, with $\gamma_{ll} = 1$. These costs are meant to capture various impediments that multinationals face when operating in a different economic, legal or social environment, as well as the various costs of technology transfer incurred by multinationals in different production locations.[5]

A firm from origin i can serve destination n by (i) producing in i and exporting to country n, by (ii) opening an affiliate in country $l \neq i, n$ and exporting from there to country n, or by (iii) opening an affiliate in n and selling the good domestically. Firms use constant returns to scale technologies, with the marginal product of labor being firm and location specific. Formally, a firm is characterized by a productivity vector $\boldsymbol{z} = (z_1, z_2, ..., z_N)$, where z_l determines the firm's productivity if it decides to produce in country l. These productivity

[5] See Kokko (1992), chapter 3, for a review of the findings of the literature measuring the costs of technology transfer to foreign affiliate firms.

vectors are allowed to vary across firms, leading firms to make different choices regarding their production locations.

Letting $\xi_{iln} \equiv \gamma_{il} w_l \tau_{ln}$, the above assumptions imply that the unit cost of a firm from i producing in location l to serve market n is $C_{iln} \equiv \xi_{iln}/z_l$. Note that all heterogeneity across firms is associated with differences in the productivity vector $\mathbf{z}$, while the trade and MP costs, $\{\tau_{ln}\}$ and $\{\gamma_{il}\}$, as well as wages (and hence ξ_{iln}), are common across firms.

We think of innovation as the process of creating firms, and assume that doing so requires f_i^e units of labor. If L_i^e units of labor are allocated to the innovation sector in country i, then the measure of firms in that country is $M_i = L_i^e / f_i^e$. We consider two opposite cases regarding worker mobility between the innovation and production sectors: no mobility and perfect mobility. With no mobility, the supply of labor to both the innovation and production sectors is perfectly inelastic.[6] Since the measure of firms in this case is pinned down by the exogenous supply of labor to the innovation sector, we refer to this case as "exogenous entry." With perfect mobility, the measure of firms is determined by the condition that workers optimally choose where to work. We refer to this case as "endogenous entry." In the theoretical analysis we restrict ourselves to the simple scenario in which workers are perfectly homogenous, implying the equalization of the wage in innovation and production. In the quantitative analysis we extend the model to allow for heterogeneity among workers, as in a Roy model – this model is discussed further in Section 5 and developed formally in Appendix B.

2.1 Firm Optimization

In this environment, firms face a simple optimization problem. First, for each market n, a firm decides which is the cheapest location from where to serve that market. Second, the firm decides which price to charge. Given our assumption on preferences, this choice leads to a mark-up of $\widetilde{\sigma} \equiv \sigma/(\sigma - 1)$ over marginal cost, so that the price is

$$p_{in} = \widetilde{\sigma} \min_l C_{iln} \,. \tag{2}$$

In Figure 3, we summarize how the price charged by a firm is determined by factors that are firm specific, i.e. the firm's productivity vector $\mathbf{z}$, and by factors which depend on the country of origin, location of production, and final sales. Third, the firm calculates the associated profits. If those profits are higher than the fixed marketing cost then the firm chooses to serve the market. Letting X_n be total expenditure in country n, the maximum unit cost under

[6] In fact, in our model there are three activities: production, marketing and entry. We always assume that there is perfect mobility between production and marketing and refer to "production workers" as those working in either production or marketing.

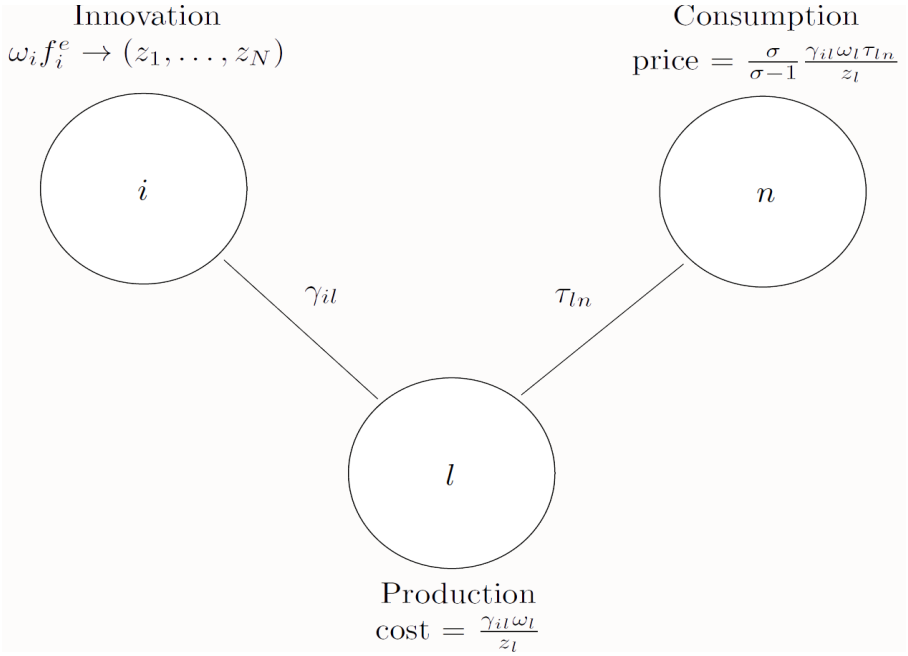


Figure 3: Firm Pricing

Note: Innovation is done in country i at cost $w_i f_i^e$, production is done in country l at unit cost $\gamma_{il} w_l / z_l$, and consumption is done in country n at price $\tilde{\sigma} \gamma_{il} w_l \tau_{ln} / z_l$.

which variable profits in market n are enough to cover the fixed cost $w_n F_n$ is defined by

$$c_n^* \equiv \left(\frac{\sigma w_n F_n}{X_n} \right)^{1/(1-\sigma)} \frac{P_n}{\widetilde{\sigma}}. \tag{3}$$

2.2 Aggregation

Although the problem for each firm is simple, our goal is to obtain analytic expressions for the aggregate variables that we can relate to the data while retaining key features of previous theories of international trade. To achieve this purpose, we consider a multivariate extension of the univariate Pareto distribution used in the Chaney (2008) version of Melitz (2003).

We assume that the productivity vector of firms in country i is randomly drawn from the multivariate distribution given by

$$\Pr(Z_1 \leq z_1, ..., Z_N \leq z_N) = G_i(z_1, ..., z_N) = 1 - \left(\sum_{l=1}^{N} \left[T_{il} z_l^{-\theta} \right]^{\frac{1}{1-\rho}} \right)^{1-\rho}, \tag{4}$$

with support $z_l \geq \widetilde{T}_i^{1/\theta}$ for all l, where $\widetilde{T}_i \equiv \left[\sum_l T_{il}^{1/(1-\rho)} \right]^{1-\rho}$, $\rho \in [0, 1)$, and $\theta > \max(1, \sigma -$

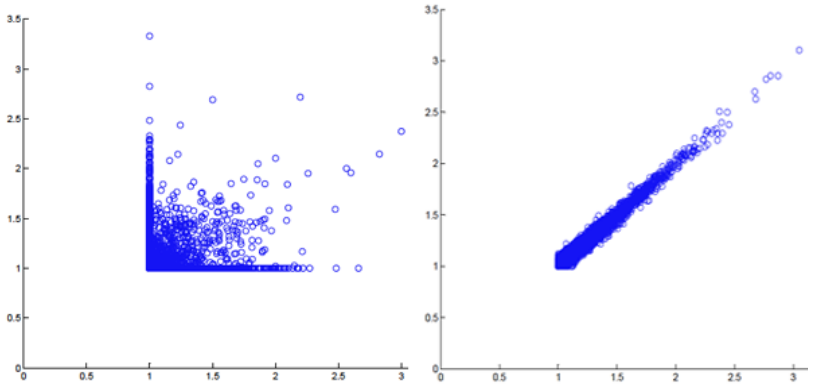


Figure 4: Simulated draws from multivariate Pareto with $\rho = 0.1$ (left panel) $\rho = 0.9$ (right panel)
Note: Simulation for 10,000 draws with distribution specified with $N = 2$, $T_1 = T_2 = 2^{\rho-1}$, $\theta = 7.2$.

1).[7] Several comments are in order regarding the properties of this distribution. First, the marginal distributions have conditional distributions that are Pareto. In particular, for $z_l \geq a > \widetilde{T}_i^{1/\theta}$ we have $\Pr(Z_l \geq z_l \mid Z_l \geq a) = (z_l/a)^{-\theta}$. Second, if $\rho \to 1$ the elements of $(Z_1, Z_2, ..., Z_N)$ are perfectly correlated. Finally, the case with $\rho = 0$ is equivalent to simply having the production location l chosen randomly with probabilities $T_{il}/\widetilde{T}_i$ among all possible locations $l = 1, ..., N$, and the productivity Z_l drawn from the Pareto distribution $1 - \widetilde{T}_i z_l^{-\theta}$ with $z_l \geq \widetilde{T}_i^{1/\theta}$. Figure 4 illustrates how the distribution depends on the value of ρ.

For the reminder of the paper, we make the following assumption.

A 1 $T_{il} = T_i^e T_l^p$.

This assumption implies that $\widetilde{T}_i = \left(\sum_l \left(T_l^p\right)^{1/(1-\rho)}\right)^{1/(1-\rho)} T_i^e$, so that we can think of T_i^e as a measure of the quality of ideas in country i, or *productivity in innovation*. In turn, T_l^p determines country l's *productivity in production*.[8] We will continue to write T_{il} rather

[7] A more detailed discussion on the properties of the distribution is provided in the Appendix. This distribution can be seen as a reformulation of the Archimedean copula of Pareto distributions. Specifically, the Archimedean copula 4.2.2 in Nielsen (2006) leads to the same function for the distribution as (4) in the two-dimensional case if z_1 and z_2 are each distributed Pareto, except that the support would be implicitly defined by $(T_1 z_1^{-\theta})^{\frac{1}{1-\rho}} + (T_2 z_2^{-\theta})^{\frac{1}{1-\rho}} \leq 1$. This distribution cannot be directly extended to $N \geq 3$ because the copula is not strict (see Nielsen (2006)). Instead, we modify the support of the distribution to make it an N-box defined by $z_l \geq \widetilde{T}_i^{1/\theta}$ for all l.

[8] This setup easily allows for splitting countries without affecting the equilibrium. For example, we could split country l into two countries, l_1 and l_2, with $T_{l_j}^e = T_l^e$ and $\left(T_{l_j}^p\right)^{1/(1-\rho)} / L_{l_j} = \left(T_l^p\right)^{1/(1-\rho)} / L_l$ for $j = 1, 2$. One can show that if there are no costs to trade and MP between l_1 and l_2 then the equilibrium is not affected by the split (the proof is available upon request).

than $T_i^e T_l^p$ for notational convenience. Note that T_i^e and f_i^e will have equivalent effects on all relevant equilibrium variables and, thus, we henceforth assume that $f_i^e = f^e$ for all i.

To guarantee that for all pairs $\{i, n\}$ there are firms from i that will decide not to serve market n, we assume that the parameters of the model (e.g., marketing costs) are such that the level of c_n^* is low enough. Formally, we make the following assumption, which we maintain throughout the rest of the paper.

A 2 $\xi_{iln} > \widetilde{T}_i^{1/\theta} c_n^*$, for all i, l, n.

The multivariate Pareto distribution together with this assumption allows us to characterize several important objects in the model, starting from the probability that a firm serves a particular destination from a certain location at some unit cost c, and the (conditional) probability that firms from i serving market n decide to do so from production location l.

Lemma 1 *The (unconditional) probability that a firm from i will serve market n from l at cost c, for* $c \leq c_n^*$, *is*

$$\Pr\left(\arg\min_k C_{ikn} = l \ \cap \min_k C_{ikn} = c\right) = \psi_{iln} \Psi_{in} \theta c^{\theta-1}, \tag{5}$$

where

$$\Psi_{in} \equiv \left[\sum_k \left(T_{ik}\xi_{ikn}^{-\theta}\right)^{\frac{1}{1-\rho}}\right]^{1-\rho} \text{ and } \psi_{iln} \equiv \left(T_{il}\xi_{iln}^{-\theta}/\Psi_{in}\right)^{\frac{1}{1-\rho}},$$

while the (conditional) probability that firms from i serving market n will choose location l for production is

$$\Pr\left(\arg\min_k C_{ikn} = l \ \mid \min_k C_{ikn} \leq c_n^*\right) = \psi_{iln}. \tag{6}$$

The formal proof of this result, as well as the rest of the results of the paper, are given in Appendix B. We now use this Lemma to analyze the model's implications for aggregate trade and MP flows. Let M_i denote the measure of firms in country i, M_{iln} denote the measure of firms from i that serve market n from location l, and X_{iln} denote the total value of the associated sales. Using the pricing rule in (2) and the cut-off rule in (3), we can compute X_{iln} by integrating firm sales using the density in (5) to obtain

$$X_{iln} = \psi_{iln} \lambda_{in}^E X_n \,, \tag{7}$$

where

$$\lambda_{in}^E \equiv \frac{\sum_l X_{iln}}{X_n} = \frac{M_i \Psi_{in}}{\sum_k M_k \Psi_{kn}}, \tag{8}$$

is the share of total expenditures in country n that are devoted to goods produced by firms from i (irrespective of where they are produced). The measure of firms behind these sales is

$$M_{iln} = \frac{\theta - \sigma + 1}{\sigma\theta} \frac{X_{iln}}{w_n F_n}. \tag{9}$$

Aggregate flows X_{iln} can be used to construct trade and MP shares. In particular, trade shares are given by expenditure shares across production locations, $\lambda_{ln}^T \equiv \sum_i X_{iln} / \sum_{i,k} X_{ikn}$, while MP shares are given by production shares across firms from different origins, $\lambda_{il}^M \equiv \sum_n X_{iln} / \sum_{j,n} X_{jln}$. Letting $Y_l \equiv \sum_{i,n} X_{iln}$ denote the value of all goods produced in country l (output), recalling that $X_n \equiv \sum_{i,l} X_{iln}$ is total expenditure by consumers in country n, and using expression (7), trade and MP shares can be written more succinctly as

$$\lambda_{ln}^T = \sum_i \frac{X_{iln}}{X_n} = \sum_i \psi_{iln} \lambda_{in}^E, \tag{10}$$

and

$$\lambda_{il}^M = \sum_n \frac{X_{iln}}{Y_l} = \frac{\sum_n \psi_{iln} \lambda_{in}^E X_n}{Y_l}. \tag{11}$$

Let Π_{iln} denote aggregate profits associated with sales X_{iln}, net of fixed marketing costs, but gross of entry costs. Given Dixit-Stiglitz preferences, variable profits associated with X_{iln} are X_{iln}/σ. The total fixed marketing costs paid by these firms are $w_n F_n M_{iln}$. Using these two expressions and (9), we obtain

$$\Pi_{iln} = \eta X_{iln}, \tag{12}$$

where $\eta \equiv 1/(\theta\tilde{\sigma})$. Therefore, total profits made in country l are a constant share of output in country l, i.e. $\sum_{i,n} \Pi_{iln} = \eta Y_l$.

2.3 Equilibrium

We start by considering the labor market clearing condition. Let L_i^p denote the amount of labor devoted to production and marketing so that

$$L_i^p = \left(1 - \eta - \frac{\theta - \sigma + 1}{\sigma\theta}\right) Y_i / w_i + \sum_{i,l} M_{iln} F_n.$$

Noting that $L_i^e = M_i f^e$, $Y_i = \sum_n \lambda_{in}^T X_n$, and using (9), we can write the labor market clearing condition for workers in production and marketing in country n as

$$\left(1 - \eta - \frac{\theta - \sigma + 1}{\sigma\theta}\right) \sum_n \lambda_{in}^T X_n + \frac{\theta - \sigma + 1}{\sigma\theta} X_i = w_i(L_i - M_i f^e). \tag{13}$$

Next, we characterize the current account balance condition.[9] For country i, total expenditure is X_i while total income equals the sum of three terms: (i) the net value of sales, which equals total sales, Y_i, minus the cost of marketing goods produced in country i, $\sum_{j,n} M_{jin} w_n F_n$; (ii) wages paid to workers engaged in marketing for sales in country i, $\sum_{j,l} M_{jli} w_i F_i$; (iii) net profits, which are equal to profits made by domestic firms, $\sum_{l,n} \Pi_{iln}$, minus profits made domestically by foreign firms, $\sum_{j,n} \Pi_{jin}$. Thus, we can write the current account balance condition as

$$X_i = Y_i - \sum_{j,n} M_{jin} w_n F_n + \sum_{j,l} M_{jli} w_i F_i + \sum_{l,n} \Pi_{iln} - \sum_{j,n} \Pi_{jin}.$$

Using (7), (12), and (13) we can rewrite this condition as

$$X_i = w_i(L_i - M_i f^e) + \eta \sum_n \lambda_{in}^E X_n. \tag{14}$$

Finally, we turn to the zero-profit condition. Total profits earned by firms created in i are $\sum_{l,n} \Pi_{iln}$ while the total cost of creating those firms is $w_i^e L_i^e$, where w_i^e is the wage paid to workers in the innovation sector. Using (12) and noting that $L_i^e = M_i f^e$, the zero-profit condition is

$$\eta \sum_n \lambda_{in}^E X_n = M_i w_i^e f^e. \tag{15}$$

Under exogenous entry, the labor supply to the production sector is exogenous, hence equations (13) and (14) constitute a system of $2N$ equations that can be used to solve for the equilibrium levels of $\boldsymbol{X}$ and $\boldsymbol{w}$ (up to a constant determined by the numeraire) given $M_i = L_i^e / f^e$ (where L_i^e is exogenous). The wage in the innovation sector, w_i^e, is recovered from (15). Under endogenous entry, perfect labor mobility between production and innovation implies that w_i^e must be equal to w_i.[10] Equations (13) , (14) and (15) with w_i^e replaced by w_i constitute a system of $3N$ equations to solve for the equilibrium levels of $\boldsymbol{X}$, $\boldsymbol{M}$, and $\boldsymbol{w}$ (up to a constant determined by the numeraire).

A key concept in the rest of the paper is the share of income earned in the innovation

[9] In this Section we impose current account balance, whereas in the quantitative section we allow for exogenous current account imbalances.

[10] We assume throughout the paper that the equilibrium with endogenous entry is interior, so that $0 < L_i^p < L_i$.

sector (or simply the innovation share), $r_i \equiv w_i^e L_i^e / \left(w_i^e L_i^e + w_i L_i^p\right)$, which is also equal to $r_i = \sum_{l,n} \Pi_{iln} / X_i$. The equilibrium conditions above imply that

$$r_i - \eta = \frac{1}{\widetilde{\sigma}} \left(\frac{X_i - Y_i}{X_i} \right) . \tag{16}$$

Therefore, the innovation share is directly related to the trade deficit, $X_i - Y_i$. In fact, in the two extreme cases of infinite MP or infinite trade costs, we must have $X_i = Y_i$ and, thus, $r_i = \eta$. The first case is discussed in more detail below.

2.4 Special Cases

In this subsection, we explore a number of special cases of the model under endogenous entry that we can characterize analytically. These cases illustrate how, in the presence of MP, comparative advantage and home market effects (HME) determine whether countries specialize in innovation or production. They also shed light on the basic forces behind the results of our quantitative analysis in Section 5.

2.4.1 Infinite MP costs - a world without MP

It is instructive to consider the case in which MP costs are infinite, i.e., $\gamma_{il} \to \infty$ for all $i \neq l$. This restriction implies that expenditure shares are equal to trade shares, $\lambda_{in}^E = \lambda_{in}^T$, and that

$$\lambda_{in}^T = \frac{M_i T_{ii} \left(w_i \tau_{in}\right)^{-\theta}}{\sum_k M_k T_{kk} \left(w_k \tau_{kn}\right)^{-\theta}}, \tag{17}$$

which is the same expression as in the Melitz/Chaney model. The equilibrium conditions under endogenous entry then imply that

$$M_i = \widetilde{M}_i \equiv \eta L_i / f^e, \tag{18}$$

so that the share of labor devoted to innovation is independent of trade costs and entry costs, and is equal to η in all countries. This is expression is the same to the one derived by Arkolakis et al. (2008) in a Melitz model with endogenous entry.[11]

[11] An equivalent result is derived by Eaton & Kortum (2001) in a setting with Bertrand competition.

2.4.2 A frictionless world - the role of comparative advantage

We now discuss the role of comparative advantage in innovation versus production. To make the analysis tractable, we focus on the case of a frictionless world, i.e., $\tau_{ln} = 1$ and $\gamma_{il} = 1$, for all i, l, n. Let $A_i \equiv \left(T_i^p\right)^{1/(1-\rho)} / L_i$ and $\delta_i \equiv L_i T_i^e / \sum_k L_k T_k^e$. A_i is an index for a country's productivity in production and δ_i is a measure of relative size. The equilibrium conditions for this case lead to the following result.

Proposition 1 *Consider a frictionless world under endogenous entry. Assume that, for all i,*

$$1 - (1-\eta)\,\widetilde{\sigma} < \frac{A_i / \left(T_i^e\right)^{\theta/(1-\rho)+1}}{\sum_k \delta_k A_k / \left(T_k^e\right)^{\theta/(1-\rho)+1}} < 1 + \eta\widetilde{\sigma}. \tag{19}$$

The share of labor devoted to innovation in country i is

$$r_i = \frac{L_i^e}{L_i} = \frac{1}{\widetilde{\sigma}} \left(1 - \frac{A_i / \left(T_i^e\right)^{\theta/(1-\rho)+1}}{\sum_k \delta_k A_k / \left(T_k^e\right)^{\theta/(1-\rho)+1}} \right) + \eta. \tag{20}$$

Condition (19) guarantees that innovation shares in (20) are interior, i.e. satisfy $r_i \in (0,1)$, so that no country is completely specialized in innovation or production, i.e., $L_i^e \neq \{0, L_i\}$, for all i.

Proposition 1 illustrates how the different parameters of the model determine whether a country specializes in innovation or production. It tells us that countries with a relatively high ratio of productivity in innovation to production (i.e., countries that have a comparative advantage in innovation) will (partially) specialize in innovation. This high ratio will be reflected in an innovation share higher than the world average, i.e., $r_i > \eta$. Countries with comparative advantage in innovation will also have a trade deficit (i.e., $X_i > Y_i$), as can be seen from equation (16).

2.4.3 A two-country world - the role of home market effects

Under endogenous entry our model exhibits HMEs, according to which the location of innovation and production across countries is affected by country size, as well as trade and MP costs. To illustrate these effects and study the resulting patterns of innovation and production in the simplest way, we consider a world with two countries.

We start with a case where the two countries have different population sizes and study two extreme cases of frictionless trade and frictionless MP, respectively.

Proposition 2 *Consider a two-country world under endogenous entry. Assume that* $A_1 = A_2$, $T_1^e = T_2^e$, *and* $L_1 > L_2$.
i) If there are no trade costs, $\tau_{12} = \tau_{21} = 1$, *and MP costs are symmetric,* $\gamma_{12} = \gamma_{21} = \gamma > 1$, *then in an interior equilibrium* $r_1 > r_2$.
ii) If there are no MP costs, $\gamma_{12} = \gamma_{21} = 1$, *and trade costs are symmetric,* $\tau_{12} = \tau_{21} = \tau > 1$, *then in an interior equilibrium* $r_1 < r_2$.

The first part of the proposition shows the existence of a home market effect (HME) in innovation. Since MP costs are positive but trade is frictionless, it makes sense to innovate in the country with the larger labor force. The opposite result arises if MP is frictionless but trade is costly, in which case the large country specializes in production. We refer to this effect as a quasi-HME, because it runs counter to the logic of the home market effect, whereby the larger country specializes in the activity with increasing returns, which here is innovation. Frictionless MP together with symmetric productivity in innovation imply that in an interior equilibrium wages must be equalized. But given trade frictions, a more than proportionate share of production will take place in the larger market.[12] It is important to note here that the quasi-HME turns out to be much weaker than the HME: our numerical simulations show that in the presence of both trade and MP costs, the large country tends to specialize in innovation. In particular, the large country specializes in innovation whenever $\tau = \gamma > 1$ and only specializes in production if γ is much smaller than τ.

HMEs and quasi-HME also arise due to differences in MP costs or trade costs, even if population sizes are the same. The following proposition illustrates these two effects.

Proposition 3 *Consider a two-country world under endogenous entry. Assume that* $A_1 = A_2$, $T_1^e = T_2^e$, *and* $L_1 = L_2$. *Then*
i) If there are no trade costs, $\tau_{12} = \tau_{21} = 1$, *and MP costs satisfy* $\gamma_{12} < \gamma_{21}$, *then in an interior equilibrium we have* $r_1 > r_2$.
ii) If there are no MP costs, $\gamma_{12} = \gamma_{21} = 1$, *and trade costs satisfy* $\tau_{12} < \tau_{21}$, *then in an interior equilibrium we have* $r_1 < r_2$.

The first result is a reflection of a standard HME: all else equal, it is more profitable to innovate in the country with higher inward MP costs. The second result is again a reflection

[12]The forces leading to quasi-HME are clearly not specific to our model. To see this, recall that the standard model used to illustrate the home-market effect has a differentiated-good sector modeled as in Krugman (1980), and a homogeneous-good sector where production takes place with constant returns to scale and perfect competition. Consider a variation of this model where instead of the homogeneous-good sector we have a sector with a continuum of goods modeled as in Eaton & Kortum (2002) – call the two sectors the Krugman and the EK sectors, respectively. One can show that if trade costs are zero for the Krugman sector and positive for the EK sector then, ceteris paribus, the large country partially specializes in the EK sector – a quasi-HME (see the online appendix).

of a quasi-HME. As above, wages are equalized in an interior equilibrium with frictionless MP and $T_1^e = T_2^e$. This implies that the country with higher inward trade costs attracts more production, pushing innovation to the country with lower inward trade costs.

The results of this section reveal how specialization in innovation or production results from the pattern of comparative advantage and home market effects. Comparative advantage is driven by differences in technological parameters such as T_i^e and T_l^p, while home-market effects are driven by the differences in country size combined with the frictions of moving ideas and goods. In fact, as we will argue in the calibration section below, various combinations of the technological parameters and the trade and MP frictions are able to generate the exact same specialization patterns and trade and MP flows. In this sense, comparative advantage and home market effects are substitutes when it comes to matching the patterns of specialization of trade and MP flows that we observe in the data.

2.5 Welfare Implications

We now turn to the model's implications for how trade and MP affect welfare in each country. We are interested both in a country's overall welfare, as measured by aggregate real income, as well as real wages of workers in innovation and production.

2.5.1 Gains from Trade and MP

We start by considering the overall gains from openness, defined as the change in aggregate real income as we move from a counterfactual equilibrium with no trade and no MP to the observed equilibrium. The gains from openness as a function of equilibrium trade and MP flows (given parameters θ, ρ, σ) are

$$GO_n = \underbrace{\left[\left(\frac{X_{nnn}}{X_n}\right)^{-\frac{1-\rho}{\theta}}\left(\frac{\sum_l X_{nln}}{X_n}\right)^{-\frac{\rho}{\theta}}\right]}_{\text{Direct Effect}}\underbrace{\left[\chi\left(\frac{\frac{1}{\tilde{\sigma}}+\frac{1}{\sigma}-\eta}{\frac{1}{\tilde{\sigma}}\frac{Y_n}{X_n}+\frac{1}{\sigma}-\eta}\right)^{1+\frac{1}{\theta}\frac{\theta-\sigma+1}{(\sigma-1)}}+(1-\chi)\left(1+\theta-\frac{Y_n}{X_n}\right)^{1/\theta}\right]}_{\text{Indirect Effect}}, \tag{21}$$

where χ is an indicator variable that takes the value of 1 under exogenous entry and 0 under endogenous entry (see Appendix B.) With no MP, we can show that the gains from openness (i.e., the gains from trade in this case) are given by $GO_n = \left(\lambda_{nn}^T\right)^{-1/\theta}$, as in Eaton & Kortum (2002), and Arkolakis, Costinot & Rodríguez-Clare (2012). With MP, the gains from openness are composed of a direct and an indirect effect, which we discuss in turn.

The understand the direct effect, consider first the simple case in which $\rho = 0$, under

which the direct effect collapses to $(X_{nnn}/X_n)^{-1/\theta}$. The term X_{nnn}/X_n is an inverse measure of the degree of openness to trade and MP of country n. As one would expect, this measure implies more openness than the typical measure of trade openness, since $X_{nnn}/X_n < \lambda_{nn}^T = \sum_i X_{inn}/X_n$. Turning to the case with $\rho > 0$, note that

$$\left(\frac{X_{nnn}}{X_n}\right)^{-\frac{1-\rho}{\theta}} \left(\frac{\sum_l X_{nln}}{X_n}\right)^{-\frac{\rho}{\theta}} = \left(\frac{\sum_l X_{nln}}{\sum_{i,l} X_{iln}}\right)^{-\frac{1}{\theta}} \left(\frac{X_{nnn}}{\sum_l X_{nln}}\right)^{-\frac{1-\rho}{\theta}}.$$

The first term on the right-hand side captures the gains for country n from being able to consume goods produced with foreign technologies (independently of where production takes place), while the second term captures the gains for country n from being able to use its own technologies abroad and import the goods back for domestic consumption. Given the equilibrium flows X_{iln}, $\rho > 0$ leads to lower gains than $\rho = 0$ since, if productivity draws are correlated, the gains associated with the second term are not as important.

The indirect effect captures the gains or losses triggered by the net inward or outward flow of profits due to MP. Countries with net outward MP flows have a net inward flow of profits and $X_n/Y_n > 1$, implying a positive indirect effect; the opposite occurs in countries with net inward MP flows. The way in which these welfare effects materialize depends on whether we assume exogenous or endogenous entry. Under exogenous entry, a net inflow of profits from MP implies a higher total income and a lower price index thanks to the effect of higher expenditures on the variety of goods available for domestic consumption. Under endogenous entry, the net inflow of profits is associated with higher entry (i.e., higher M_n), which increases welfare by inducing a better selection of varieties in the domestic market.

It is useful to compare our result for gains from openness with those in the perfectly competitive setting of Ramondo & Rodríguez-Clare (2013). Although these authors did not derive it explicitly, one can show that their model leads to an expression for the gains from openness equal to the direct effect in (21).[13] Thus, given trade and MP flows, the difference between the two models is captured entirely by the indirect effect. We can then conclude that our monopolistic competition setup implies larger gains from openness than the perfect competition model of Ramondo & Rodríguez-Clare (2013) for countries with a net outflow of MP, while the opposite is true for countries with a net inflow of MP.

In addition to the gains from openness, we are also interested in the separate welfare effects of trade and MP. The *gains from trade*, GT, are defined as the ratio of real income (X_i/P_i) between the calibrated equilibrium and a counterfactual equilibrium where there is

[13] In Ramondo & Rodríguez-Clare (2013) the parameters θ and ρ play analogous roles as in our model, except that in their case those parameters are associated with a multivariate Frechet distribution rather than a multivariate Pareto distribution.

no trade, computed by letting $\tau_{ln} \to \infty$ for $l \neq n$. Analogously, the *gains from MP*, GMP, are defined as the ratio of real income between the calibrated equilibrium and a counterfactual equilibrium with no MP, computed by letting $\gamma_{il} \to \infty$ for $i \neq l$.

2.5.2 Multinational Production and Real Wages

As mentioned in the Introduction, there is a widespread concern that globalization of production may have a detrimental effect on workers in rich countries. In this subsection we use our model to explore this possibility. In particular, we study the effect of a decline in outward MP costs on the real wage in a country that has a relative abundance of innovation labor and or a high productivity in innovation (under exogenous entry), or a comparative advantage in innovation (under endogenous entry). To make the analysis tractable, we focus on the comparative statics of moving from a situation with frictionless trade but no MP to a situation with both frictionless trade and MP.

For the case of exogenous entry, we assume that $\rho \to 1$. This assumption makes it more likely that MP will hurt workers in rich countries, since the gains from MP arising from differences in firm productivity across countries are not present in this case. By rich countries, in this context, we mean countries that have a relative abundance of high-productivity firms, i.e., a relatively high ratio $m_i \equiv T_i^e L_i^e / L_i^p$, whereas productivity in production is assumed to be the same across countries, $\widetilde{A}_i \equiv \left(T_i^p\right)^{1/(1-\rho)} / L_i^p = \widetilde{A}$ for all i.

Proposition 4 *Consider a world under exogenous entry. Assume that $\widetilde{A}_i = \widetilde{A}$ and $T_i^e = T^e$ for all i, that $\rho \to 1$, and that $m_j = \hat{m}$, for all $j \neq i$, and $m_i = \hat{m} + \varepsilon$, for ε small enough. Consider a switch from frictionless trade but no MP to frictionless trade and MP. This switch*
i) increases real production wages iff $\sigma < \bar{\theta} \equiv \frac{(1+\theta)^2}{1+\theta+\theta^2}$,
ii) increases real wages for innovation workers and aggregate real income, for any value of σ.

By giving firms the ability to locate production in low-wage countries, MP exerts a downward pressure on production wages in rich countries. The same forces lead to an increase in innovation wages and total income, and this increases the variety of goods available for consumption and decreases the price index. If the elasticity of substitution is low enough, this increase in variety will have a large downward effect on the price index, which more than compensates for the decrease in nominal wages, allowing real production wages to increase.

Under endogenous entry, labor can be reallocated to innovation and this leads to a more beneficial effect of MP on real wages.

Proposition 5 *Consider the case of endogenous entry and assume that condition* (19) *holds, so that the equilibrium in a frictionless world is an interior equilibrium. Consider a switch from frictionless trade but no MP to both frictionless trade and MP. This switch increases real wages and real income.*

Comparing Propositions 4 and 5 reveals that the effect of a decline in MP costs critically depends on whether entry is exogenous or endogenous. With endogenous entry, MP openness induces higher demand for labor in the innovation sector in the rich country. This leads to a contraction in employment in the production sector, which results in an improvement in the country's terms of trade (i.e., the relative wage) and an increase in the real wage.

2.5.3 Can Countries Lose from MP?

The strong positive result in Proposition 5 depends on the assumption that trade is frictionless and that all barriers to MP are eliminated. In fact, in our calibrated examples in Section 5 we find that some countries (e.g., Greece) lose from MP, i.e. $GMP < 1$. To understand this possibility, it is useful to start by discussing a simpler result, namely that a country can lose from *unilateral* MP liberalization (i.e., a decline in inward MP costs). Consider a perfectly symmetric two-country world with frictionless trade. In this context, it can be shown that unilateral MP liberalization leads to a decline in innovation and welfare in the liberalizing country (see the online appendix). This result resonates with the well-known result of Venables (1987) that unilateral liberalization can decrease welfare in a Krugman (1980) model with a homogeneous-good sector, but his mechanism is different. The welfare effect in Venables (1987) is caused by the delocation of firms away from the liberalizing country and the resulting increase of its differentiated-goods price index. In contrast, in our model the price index falls in the liberalizing country, but its welfare declines because of a deterioration in its terms of trade caused by the expansion of employment in the production sector.

Can a country lose from *multilateral* MP liberalization? Resorting to numerical examples in the simple case of two countries we find that this is indeed possible.[14] The logic is as above: if MP liberalization triggers home market effects that push innovation in country i below its no-MP level, i.e. $r_i < \eta$, the deterioration of country $i's$ terms of trade may dominate the direct MP gains from the use of foreign ideas, implying loses from MP, $GMP_i < 1$.

The previous discussion may suggest the possibility that a country loses from openness, $GO_i < 1$. But our numerical simulations for two countries never lead to such a result: even if openness leads to a decline in innovation below its autarky value, i.e. $r_i < \eta$, the direct gains from openness always outweigh the indirect loses through a decline in innovation. The key

[14] For the numerical example we set θ, σ and ρ as calibrated in Section 4, together with $\tau_{12} = \tau_{21} = 3$ and $\gamma_{12} = 3$ and $\gamma_{21} = 4$.

insight here is that trade and MP are substitutes in the sense that, if one of these channels is present, then adding the other channel leads to small additional direct gains (see Ramondo & Rodríguez-Clare (2013)) which may not be enough to compensate for the loses arising from the fall in innovation.

2.5.4 Can Countries Lose from Trade?

So far we have focused on the welfare implications of opening up to MP when trade is already feasible. We now consider the opposite case in which countries open up to trade when MP is already feasible. The question is whether opening up to trade can generate a decline in a country's real income, as reflected in $GT_i < 1$. In fact, in the counterfactual analysis with the calibrated model in Section 5 we find that some countries (e.g., Brazil and China) do indeed lose from trade in this sense.

To understand this finding, we turn again to the simple two-country numerical example discussed above, and study the consequences of multilateral trade liberalization in a setting where everything is symmetric except for MP costs.[15] In particular, we assume that inward MP costs are lower than outward MP costs for country 2, i.e. $\gamma_{12} < \gamma_{21}$. This asymmetry implies that multilateral trade liberalization leads country 1 to specialize in innovation and country 2 to specialize in production ($r_1 > \eta > r_2$). For certain parameter values we find that moving from trade autarky to a situation with symmetric trade costs ($\tau_{12} = \tau_{21}$) leads to a decline in the real wage in country 2, again the country that experiences a decline in innovation. Interestingly, the relative wage of country 2 increases as we open up trade, but prices increase even more. Reminiscent of Venables (1987), the delocation of firms from country 2 to country 1 leads to an increase in the price index in country 2 as it must now incur in MP costs for more of the ideas used domestically.

3 Empirical Estimates

In this Section, we use data on production, trade, and multinational sales to provide empirical estimates that are informative for the main two parameters of our theory that we need to calibrate, θ and ρ. Loosely speaking, the value of θ governs the substitutability across products of heterogeneous firms from a given origin, while the value of ρ governs the substitutability across different production locations for a given firm. To infer the value of these

[15] The numerical example we use here again has the θ, σ, and ρ as calibrated in Section 4 and have countries symmetric in all respects except for MP costs. We set $\tau_{12} = \tau_{21} = 3$ while $\gamma_{12} = 3$ and $\gamma_{21} = 4$.

parameters, we consider the estimates of the trade elasticity from two distinct gravity equations.

The first gravity equation is defined over X_{iln}, the sales volumes of the set of firms that originate in country i, produce in country l, and sell in country n. Because this gravity equation is defined over trade flows conducted by firms that originate from a single i, we refer to this equation as "restricted gravity." The second gravity equation is defined over $X_{ln} \equiv \sum_i X_{iln}$, the sales to n from *all* firms operating in country l. Because this gravity equation is defined over trade flows by firms from all countries, we refer to this equation as "unrestricted gravity."

3.1 Restricted Gravity

To estimate the restricted gravity equation, we use expression (7) –see also (B.4) in the Appendix– and take logarithms to obtain

$$\ln X_{iln} = \alpha_{il}^r + \mu_{in}^r - \frac{\theta}{1-\rho} \ln \tau_{ln}, \tag{22}$$

where α_{il}^r and μ_{in}^r are fixed effects.[16] Equation (22) relates sales of firms from i producing in l and selling to n to a production-location and a destination fixed effect as well as to the trade friction between l and n, τ_{ln}, with an elasticity of $-\theta/(1-\rho)$.

A difficulty of operationalizing (22) is that we must have an accurate measure of the relative size of trade frictions between countries l and n. The use of proxies for τ_{ln}, such as distance or shared language, cannot reveal the structural parameters of interest as the coefficient estimate would conflate the trade elasticity with the way in which τ_{ln} varies with the proxy.[17] We instead rely on a measure of the size of trade costs that is directly related to a critical component of τ_{ln}, which is the asymmetric treatment across locations of production in the tariffs applied to goods. Specifically, we operationalize the "restricted" gravity equation

[16] Given i, the fixed effect captured by α_{il}^r varies by location of production and corresponds (in the model) to $\alpha_{il}^r = \ln\left(M_i \left[T_i^e T_l^p (w_l \gamma_{il})^{-\theta}\right]^{\frac{1}{1-\rho}}\right)$, while the fixed effect captured by μ_{in}^r varies by country of destination and corresponds (in the model) to $\mu_{in}^r = \ln\left(X_n \Psi_{in}^{\frac{-\rho}{1-\rho}} / \sum_k M_k \Psi_{kn}\right)$.

[17] This feature of the gravity model of trade led Eaton & Kortum (2002), Donaldson (2014), and Simonovska & Waugh (2013) to use price gaps of homogeneous goods between locations to back out measures of τ_{ln}. In our monopolistically competitive model we cannot use these price variations for that same purpose so we need to resort to different ways of measuring τ_{ln}.

by parameterizing trade costs so that

$$\ln X_{iln} = \alpha_l^r + \mu_n^r + \beta^r \ln(1 + t_{ln}) + \sum_k \delta_k^r [1|d_{ln} \in d_k] + \Theta^r H_{ln} + \varepsilon_{ln}, \tag{23}$$

where t_{ln} is the simple average tariff applied by n on goods from l, $[1|d_{ln} \in d_k]$ is an indicator variable for distance between n and l –whose marginal effect on trade costs is given by δ_k^r–, and H_{ln} is a vector of standard gravity controls, including a shared language, a shared border, and an indicator variable, called *self*, that is equal to one if $l = n$. To the extent that constructed measures of t_{ln} accurately capture variation in asymmetric trade frictions between countries, the coefficient estimate $\hat{\beta}^r$ has the structural interpretation of the parameter ratio $\theta/(1-\rho)$. The coefficients on the other, more standard, proxies for trade costs such as the distance indicator variables, do not have a direct structural interpretation as they are a mixture of $-\theta/(1-\rho)$ and the effect of the variable on the size of trade costs. Because in our data there are multiple observations for each production location l and for each destination country n, we can estimate (23) via ordinary least squares (OLS) with dummy variables.

We use data on the operation of U.S. manufacturing firms across multiple locations constructed from the 1999 benchmark survey of the Bureau of Economic Analysis (BEA) on the operations of U.S. multinationals abroad. For each country l, we observe sales of U.S. multinationals in their host country and their exports to the United States, Canada, Japan, the United Kingdom, and a composite of fourteen European Union countries. We also observe the domestic sales of U.S. firms in the United States (net of the sales of foreign affiliates in the United States) and their exports to each country in the data set. Details about the construction of the data can be found in the Appendix A.

In our sample of the global operations of U.S. multinationals, there are two forms of variation in t_{ln} that identify β^r. The first type of variation is due to the fact that firms that open a local affiliate avoid all trade costs (i.e. $t_{nn} = 0$), while firms from another country generally must pay the applied MFN tariff rate. A second source of variation in t_{ln} is due to the fact that some l and n belong to common preferential trade agreements (and so $t_{ln} = 0$), while others do not (so that exports from l pay country n's MFN applied tariff rates).[18]

There are several concerns that arise in using tariff data to estimate the trade elasticity. First, there is the problem of endogeneity since country pairs for which there is a natural affinity for trade are more likely to agree to preferential trading arrangements. The standard gravity controls in (23) proxy for this affinity. To the extent that there are other determinants of preferential trading agreements that are excluded from (23), the trade elasticity may be

[18]There is also some variation in constructed tariff measures due to the fact that developed countries extend Generalized System of Preference tariffs to a number of developing countries.

biased upward. A second potential problem arises because the model does not suggest an appropriate way to aggregate tariffs across industries. We have chosen a simple average of applied tariffs because other aggregation schemes are either ad hoc or have an element of endogeneity to them.[19] If t_{ln} is seriously mismeasured, our estimate of β^r is biased downward. We include *self* to control for the variation in t_{ln} that is due to unmeasured trade costs, such as administrative and information frictions, that local production avoids.

3.2 Unrestricted Gravity

The "unrestricted" gravity equation has the same form as the "restricted" gravity equation, but it is estimated on the bilateral sales of *all* firms located in country l selling to country n. Specifically, we estimate

$$\ln X_{ln} = \alpha_l^u + \mu_n^u + \beta^u \ln(1 + t_{ln}) + \sum_k \delta_k^u [1|d_{ln} \in d_k] + \Theta^u H_{ln} + v_{iln}. \tag{24}$$

The coefficient estimate $\hat{\beta}^u$ does not have a structural interpretation, but it still provides information on the relative magnitudes of θ and ρ. To see this, recall that if MP is not possible then all exports are done by local firms, the correlation of the firm productivity shocks determined by ρ is irrelevant, and the coefficient on tariffs is equal to θ (see equation 17). In the data most exports are done by domestic firms so that X_{ln} disproportionately contains information on the operations of domestic firms. This suggests that $\hat{\beta}^u$ is closer to θ than $\hat{\beta}^r$, which is equal to $\theta/(1-\rho)$. In summary, the model implies the following restriction on parameters: $\hat{\beta}^r = -\theta/(1-\rho) < \hat{\beta}^u < -\theta < 0$.

We estimate (24) using data on trade volumes of manufacturing industries and domestic absorption again for 1999. To ensure comparability between the coefficients, we restrict the sample so that the country pair coverage in the restricted and unrestricted samples is the same.

3.3 Results

The coefficient estimates for the two OLS regressions are reported in Table 1 (with robust standard errors in parenthesis). Each column of the Table corresponds to a dependent variable, while the first and second rows correspond to the restricted and unrestricted specifications, respectively. Of most relevance to our analysis are the elasticity estimates for *tariff* shown in the first column. Note that the underlying data in both specifications have 317 ob-

[19] We report robustness with an alternative weighting scheme in the online Appendix.

	Tariff	Distance Dummies					Other Gravity Controls			R-sq.
		D2	D3	D4	D5	D6	Self	Border	Lang.	
Restricted	-10.9	-0.4	-0.4	-1.3	-2.9	-2.7	2.1	0.2	0.7	0.84
	(-3.5)	(0.5)	(0.6)	(0.6)	(0.5)	(0.6)	(0.7)	(0.6)	(0.3)	
Unrestricted	-4.3	-0.7	-0.7	-0.7	-1.9	-1.7	3.3	1.0	0.0	0.89
	(1.8)	(0.3)	(0.3)	(0.3)	(0.3)	(0.3)	(0.4)	(0.3)	(0.1)	

Table 1: Restricted and Unrestricted Gravity: OLS Estimates

Notes: Robust standard errors in parenthesis.

servations. The trade elasticity in the restricted regression of 10.9 is our estimate of $\theta/(1-\rho)$. The trade elasticity in the unrestricted regression is significantly smaller at 4.3, as expected.[20]

The coefficient estimate of 4.3 obtained from the unrestricted gravity estimation is in the ballpark of estimates of the trade elasticity obtained by the trade literature. Anderson & Van Wincoop (2004) reports that estimates of the trade elasticity using price differentials, or tariffs, across countries range from five to ten. More recent estimates using tariffs, such as Romalis (2007) and Caliendo & Parro (2014), fall within the same range. Simonovska & Waugh (2013), using price differentials, and a refinement of the methodology of Eaton & Kortum (2002), bring these estimates closer to 4.

4 Calibration

We restrict our analysis to 26 countries for which we have good data for both trade, output and MP.[21] For trade we use World Input-Output Database (WIOD) data on manufacturing trade flows from any country l to country n, including home sales, as the empirical counterpart of bilateral trade in the model.[22] Using this information, we construct the $N \times N$ matrix of trade shares, λ_{ln}^T, and the $N \times 1$ vector of aggregate expenditures, X_n. In addition, we use data from Ramondo, Rodriguez-Clare & Tintelnot (2014) on the gross value of production for multinational affiliates from country i in country l to construct the empirical counter-

[20] OLS estimates maybe biased because of heteroscedasticity (see Silva & Tenreyro (2006)). To study the effect of this bias, we re-estimate both gravity models using a Poisson pseudo-maximum-likelihood estimator. We obtain the alternative estimates of 8.4 for the restricted regression and 5.4 for the unrestricted regression. The results for our estimation and the counterfactual experiments that we perform below are very similar if we use these estimates.

[21] These countries are Australia, Austria, Benelux (Belgium, Luxembourg, and Netherlands aggregated as one country), Brazil, Canada, China, Cyprus, Denmark, Spain, Finland, France, United Kingdom, Germany, Greece, Hungary, Ireland, Italy, Japan, South Korea, Mexico, Poland, Portugal, Sweden, Turkey, and the United States.

[22] See Timmer (2012) and Costinot & Rodriguez-Clare (2014) for the description of the WIOD database.

part of bilateral MP flows and obtain an $N \times N$ matrix of production shares, λ_{il}^{M}.[23] Finally, we measure the $N \times 1$ vector of labor endowments, L_i, as equipped labor, from Klenow & Rodríguez-Clare (2005), multiplied by the share of employment in the manufacturing sector, from UNIDO; this is also the variable we refer to as country size. All the data refer to an average over 1996-2001.

We set the value of $\theta/(1-\rho) = 10.9$ to match the restricted gravity elasticity shown in Table 1. To disentangle ρ from θ, we use the predictions of the model regarding the unrestricted gravity regression coefficient. Computing that same elasticity in the model, however, requires calculating the model's equilibrium to generate a dataset of trade flows and trade costs, a procedure that we describe below.

We also need a value for the parameter σ, which determines the markup in the model, $\widetilde{\sigma}$. Estimates for the average mark-up for manufacturing across OECD countries by Martins, Scarpetta & Pilat (1996) are around 20 percent, while estimates by Domowitz, Hubbard & Petersen (1988) for the U.S. manufacturing sector bring this number closer to 37 percent. Since we need to satisfy the restriction $\theta > \sigma - 1$, we set $\sigma = 4$, which leads to markups of 33 percent, closer to the high end of the empirical estimates.[24]

The calibration of the rest of the parameters proceeds in three steps. The first step involves computing τ's and γ's by using the trade and MP shares from the data. Our procedure is an extension of the approach in Head & Ries (2001) to a setting with MP. Head & Ries (2001) show that in a gravity model of trade, if one assumes that trade costs are symmetric ($\tau_{nl} = \tau_{ln}$) and that $\tau_{nn} = 1$ for all n then trade costs can be obtained as $\widehat{\tau}_{ln} = \left(\frac{\lambda_{ln}^{T}\lambda_{nl}^{T}}{\lambda_{ll}^{T}\lambda_{nn}^{T}}\right)^{-1/2\theta}$. In our model, if $\rho = 0$ then we could follow the exact same strategy to estimate τ_{ln}, and analogously we could estimate MP costs as $\widehat{\gamma}_{il} = \left(\frac{\lambda_{il}^{M}\lambda_{li}^{M}}{\lambda_{ii}^{M}\lambda_{ll}^{M}}\right)^{-1/2\theta}$ (assuming $\gamma_{ii} = 1$ for all i). If $\rho > 0$ then these formulas are no longer valid, but we can still apply the same basic idea. In Appendix B we show that given the data on bilateral trade and MP flows, our model determines all trilateral flows, X_{iln}, which in turn, imposing symmetry on trade and MP costs and $\tau_{ll} = \gamma_{ll} = 1$ for all l, can be used to construct $\widehat{\tau}_{ln}$ and $\widehat{\gamma}_{il}$.

In the second step we calibrate the parameter values for T_l^p and T_i^e. We normalize $T_i^e =$

[23] Since our quantitative analysis is restricted to the manufacturing sector, while the MP data from UNCTAD includes all MP flows, we rely on the following approximation. We observe that, for the United States, MP flows in manufacturing account for approximately one half of overall MP flows, while manufacturing gross output is approximately one half of overall GDP. Thus, we take overall manufacturing MP flows divided by GDP (taken from the World Development Indicators) as an approximation of manufacturing MP flows as a share of gross production in manufacturing.

[24] Given our calibration, the profit share, $\eta = 1/(\tilde{\sigma}\theta)$, is 18.75% (15.15% for $\theta = 4.95$). The profit share equals the share of payments to workers in the innovation sector. This number is close to the 15-percent share of income accrued to intangible capital in the United States, a number calculated by Corrado, Hulten & Sichel (2009).

$T_i^p = 1$ for $i = USA$ and we set $T_i^p = (L_i)^{1-\rho}$ for all i so that deviations of T_i^e from one give us the strength of a country's comparative advantage in innovation relative to the US (see Proposition 1). We then pick T_i^e so that the model exactly matches the innovation shares, r_i, in (16), calculated using trade, MP, and absorption data. The calibrated T_i^p, T_i^e and r_i's are illustrated in Table 2. Crucially, our calibration procedure allows us to disentangle T_i^e from γ_{il} and, thus, comparative advantage from HMEs, as we further explain below.[25]

The final step of the algorithm requires estimating an unrestricted gravity regression in which the dependent variable is the model-generated trade share from l to n and the regressors are the calibrated trade costs from l to n, and exporting and importing fixed effects. For $\theta = 4$ and $\rho = 0.63$, which satisfy $\theta/(1-\rho) = 10.9$ (our target from Section 3), the model predicts an OLS unrestricted trade elasticity of 4.3. To illustrate the role of ρ in determining the export behavior of multinationals we perform an additional calibration where we set $\rho = 0$ and redo steps 1 and 2 (step 3 is trivial because for $\rho = 0$, the trade elasticity from the unrestricted gravity equation is equal to θ). The fit for trade and MP data is equally good under this specification, except of course that it does not line up with the trade elasticity in the restricted gravity equation as estimated in Section 3.

We assess the fit of the model in Figures (5) and (6), which plot aggregate and bilateral trade and MP. Both aggregate inward and outward shares for trade and MP as well as bilateral shares were not directly targeted by our calibration procedure. Figure 6 reveals that the model delivers slightly higher MP shares than the ones observed the data, while the opposite is true for trade shares.

4.1 Innovation, Comparative Advantage and HMEs

We next discuss the role of comparative advantage and HMEs in explaining the variation in innovation rates across countries. To isolate the role of HME we use the calibrated model and shut down the comparative advantage force by setting $T_i^e = 1$ for all i. The results are shown in column 4 of Table 2—column 3 shows the baseline values for r. The results reveal a significant role for comparative advantage in explaining country innovation. The calibrated T_i^e and r_i in the model exhibit a positive correlation of 0.24, indicating that part of the innovation potential of a country is explained by a (relatively) higher productivity in the innovation sector. A country's r depends on its own T^e and also that of other countries, especially the

[25] An alternative calibration method can be found in Arkolakis, Ramondo, Rodríguez-Clare & Yeaple (2013) (Burstein & Vogel (2012) and Allen, Arkolakis & Takahashi (2014) develop similar procedures for trade models). In that calibration we show that the model can fit exactly the trade and MP data if symmetry on trade and MP costs is not assumed, exploiting the fact that the number of parameters used is at least as large as the bilateral trade and MP shares. However, that algorithm cannot disentangle comparative advantage and HMEs since T_i^e's cannot be identified separately from MP costs.

	T^p	T^e	r	
			baseline T^e	$T^e = 1$
	(1)	(2)	(3)	(4)
AUS	0.34	2.20	0.15	0.07
AUT	0.27	2.05	0.16	0.01
BNL	0.39	1.22	0.24	0.17
BRA	0.54	0.75	0.16	0.17
CAN	0.42	1.26	0.15	0.03
CHN	1.16	0.19	0.16	0.19
CYP	0.10	0.05	0.18	0.19
DNK	0.25	3.12	0.21	0.07
ESP	0.40	1.49	0.15	0.13
FIN	0.23	3.58	0.23	0.06
FRA	0.53	1.54	0.20	0.13
GBR	0.54	1.36	0.18	0.12
GER	0.70	1.07	0.20	0.22
GRC	0.20	2.41	0.17	0.14
HUN	0.29	1.06	0.05	0.06
IRL	0.19	1.85	0.10	0.00
ITA	0.46	1.68	0.17	0.13
JPN	0.79	1.49	0.20	0.17
KOR	0.55	1.52	0.18	0.15
MEX	0.62	0.75	0.14	0.16
POL	0.47	0.51	0.13	0.17
PRT	0.29	1.36	0.10	0.09
ROM	0.37	0.38	0.17	0.19
SWE	0.28	2.31	0.19	0.06
TUR	0.44	0.88	0.17	0.18
USA	1.00	1.00	0.23	0.24

Table 2: Comparative Advantage vs Home-Market Effects

Note: Column 1 shows $T_i^p = L_i^{1-\rho}$, where L_i is from data on equipped labor, while column 2 shows the calibrated T^e. Columns 3 and 4 show innovation shares (r) coming from equilibria with calibrated T^e's and $T_i^e = 1$, for all i, respectively.

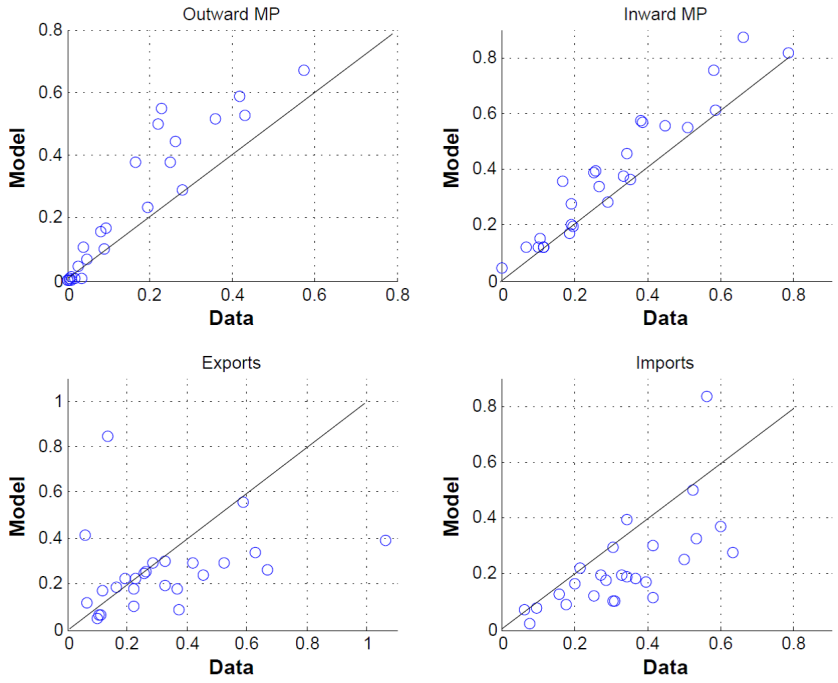


Figure 5: Average outward and inward trade and MP shares, model and data.
Note: Imports and exports for country n are normalized by absorption in country n. Outward (inward) MP shares refer total sales of foreign affiliates from (into) country n normalized by gross production in country n.

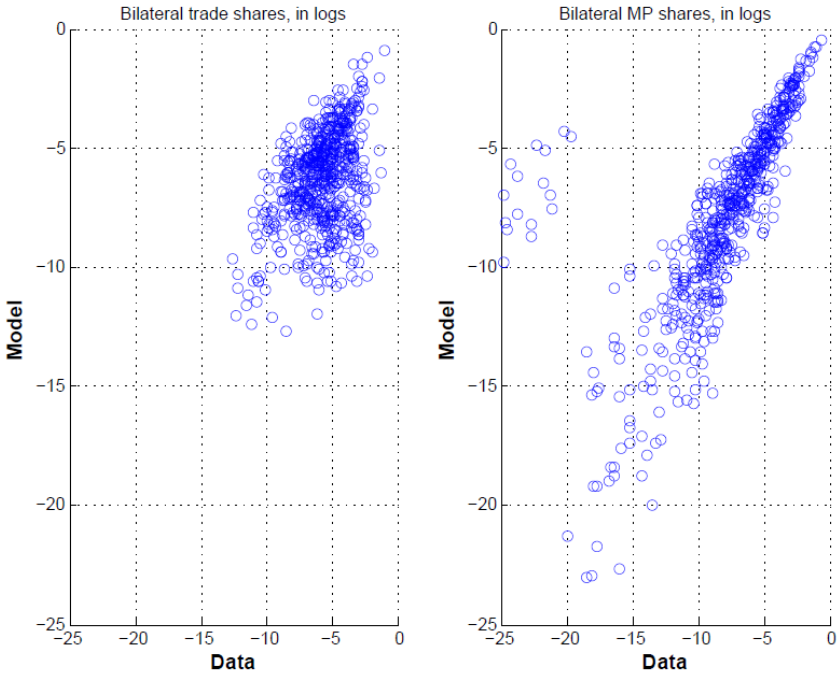


Figure 6: Bilateral trade and MP shares, model and data.
Note: Trade flows are normalized by absorption. MP shares are normalized by gross production.

ones with whom it has stronger connections (as implied by the trade and MP costs). Countries with a high T^e, such as Denmark and Finland, suffer a marked decline in r when T^e becomes 1 everywhere (from more than 20% to less than 10%), implying that the specialization in innovation in these countries is driven by comparative advantage. On the contrary, Germany and the United States, still experience an increase in r, implying that specialization in innovation in these countries is driven to a large extent by HMEs. Another informative example is Canada where despite the small decline in T^e (from 1.26 to 1) innovation drastically falls from 15% to 3%. This large change suggests that, as a result of its strong exposure to the US market and its relatively small size, HMEs work strongly towards specialization in production for Canada.

4.2 Additional Implications

The calibrated model has additional implications that we now contrast with data not used in the calibration algorithm.

R&D employment shares. Figure 7 shows the innovation share in the model and in the data relative to the United States. The innovation share in the model is r_i from Table **2,** while in the data it corresponds to R&D expenditures relative to local value-added for the manufacturing sector (as in Figure 1). There is a strong positive association between the two variables, in spite of the fact that R&D data were not used in the calibration of the model. This positive association suggests that the model does a good job in capturing the observed relationship between trade, MP and innovation.[26]

Trade and MP costs and distance. The estimated trade and MP costs should correlate to geographic variables such as bilateral distance. To evaluate this relationship we regress the logarithm of estimated trade and MP costs on origin and destination fixed effects, the indicator variable for distance, as in the gravity regressions above, and border and language dummies. The coefficients are all significant at the one-percent level and have the expected signs. In particular, the results imply that trade and MP costs increase with distance and are lower for country pairs that share a border or a language.

Bridge MP. As discussed above, our calibration procedure implies a unique mapping from observed bilateral trade and MP shares to simulated trilateral flows, X_{iln}. As noted

[26]The two variables are quite different in levels: The observed share of labor employed in R&D is an order of magnitude lower than the model's implied share, which revolves around 17%. There are two explanations for this. First, the R&D data are for the whole economy, whereas R&D in manufacturing (which is what we focus on) is higher. For example, the ratio of R&D to value added in the United States is 2.7% in the whole economy and 8.7% in manufacturing. Second, R&D in the data captures only a small part of what constitutes innovation in the model. A more interesting feature is whether the model captures the cross-country pattern observed for this variable, which is what we focus on.

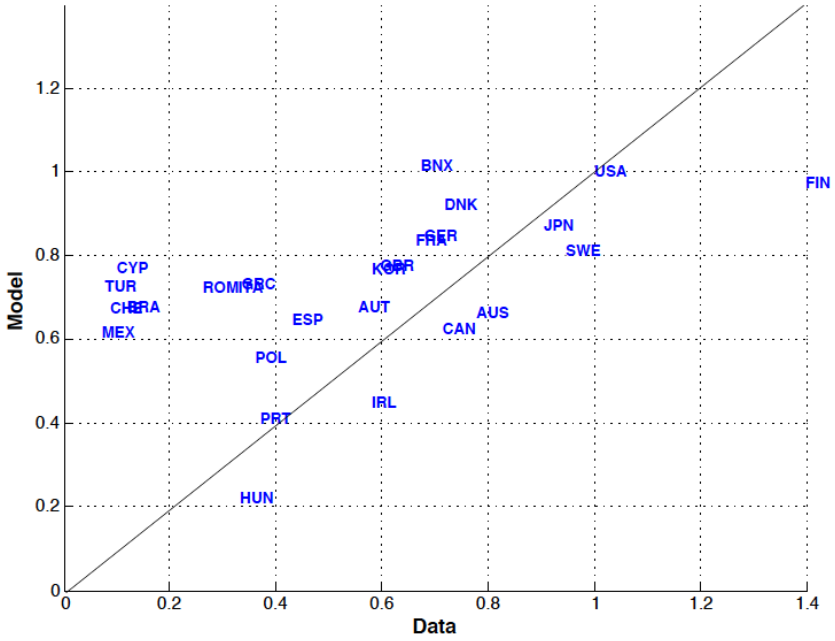


Figure 7: Employment Shares in R&D: Model and Data.

Note: Employment shares in R&D are: in the model, the (equilibrium) variable r; in the data, employment in R&D, as a share of total employment, from the World Development Indicators, an average over the nineties.

above, the model does a good job in matching the aggregate trade and MP flows. We now assess the ability of our model to predict these trilateral flows for the case of $i = US$, which is the only case for which we have the necessary data. The analysis in this section motivates the alternative calibration with $\rho = 0$.[27] With that calibration the export behavior of foreign affiliates and domestic firms is the same and it is easy to show that

$$X_{iln}|_{\rho=0} = \lambda_{il}^{M} \lambda_{ln}^{T} X_n. \tag{25}$$

This expression reveals that when $\rho = 0$ the cross country structure of the sales of affiliates from country i producing in country l is the same for the sales of all firms operating in country l. In the model deviations from this benchmark are driven by differences in the restricted and unrestricted trade elasticities.

Following Ramondo & Rodríguez-Clare (2013), we refer to MP sales sold outside of the local market as "bridge" MP (BMP) flows, since firms from i use l as a bridge to reach another location n. In turn, BMP shares are defined as the ratio of total BMP flows to total MP flows

[27] As explained in Section 3.1, our data for trilateral flows come from the BEA and are restricted to X_{iln} for $i = US$ and n equal to the top five markets for US sales (Canada, European Union, Great Britain, Japan, and the United States). We emphasize that these data were used in the calibration of the model only to pin down one parameter, ρ, but they were not used in the calibration of trade and MP costs.

($\sum_{n \neq l} X_{iln} / \sum_n X_{iln}$). The BMP shares predicted by the model are lower than those in the data: averaging across all production locations in our data set, the average BMP share in the model is 0.03 while in the data this is 0.39. This behavior implies that when $\rho > 0$ multinationals tend to serve foreign markets mostly through exports or MP and less through BMP. Indeed, when we switch to the alternative calibration with $\rho = 0$, BMP shares increase dramatically to 0.30, which is very close to what we see in the data. Thus, even without fixed costs of setting up foreign affiliates, our model can generate the levels of BMP observed in the data, but doing so comes at the cost of not matching the trade elasticities estimated for the restricted and unrestricted gravity equations in Section 3.

Another criteria to assess the predictions of the model with respect to BMP is the pattern of BMP flows across destinations. To assess this fit, we adjust actual and predicted BMP flows for country sizes (i.e. gravity), calculating the normalized BMP flows as $X_{iln} / (Y_l X_n)$. The correlation between normalized BMP flows in the model and the data is 0.87 for the baseline calibration and 0.80 with $\rho = 0$. We conclude that the model performs well in matching the pattern of normalized BMP flows and that calibrating the model to the differences in trade elasticities between unrestricted and restricted gravity improves this fit.

5 Counterfactual Experiments

Armed with our calibrated model, we perform a series of counterfactual experiments to understand the effects of openness on innovation and welfare across and within countries. We first calculate the gains from openness as well as the gains from trade and the gains from MP, according to the definitions in Section 2.5. Next, we compute the effect of a decline in MP costs on innovation, real income, and real wages of workers in the innovation and production sectors. Finally, we analyze the effects of the integration to the world economy of China, a country with a large labor force, motivated by the emergence of this country as a key location for manufacturing production over the last two decades.[28]

For these counterfactual exercises, in addition to the two versions of the model presented in Section 2, we consider a version of the model with endogenous entry but heterogeneous workers as in Roy (1951) (see Appendix B for details). This version of our model allows us to address questions concerning the effect of globalization on innovation, aggregate welfare, and the distribution of income between innovators and workers in a single framework. A critical parameter, which we label κ, captures the extent to which workers differ in their

[28]The reader should keep in mind that our model was calibrated to manufacturing, hence the welfare implications correspond to changes in manufacturing consumption only.

relative productivities in the two activities.[29] The model of Section 2 with endogenous entry and homogeneous workers corresponds to a special case of this model with $\kappa \longrightarrow \infty$. We have some guidance from recent quantitative work for the value of this parameter, such as Lagakos & Waugh (2013), and Hsieh, Hurst, Jones & Klenow (2013). Based on those papers, we set $\kappa = 3$.

In the rest of this Section we refer to the three versions of the model as the linear model (endogenous entry with homogeneous labor, $\kappa \to \infty$), the Roy model (endogenous entry with heterogeneous labor, $\kappa = 3$) and the model with exogenous entry. The calibration of each of these versions of the model is done exactly as described in Section 4: given that the three versions of the model differ only in the specification of the labor market, the implied trade and MP costs are the same across specifications – see the Online Appendix for details.

5.1 The Gains from Openness

Table 3 presents the innovation rate r in the baseline equilibrium as well as the gains from openness decomposed into the direct and indirect effects, as discussed in Section 2.5.[30] A small and open economy like Benelux derives enormous gains of moving from isolation to the baseline equilibrium. These gains primarily stem from the direct effect, but being specialized in innovation (as reflected in its high r), part of these gains come from a positive indirect effect. Different is the case of Mexico, which specializes in production (as reflected in its low r) and has a negative indirect effect. Germany, Japan and the United States exhibit a pattern like that in Benelux, namely specialization in innovation ($r > \eta$) and a positive indirect effect, whereas Canada, Spain and Portugal are examples of countries that, like Mexico, are specialized in production ($r < \eta$) and have a negative indirect effect.

In Table 3 we also present the gains from trade and the gains from MP. These welfare changes tend to be low (relative to the gains from openness), because trade and MP are substitutes: once an economy has access to either trade or MP, then adding the other channel does not generate large additional gains (see Ramondo & Rodríguez-Clare (2013)). Can trade or MP lead to welfare losses? Our model yields an affirmative answer. Brazil, China, Hungary, Poland and Portugal lose from trade, while Greece and Romania lose from MP (although these loses are very small). This is in contrast to Ramondo & Rodríguez-Clare (2013), where gains from trade and gains from MP were always positive. The possibility of losses in our model is explained in Section **2.5.** If calibrated inward MP costs are low relative to cal-

[29] Formally, we assume that workers independently draw their productivity in innovation and production from a Fréchet distribution with shape parameter κ.

[30] The results in Table 3 correspond to the linear model – results for the Roy and exogenous entry versions of the model are very similar and are therefore not reported.

	r	GO			GT	GMP
		overall	direct	indirect		
AUS	0.15	1.14	1.20	0.95	1.01	1.06
AUT	0.16	1.48	1.54	0.96	1.05	1.15
BNL	0.24	1.83	1.73	1.06	1.10	1.34
BRA	0.16	1.04	1.09	0.96	0.987	1.003
CAN	0.15	1.66	1.77	0.94	1.11	1.08
CHN	0.16	1.02	1.07	0.96	0.974	0.9997
CYP	0.18	1.71	1.73	0.99	1.31	1.02
DNK	0.21	1.41	1.37	1.03	1.10	1.09
ESP	0.15	1.15	1.21	0.95	1.003	1.03
FIN	0.23	1.38	1.31	1.05	1.08	1.11
FRA	0.20	1.25	1.24	1.01	1.05	1.09
GBR	0.18	1.35	1.36	0.99	1.05	1.17
GER	0.20	1.23	1.21	1.01	1.03	1.13
GRC	0.17	1.17	1.20	0.98	1.07	0.999
HUN	0.05	1.52	2.10	0.72	0.999	1.18
IRL	0.10	2.25	2.60	0.86	1.08	1.32
ITA	0.17	1.16	1.19	0.97	1.03	1.02
JPN	0.20	1.08	1.06	1.02	1.04	1.04
KOR	0.18	1.06	1.07	0.99	1.01	1.02
MEX	0.14	1.11	1.19	0.94	1.004	1.001
POL	0.13	1.08	1.18	0.91	0.966	1.005
PRT	0.10	1.32	1.56	0.84	0.986	1.07
ROM	0.17	1.07	1.10	0.97	1.01	0.997
SWE	0.19	1.53	1.53	1.00	1.08	1.16
TUR	0.17	1.06	1.09	0.98	1.01	0.993
USA	0.23	1.14	1.08	1.06	1.06	1.08
Avg	0.17	1.32	1.38	0.97	1.05	1.08

Table 3: Gains from Openness, Trade, and MP.

Note: The gains from openness refer to the gains of moving from isolation to the calibrated equilibrium. The gains from trade (MP) refer to the gains of moving from an equilibrium with only MP (trade) to the calibrated equilibrium with both trade and MP. "Direct" and "indirect" refer to the first and second terms, respectively, on the right-hand side of (21). The overall effect is the full right-hand side of (21).

ibrated outward MP costs, then lowering trade costs from infinity to their calibrated values (while leaving MP costs at their calibrated levels) leads to a reallocation of resources from innovation to production, a de-location effect that may increase prices as more ideas used in production must now bear the MP costs. On the other hand, lowering MP costs from infinity to their calibrated values (while leaving trade costs at their calibrated levels) also leads to a decline in innovation, triggering a deterioration in the country's terms of trade. As shown in the first column of Table 3, equilibrium innovation in all the countries that lose from MP or trade is lower than its equilibrium level in the counterfactual with no MP or with no trade, i.e., $r_i < \eta$.

We now quantify the gains from lowering all MP barriers by five percent. Columns 1-2, 3-4 and 7-8 of Table 4 show the percentage change in r_i and X_i/P_i for the three different models. Columns 5-6 and 9-10 present the percentage change in the real wage in production (w_i/P_i) and innovation (w_i^e/P_i) in the Roy and exogenous-entry models.

MP liberalization generates a large reallocation of worldwide innovation efforts towards the countries that have the highest innovation rates in the baseline equilibrium. This is especially noticeable in Benelux, where r increases by 41% in the linear model, 13% in the Roy model and 6% in the exogenous-entry model. The strong increase in innovation in Benelux is caused by the non-linear response of MP flows to changes in MP costs – a five-percent decline in γ_{il} when it is already low has a stronger effect than the same percentage decline in γ_{il} when it is high. The effect is much smaller in the Roy model because of the increase in the cost of innovation that results as innovation expands and absorbs workers with lower relative productivity in that activity. In the exogenous-entry model the increase in r is even smaller – since there is no labor reallocation in this case, the increase in r is purely a reflection of higher wages in innovation relative to production.

Turning to the welfare implications of MP liberalization, we see that only Romania and Turkey experience overall loses, and they are very small and occur only in the linear model. As explained in Section 2.5, aggregate loses arise because of the reallocation of resources from innovation to production and the associated terms of trade or delocation effects. Since this reallocation is smaller, or non-existent, in the Roy and exogenous-entry models, loses are less likely.

What are the real wage implications for workers in production and innovation? The results in Table 4 show that real wages of production workers increase with MP liberalization in all the countries in our sample for both the exogenous entry and the Roy model. Thus, despite the possibility raised in Proposition 4 that, under exogenous entry, real production wages could fall with MP liberalization, we see that the overall efficiency gains always dom-

inate any negative effects coming from a decline in relative production wages. We see a very different picture when we look at the changes in real wages for innovation workers. These wages tend to fall with MP liberalization in the countries which are net recipients of MP – for example, real innovation wages decrease by 2.5% in Mexico and by around 1% in Brazil and China according to the Roy model and even more according to the exogenous-entry model. More broadly, we see that real innovation wages tend to increase by more than real production wages in the countries exhibiting net outflows of MP, whereas the opposite happens in countries exhibiting net inflows of MP.

	Linear Model		Roy Model				Exogenous-Entry Model			
% change in:	r	X/P	r	X/P	w/P	w^e/P	r	X/P	w/P	w^e/P
AUS	-6.30	1.78	-3.48	1.85	2.08	0.57	-1.81	1.90	2.25	0.06
AUT	-1.88	3.56	-0.30	3.53	3.57	3.35	0.55	3.46	3.35	4.03
BNL	40.73	6.40	13.45	6.09	4.60	10.62	6.07	5.66	3.76	12.08
BRA	-5.57	0.23	-4.57	0.31	0.60	-1.26	-3.59	0.46	1.14	-3.15
CAN	-19.64	3.32	-4.47	3.50	3.76	1.93	-0.42	3.44	3.51	3.00
CHN	-5.33	0.25	-4.64	0.29	0.58	-1.29	-3.76	0.42	1.13	-3.36
CYP	-2.59	0.03	-1.98	0.37	0.51	-0.30	-1.58	0.51	0.85	-1.08
DNK	3.25	2.78	3.05	2.78	2.52	3.71	1.91	2.62	2.09	4.59
ESP	-8.90	1.07	-6.06	1.22	1.60	-0.94	-3.95	1.37	2.09	-2.63
FIN	4.01	3.36	3.94	3.26	2.90	4.49	2.46	3.06	2.33	5.60
FRA	0.03	2.76	1.71	2.71	2.58	3.26	1.51	2.59	2.21	4.14
GBR	1.71	3.87	3.80	3.93	3.64	5.20	2.88	3.81	3.14	6.79
GER	-1.15	3.17	1.31	2.98	2.87	3.41	1.27	2.81	2.50	4.12
GRC	-5.28	0.10	-3.15	0.13	0.40	-1.18	-3.12	0.25	0.89	-2.88
HUN	-46.81	4.57	-16.42	4.44	4.76	-1.78	-6.82	4.20	4.61	-2.91
IRL	71.76	6.29	13.74	5.47	4.91	10.04	6.56	5.18	4.29	12.08
ITA	-5.95	1.12	-3.39	1.18	1.43	-0.05	-1.88	1.26	1.65	-0.64
JPN	1.12	1.09	1.60	1.11	0.98	1.60	1.31	1.04	0.70	2.36
KOR	-1.85	0.60	-1.08	0.63	0.73	0.19	-0.85	0.66	0.85	-0.19
MEX	-11.71	0.57	-9.17	0.69	1.20	-2.49	-6.43	0.88	1.98	-5.61
POL	-13.94	0.88	-10.15	0.93	1.44	-2.63	-7.00	1.13	2.20	-5.95
PRT	-16.95	3.37	-10.76	3.20	3.61	-0.78	-6.24	3.13	3.84	-3.30
ROM	-4.70	-0.04	-4.50	0.04	0.35	-1.49	-3.79	0.18	0.95	-3.62
SWE	1.95	4.20	4.36	4.22	3.88	5.65	3.01	4.05	3.31	7.19
TUR	-4.51	-0.05	-3.57	0.04	0.30	-1.23	-3.06	0.17	0.79	-2.90
USA	3.87	1.76	2.22	1.65	1.38	2.54	1.73	1.48	0.96	3.24
average	-1.33	2.19	-1.48	2.18	2.20	1.58	-0.96	2.14	2.21	1.19

Table 4: MP Liberalization: Linear, Roy, and Exogenous-Entry Model.
Note: MP liberalization refers to a five-percent decrease in all MP costs. w^e is the wage per efficiency unit in the innovation sector. For the linear model, changes in X/P are equivalent to changes in w/P.

We explored how the results presented in Tables 3 and 4 change under our alternative

calibration with $\rho = 0$. The average country gains 21% versus 30% in our baseline calibration and as expected the difference is almost entirely due to the direct effect. Moreover, the gains from MP are positive for all countries, except Cyprus, and no country loses from opening up to trade. While the gains from MP are larger, on average, in our baseline calibration (8.2% vs 6.2%), the average gains from trade are more than twice as large for $\rho = 0$ (10.5% vs 4.6%). This is not surprising, however, given that the elasticity elasticity $(1-\rho)/\theta$ when $\rho = 0$ is $1/4$ vs $1/10.9$ in our baseline calibration with $\rho = 0.63$. Regarding the MP liberalization exercise, the alternative calibration delivers very similar results to those in Table 4.

5.2 The Rise of the East

Perhaps the single most important recent event relevant to the questions addressed in this paper is the emergence of China as a manufacturing center for firms from all over the world. In our last counterfactual exercise, we analyze how China's emergence may have affected innovation in different countries, as well as whether there could be a negative effect on workers in rich countries. For the simple scenarios in which we could explore this question analytically, Propositions 4 and 5 showed that the effects of a decline in MP costs on real wages depend on whether entry is endogenous or exogenous. In particular, it was shown how workers could lose under exogenous entry but would always gain under endogenous entry—in the latter case, the effect of integration would be an increase in innovation in the rich country with no negative effect on real wages. We now use the calibrated model to explore these questions at a quantitative level.

In the baseline equilibrium, China is partially specialized in production, as reflected in the fact that $r_{CHN} = 16\% < \eta = 19\%$. Although the gap between r_{CHN} and η may not seem large, one should keep in mind that China represents 28% of the equipped labor among the countries in our sample, and hence three percentage-points difference in this rate is larger than the entire equipped labor of several countries. Moreover, the data that we have used in our calibration is for the late 1990s. While we do not have more recent MP data for China, there is suggestive evidence that specialization of China in production has increased since then. For example, according to UNCTAD and for the period 2003-2013, the cumulative value of greenfield investments made by foreigners into China was $1,159 billions while the corresponding value for Chinese investments abroad was $251 billions.

We consider three counterfactual exercises: (i) China in autarky ($\tau'_{l,CHN}$, $\tau'_{CHN,l}$, $\gamma'_{i,CHN}$, $\gamma'_{CHN,i} \longrightarrow \infty$ for all $l, i \neq CHN$); (ii) a unilateral MP liberalization of 10% into China (i.e., $\gamma'_{i,CHN}/\gamma_{i,CHN} = 0.9$ for all $i \neq CHN$); and (iii) frictionless MP from the United States into China (i.e., $\gamma'_{USA,CHN} = 1$). We explore the implications of each of these counterfactual

changes on innovation (r), aggregate welfare (X/P), real wages for workers in production (w/P) and real wages for workers in innovation (w^e/P). Table 5 presents the results for China, the United States, Japan, Mexico and Ireland for the linear and the Roy versions of the model.

	China	USA	Japan	Mexico	Ireland
	% changes				
China in autarky					
linear model					
r	19.34	-8.58	-6.42	3.72	14.74
X/P	-2.01	-2.48	-2.46	-0.24	-1.44
Roy model					
r	19.33	-7.48	-6.49	1.86	-2.23
X/P	-2.16	-2.48	-2.50	-0.24	-1.32
w/P	-3.35	-1.75	-1.96	-0.34	-1.23
w^e/P	3.78	-4.98	-4.66	0.37	-2.06
Unilateral MP lib into China					
linear model					
r	-11.46	3.61	3.55	-1.69	-3.02
X/P	0.69	0.95	0.93	0.04	0.65
Roy model					
r	-10.04	2.91	3.15	-0.82	1.70
X/P	0.78	0.89	0.86	0.05	0.52
w/P	1.41	0.59	0.59	0.09	0.45
w^e/P	-2.71	1.86	1.91	-0.23	1.09
Frictionless MP into China from USA					
linear model					
r	-100.00	66.01	-9.02	-21.69	-100.00
X/P	61.3	12.6	0.2	2.0	4.2
Roy model					
r	-97.48	49.69	-5.74	-9.53	-22.63
X/P	43.7	13.6	0.2	1.2	2.0
w/P	51.9	7.6	0.7	1.8	2.9
w^e/P	-57.9	29.9	-1.8	-2.1	-6.4

Table 5: The Rise of China.

(i) China in autarky. Moving China back to autarky implies a major reduction in r for countries that are specialized in innovation in the baseline equilibrium. For example, r falls by between 6% and 9% in the United States and Japan, in the linear and Roy versions of the model. On the contrary, countries that are specialized in production in the baseline equilib-

rium (e.g., Ireland and Mexico) experience an increase in r according to the linear model. This might seem counterintuitive: if Ireland and China compete for inward MP, then it would be natural to expect China's disappearance from the scene to lead to *more* MP towards Ireland, leading to even further specialization in production and a decline in r. The difference here arises because innovation is *a non-rival activity*, so its decline in the United States leads to fewer firms created there, a lower demand to open up affiliates abroad, and an increase in r everywhere else. Of course, there are also direct effects that tend to decrease r in all countries other than China, and these opposing forces may result in different net effects in the alternative models – this explains the ambiguous effects in Ireland, where r increases by 15% according to the linear model but decreases by 2.2% according to the Roy model.[31]

Aggregate welfare falls everywhere, with the average loss being 1.5% in both the linear and Roy versions of the model. In the Roy model we can further study how China's move to autarky differentially affects production and innovation workers. As one would expect, given that $r_{CH} < \eta$ in the baseline equilibrium, the relative wage of innovation workers (i.e., w^e/w) in China increases by around 7% when it moves to autarky. On average, w^e/w falls in the rest of the world, but the effect in each country depends on whether r increases or decreases: w^e/w falls by around 3% in the US and Japan whereas it increases by around 0.7% in Mexico. The large increase in their relative wage implies that workers in the innovation sector actually benefit from China's move to autarky. But the converse is not true: real wages for production workers do not increase in countries that experience a decline in specialization in innovation – for example, w/P falls by around 2% in the United States and Japan. Of course, workers in innovation suffer even more (around 5% in the United States and Japan).

(ii) Unilateral MP Liberalization into China. The second counterfactual exercise is a 10% decline in MP costs from all countries into China (so that symmetry in γ's for China is relaxed). This reduction leads China to specialize even more in production, with a decline in r of around 10%, while the United States and Japan further specialize in innovation, with an increase in r of around 3%. In turn, higher innovation in the United States increases outward MP to Mexico, which deepens its specialization in production, as reflected in a decline in r of around 1%.[32]

Aggregate welfare increases by 0.7% in China and by around 0.9% in the United States and Japan, while Mexico and Ireland also gain but by smaller amounts. Interestingly, there are welfare losses in Cyprus, Greece, Romania and Turkey (see Tables in the Online Appen-

[31] With two countries r would always move in the same directions in the two versions of the model, so this flip in the sign of $r' - r$ must be related to third country effects. We also see that the sign of $r' - r$ flips as we go from the linear to the Roy model for Austria, France, and Great Britain.

[32] The effects on Ireland are again ambiguous: r decreases by 3% according to the linear model but increases by 1.7% according to the Roy model.

dix). These four countries experience declines in innovation, and hence, by the results of Section 2.5, they are candidates for welfare losses if this is caused by the intensification of home market effects associated with MP liberalization in China. Specifically, countries that are initially well linked to China benefit disproportionally from the new opportunities for outward MP generated by China's MP liberalization, leading to a reallocation of innovation to these countries and away from countries with weak links to China, who may then suffer welfare losses.

Turning to the distributional implications, the gains from MP liberalization in China are captured by production workers, who gain 1.4%, while innovation workers actually experience losses of almost 3%. Contrary to popular fears, production workers gain in the United States and Japan, although as expected their gains are lower than those of innovation workers. Interestingly, the decline in innovation in Mexico leads to losses for innovation workers there.

(iii) Frictionless MP from the United States into China. The last counterfactual exercise is the complete removal of frictions for MP into China from the United States. Such a radical MP liberalization has major effects all over the world. First, there is a dramatic reallocation of innovation from China to the United States, with innovation collapsing to nothing in China and increasing by 66% in the United States according to the linear version of the model. Second, the increase in innovation in the United States leads to a strong displacement of innovation away from countries with strong links to the United States, such as Ireland and Mexico. Third, there is a decline in r in other high-innovation countries such as Japan, Germany and Benelux (see Tables in the Online Appendix). This happens for two reasons: first, these countries now face much worse conditions for doing MP in China as wages there increase thanks to the higher demand for labor by US multinationals; second, the increase in innovation in the United States leads their multinationals to open foreign subsidiaries everywhere, extending the pressure across the globe and reducing incentives for innovation everywhere else.[33]

Aggregate welfare increases by between 44% and 61% in China and by around 13% in the United States. Other countries also gain, with few exceptions – for example, Germany loses by around 1.5%. These welfare gains are shared very unequally across production and innovation workers, for example innovation workers experience large loses in China (-58%). However, the fear that a reallocation of production by US multinationals to China could hurt workers in the United States does not materialize in our calibrated model – production workers in the United States actually experience an increase in their real income of around 8%, although of course innovation workers gain even more (30%).

[33] Again, third-country effects help some countries experience increases in innovation – for example, r increases in Austria, Hungary and Spain.

6 Conclusion

The decline in the costs of multinational production (MP) has allowed some countries to specialize in innovation while others specialize in the production of goods and services created elsewhere. To quantify the aggregate and distributional implications of this phenomenon, we develop a quantifiable general equilibrium model where firms can serve a market by exporting from their home country, by producing in the foreign market, or by exporting from a third location. In making their location decisions, firms face a proximity-comparative advantage tradeoff: trade costs may induce firms to open foreign affiliates near to their foreign customers, but this may imply not minimizing production costs. In the aggregate, countries that have a high productivity in innovation relative to production tend to specialize in innovation, but home market effects lead production to concentrate in countries with large "market potential" and draw innovation towards countries with large "production potential."

The model yields simple structural expressions for bilateral trade and MP that we use to calibrate it across a set of OECD countries. The calibration reveals that HMEs play a critical role in determining the structure of global innovation and production. We use the calibrated model to perform a series of comparative statics exercises designed to study the welfare implications of shocks driving increased specialization in innovation and production, including a generalized reduction in the cost of transferring technologies across borders, the integration of China into the global economy, and the impact of selective integration between countries. We find that countries that are driven to specialize in production due to HMEs may experience aggregate losses, although these losses tend to be very small. Contrary to popular fears, we find that production workers gain even in countries that further specialize in innovation.

References

Allen, T., Arkolakis, C. & Takahashi, Y. (2014), 'Universal gravity', *mimeo* .

Anderson, J. E. & Van Wincoop, E. (2004), 'Trade costs', *Journal of Economic Literature* **42**(3), 691–751.

Arkolakis, C., Costinot, A. & Rodríguez-Clare, A. (2012), 'New trade models, same old gains?', *American Economic Review* **102**(1), 94–130.

Arkolakis, C., Demidova, S., Klenow, P. J. & Rodríguez-Clare, A. (2008), 'Endogenous variety and the gains from trade', *American Economic Review, Papers and Proceedings* **98**(4), 444–450.

Arkolakis, C., Ramondo, N., Rodríguez-Clare, A. & Yeaple, S. (2013), 'Innovation and production in the global economy', *NBER* **18792**.

Autor, D., Dorn, D. & Hanson, G. H. (2012), The china syndrome: Local labor market effects of import competition in the united states, Technical report, National Bureau of Economic Research.

Burstein, A. & Monge-Naranjo, A. (2009), 'Foreign know-how, firm-control, and the income of developing countries', *Quarterly Journal of Economics* **124**(1), 149–195.

Burstein, A. & Vogel, J. (2012), 'International trade, technology, and the skill premium'. Manuscript, Columbia University and UCLA.

Caliendo, L. & Parro, F. (2014), 'Estimates of the trade and welfare effects of NAFTA'. forthcoming, Review of Economic Studies.

Chaney, T. (2008), 'Distorted gravity: The intensive and extensive margins of international trade', *American Economic Review* **98**(4), 1707–1721.

Corrado, C., Hulten, C. & Sichel, D. (2009), 'Intangible capital and u.s. economic growth', *The Review of Income and Wealth* **55**(3), 661–685.

Costinot, A. & Rodriguez-Clare, A. (2014), Trade theory with numbers: Quantifying the consequences of globalization, *in* G. Gopinath, E. Helpman & K. Rogoff, eds, 'Handbook of International Economics', Vol. 4, chapter 4.

Domowitz, I., Hubbard, R. G. & Petersen, B. C. (1988), 'Market structure and cyclical fluctuations in us manufacturing', *Review of Economics and Statistics* **70**(1), 55–66.

Donaldson, D. (2014), 'Railroads of the raj: Estimating the economic impact of transportation infrastructure'. forthcoming, American Economic Review.

Eaton, J. & Kortum, S. (2001), 'Technology, trade, and growth: A unified framework', *European Economic Review* **45**(4-6), 1741–1780.

Eaton, J. & Kortum, S. (2002), 'Technology, geography and trade', *Econometrica* **70**(5), 1741–1779.

Eaton, J. & Kortum, S. (2007), Innovation, diffusion, and trade, *in* E. Sheshinski, R. J. Strom & W. J. Baumol, eds, 'Entrepreneurship, Innovation, and the Growth Mechanism of the Free-Enterprise Economies', Princeton University Press, pp. 276–299.

Fajgelbaum, P. D., Grossman, G. M. & Helpman, E. (2014), 'A linder hypothesis for foreign direct investment', *forthcoming, Review of Economic Studies* .

Feenstra, R. C. & Hanson, G. H. (1999), 'The impact of outsourcing and high-technology capital on wages: estimates for the united states, 1979–1990', *The Quarterly Journal of Economics* **114**(3), 907–940.

Feenstra, R. C., Romalis, J. & Schott, P. K. (2002), 'US imports, exports, and tariff data, 1989-2001', *NBER Working Paper* **9387**.

Grossman, G. & Helpman, E. (1991), *Innovation and Growth in the Global Economy*, MIT press, Cambridge, Massachussets.

Grossman, G. M. & Rossi-Hansberg, E. (2008), 'Trading tasks: A simple theory of offshoring', *American Economic Review* **98**(5), 1978.

Head, K. & Ries, J. (2001), 'Increasing returns versus national product differentiation as an explanation for the pattern of u.s.-canada trade', *American Economic Review* **91**(4), 858–876.

Helpman, E. (1984), 'A simple theory of international trade with multinational corporations', *Journal of Political Economy* **92**(3), 451–471.

Helpman, E., Melitz, M. J. & Yeaple, S. R. (2004), 'Export versus FDI with heterogeneous firms', *American Economic Review* **94**(1), 300–316.

Hsieh, C.-T., Hurst, E., Jones, C. I. & Klenow, P. J. (2011), The allocation of talent and us economic growth, Technical report, Working Paper.

Hsieh, C.-T., Hurst, E., Jones, C. I. & Klenow, P. J. (2013), The allocation of talent and us economic growth, Technical report, National Bureau of Economic Research.

Irarrazabal, A., Moxnes, A. & Opromolla, L. D. (2013), 'The margins of multinational production and the role of intra-firm trade', *Journal of Political Economy* **121**(1), 74–126.

Klenow, P. & Rodríguez-Clare, A. (2005), Externalities and growth, *in* P. Aghion & S. Durlauf, eds, 'Handbook of Economic Growth', Vol. 1A, Elsevier, chapter 11, pp. 817–861.

Kokko, A. (1992), *Foreign Direct Investment, Host Country Characteristics, and Spillovers*, Stockholm School of Economics, Economic Research Institute, Stockholm.

Krugman, P. (1980), 'Scale economies, product differentiation, and the pattern of trade', *American Economic Review* **70**(5), 950–959.

Lagakos, D. & Waugh, M. E. (2013), 'Specialization, agriculture, and cross-country productivity differences', *American Economic Review* **103**(2), 948–980.

Markusen, J. R. (1984), 'Multinationals, multi-plant economies, and the gains from trade', *Journal of International Economics* **16**(3), 205–226.

Markusen, J. R. (2002), *Multinational Firms and The Theory of International Trade*, MIT Press, Massachussetts.

Markusen, J. R. & Venables, A. (1998), 'Multinational firms and the new trade theory', *Journal of International Economics* **46**(2), 183–204.

Markusen, J. R. & Venables, A. (2000), 'The theory of endowment, intra-industry, and multinational trade', *Journal of International Economics* **52**(2), 52.

Martins, J. O., Scarpetta, S. & Pilat, D. (1996), 'Mark-up ratios in manufacturing industries: Estimates for 14 OECD countries', *OECD Economics Department Working Paper* **162**.

McGrattan, E. (2011), 'Transition to FDI openness: Reconciling theory and evidence', *Minneapolis Federal Reserve Staff Report* **455**.

McGrattan, E. & Prescott, E. C. (2010), 'Technology capital and the US current account', *American Economic Review* **100**(4), 1493–1522.

Melitz, M. J. (2003), 'The impact of trade on intra-industry reallocations and aggregate industry productivity', *Econometrica* **71**(6), 1695–1725.

Nielsen, R. B. (2006), *An Introduction to Copulas*, Spinger, New York.

Ramondo, N. (2012), 'A quantitative approach to multinational production', *Journal of International Economics* .

Ramondo, N. & Rodríguez-Clare, A. (2013), 'Trade, multinational production, and the gains from openness', *Journal of Political Economy* .

Ramondo, N., Rodriguez-Clare, A. & Tintelnot, F. (2014), 'Multinational production data set'.

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

Romalis, J. (2007), 'NAFTA's and CUFTA's impact on international trade', *Review of Economics and Statistics* **89**(3), 416–435.

Roy, A. D. (1951), 'Some thoughts on the distribution of earnings', *Oxford economic papers* **3**(2), 135–146.

Silva, J. M. C. S. & Tenreyro, S. (2006), 'The log of gravity', *Review of Economics and Statistics* **88**(4), 641–658.

Simonovska, I. & Waugh, M. (2013), 'The elasticity of trade: Estimates and evidence', *Journal of International Economics* **92**, 34–50.

Timmer, M. P. (2012), 'The world input-output database (wiod): Contents, sources and methods', *WIOD Working Paper Number 10* .

Tintelnot, F. (2014), 'Global production with export platforms', *mimeo* .

Venables, A. J. (1987), 'Trade and trade policy with differentiated products: A chamberlinian-ricardian model', *The Economic Journal* pp. 700–717.

Supplementary Appendix not for Publication

A Data Appendix

Data used for the restricted and unrestricted gravity estimations. The production data for the *restricted sample* (X_{iln}, where i = U.S.) were assembled from several sources that depend on the location of production l. For the case of $l \neq U.S.$ (U.S. MP abroad), our data are from the confidential 1999 survey of the BEA of U.S. direct investment abroad. This legally mandatory survey identifies all U.S. firms that own productive facilities abroad. The survey requires firms to report for their majority-owned, manufacturing affiliates the location of the affiliates l, the sales of these affiliates to customers in their host country ($l = n$) and their sales to customers in the United States, Canada, Japan, the United Kingdom, and an aggregation of a subset of countries in the European Union ($l \neq U.S., n$).[34] For the case of $l = U.S.$, the data was constructed using a mixture of publicly available data and a confidential survey conducted by the BEA on the activities of the U.S. affiliates of foreign firms. Aggregate bilateral trade volumes in manufactures and aggregate domestic manufacturing sales were collected from Feenstra, Romalis & Schott (2002) and the Census of Manufacturing respectively. From these aggregates we subtracted the total contribution of foreign firms to these sales using the BEA data set.

The data for the *unrestricted sample* ($\sum_i X_{iln}$) were also constructed using data from several sources. The bilateral trade data ($l \neq n$) came from Feenstra et al. (2002) for the year 1999. The domestic production data ($l = n$) was collected from the OECD for most developed countries, from the INSTAT database maintained by UNIDO for many of the developing countries, and for a few additional countries the domestic absorption data was obtained from the estimates found in Simonovska & Waugh (2013). In the estimation we use only those bilateral pair observations for which both X_{iln} and X_{ln} are both nonzero and non-missing, yielding a sample size of 316.

The data for trade frictions was drawn from several sources. The raw tariff data was obtained from either the WTO or from the WITS web-site maintained by the World Bank. Tariffs applied by a given country n can differ from their MFN levels across exporting countries l either because no tariff is applied, as when $n = l$ or n and l are both in a free trade agreement or customs union, or because country n extends GSP tariffs to a developing country l. Data for distance (d_{ln}) and for the standard gravity controls (H_{ln}) are from the *Centre d'Etudes*

[34] These countries are Austria, Belgium, Denmark, France, Finland, Germany, Greece, Luxembourg, Ireland, Italy, Netherlands, Portugal, Spain, and Sweden. The BEA data for affiliate exports contains information on the destination for only these four countries and for seven regions in total. Of these regions, only the European countries share a common tariff.

		Distance Dummies					Other Gravity Controls			
	Tariff	D2	D3	D4	D5	D6	Self	Border	Lang.	Chi-sq.
Restricted	-8.4	-0.4	-1.7	-0.4	-1.6	-1.2	2.2	0.3	-0.2	6006
	(2.6)	(0.3)	(0.4)	(0.2)	(0.4)	(0.6)	(0.5)	(0.4)	(0.2)	
Unrestricted	-5.4	-0.6	-0.6	-0.8	-1.7	-1.3	2.8	0.8	-0.2	34975
	(1.7)	(0.1)	(0.3)	(0.1)	(0.2)	(0.2)	(0.2)	(0.2)	(0.2)	

Table 6: Restricted and Unrestricted Gravity: Poisson PML Estimator
Notes: Robust standard errors in parenthesis.

Prospectives et Informations Internationales (CEPII) web-site. To allow for non-linearities in the effect of distance on trade cost, we constructed six categorical variables ($D1$ through $D6$) defined by the size of the distance.[35] Finally, a dummy variable was included that takes a value of one for the case in which $l = n$ and a value of zero for the case $l \neq n$.

Poisson Estimation

Estimating the gravity equations via OLS, we assume that any source of variation in the error term, ε_{ln}, is orthogonal to the independent variables. This is a strong assumption, but one that is routinely maintained in gravity equation estimation. However, if ε_{ln} is heteroscedastic, OLS estimates may be biased (see Silva & Tenreyro (2006)). To study the effect of this bias, we re-estimate both gravity models using the Poisson pseudo-maximum-likelihood (PML) estimator proposed by Silva & Tenreyro (2006) and report the coefficient estimates in Table 6. The results are somewhat affected by the alternative estimation procedure, with the trade elasticity slightly smaller for the restricted gravity case and somewhat larger for the unrestricted case. These Poisson estimates provide us with an alternative parameter choice to our baseline calibration.

Calibrated trade and MP costs and geographical distance

Figure 8 illustrates the strong positive relationship between the calibrated trade and MP costs with geographical distance.

B Theory Appendix

B.1 Properties of the Multivariate Pareto

(i) We show that with $\rho \to 1$ the elements of $\mathbf{z}$ are perfectly correlated, i.e. $\lim_{\rho \to 1} G_i(z_1, ..., z_N) = 1 - \max_l T_{il} z_l^{-\theta}$. Let $x \equiv \max_l T_{il} z_l^{-\theta}$ and note that $G_i(z_1, ..., z_N) = 1 - x \left(\sum_{l=1}^{N} \left(T_{il} z_l^{-\theta} / x \right)^{\frac{1}{1-\rho}} \right)^{1-\rho}$.

[35] The categories are less than 1,000km, between 1,000 and 3,000km, between 3000 and 6000km, between 6000 and 9000km, between 9,000 and 12,000km, and greater than 12,000 km.

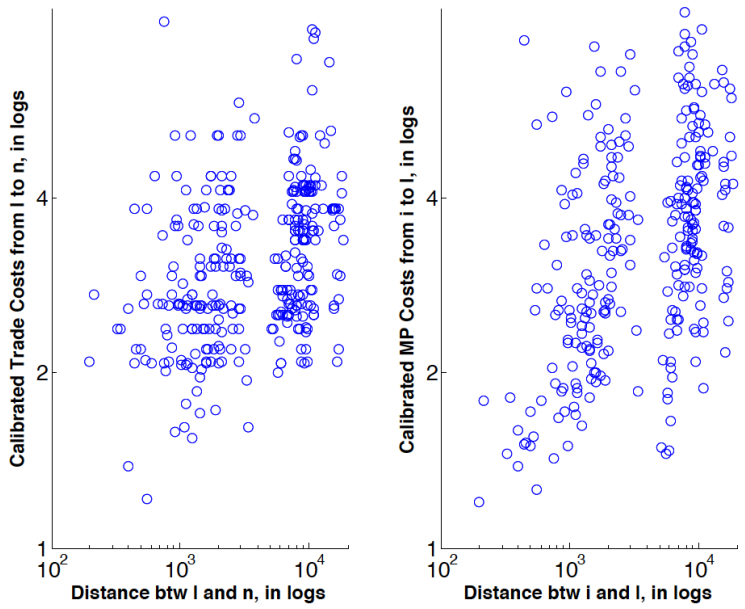


Figure 8: Calibrated Trade and MP Costs and Geographical Distance.

As $\rho \to 1$ then $\sum_{l=1}^{N} \left(T_{il} z_l^{-\theta} / x \right)^{\frac{1}{1-\rho}} \to 1$, proving the result.

(ii) We also show that $\rho = 0$ is equivalent to the case of the production location l chosen randomly with probabilities $T_{il}/\widetilde{T}_i$ among all possible locations $l = 1, ..., N$, and the productivity Z_l chosen from the Pareto distribution $1 - \widetilde{T}_i z_l^{-\theta}$ with $z_l \geq \widetilde{T}_i^{1/\theta}$. We simply need to prove that for $l \neq k$ we have $\Pr(Z_l > \widetilde{T}_i^{1/\theta} \cap Z_k > \widetilde{T}_i^{1/\theta}) = 0$, and $\Pr(Z_l \leq z_l \cap Z_k = \widetilde{T}_i^{1/\theta}$ for all $k \neq l) = \left(T_{il}/\widetilde{T}_i \right) \left(1 - \widetilde{T}_i z_l^{-\theta} \right)$. Note that with $\rho = 0$ the density associated with the distribution above is zero, if it is evaluated at a point with $Z_v > \widetilde{T}_i^{1/\theta}$ for two or more v, while $\Pr(Z_l \leq z_l \cap Z_k = \widetilde{T}_i^{1/\theta}$ for all $k \neq l) = 1 - \left[\sum_{k \neq l}^{N} T_{ik}/\widetilde{T}_i + T_{il} z_l^{-\theta} \right] = \left(T_{il}/\widetilde{T}_i \right) \left(1 - \widetilde{T}_i z_l^{-\theta} \right)$ proving the result.

B.2 Proof of Lemma 1

The (unconditional) probability that a firm from i will serve market n from l is

$$\Pr \left(\arg\min_k C_{ikn} = l \cap \min_k C_{ikn} \leq c_n^* \right).$$

To compute this probability, note that,

$$\Pr \left(C_{i1n} \geq c_{i1n}, ..., C_{iNn} \geq c_{iNn} \right) = \Pr \left(Z_1 \leq \frac{\xi_{i1n}}{c_{i1n}}, ..., Z_N \leq \frac{\xi_{iNn}}{c_{iNn}} \right).$$

Assuming that $c_{ikn} \leq \xi_{ikn}\widetilde{T}_i^{-1/\theta}$ for all k, then our assumption regarding the distribution of $\boldsymbol{z}$ for firms in country i implies that

$$\Pr\left(Z_1 \leq \frac{\xi_{i1n}}{c_{i1n}}, ..., Z_N \leq \frac{\xi_{iNn}}{c_{iNn}}\right) = 1 - \left(\sum_{k=1}^{N}\left[T_{ik}\left(\frac{\xi_{ikn}}{c_{ikn}}\right)^{-\theta}\right]^{\frac{1}{1-\rho}}\right)^{1-\rho}. \tag{B.1}$$

But we know that

$$\Pr\left(C_{i1n} \geq c_{i1n}, ..., C_{iln} = c_{iln}, ..., C_{iNn} \geq c_{iNn}\right) = -\frac{\partial \Pr\left(C_{i1n} \geq c_{i1n}, ..., C_{iln} = c_{iln}, ..., C_{iNn} \geq c_{iNn}\right)}{\partial c_{iln}},$$

hence from $(B.1)$ we get

$$\Pr\left(C_{i1n} \geq c_{i1n}, ..., C_{iln} = c_{iln}, ..., C_{iNn} \geq c_{iNn}\right) = \theta\left(\sum_{k=1}^{N}\left[T_{ik}\left(\frac{\xi_{ikn}}{c_{ikn}}\right)^{-\theta}\right]^{\frac{1}{1-\rho}}\right)^{-\rho}\left(T_{il}\xi_{iln}^{-\theta}\right)^{\frac{1}{1-\rho}} c_{iln}^{\theta/(1-\rho)-1}.$$

Notice also that if $c < \xi_{ikn}\widetilde{T}_i^{-1/\theta}$ for all k,

$$\begin{aligned}\Pr\left(\arg\min_k C_{ikn} = l \cap \min_k C_{ikn} = c\right) &= \Pr\left(C_{i1n} \geq c, ..., C_{iln} = c, ..., C_{iNn} \geq c\right) \\ &= \theta\Psi_{in}^{-\frac{\rho}{1-\rho}}\left(T_{il}\xi_{iln}^{-\theta}\right)^{\frac{1}{1-\rho}} c^{\theta-1} = \psi_{iln}\Psi_{in}\theta c^{\theta-1}.\end{aligned}$$

Given Assumption 1 we know that $c_{in}^* < \xi_{ikn}\widetilde{T}_i^{-1/\theta}$ so that we can integrate the previous expression over c from 0 to c_n^* to show that the probability that firms from i serving market n will choose location l for production is

$$\Pr\left(\arg\min_k C_{ikn} = l \cap \min_k C_{ikn} \leq c_n^*\right) = \psi_{iln}\Psi_{in}\left(c_n^*\right)^{\theta}, \tag{B.2}$$

while

$$\Pr\left(\min_k C_{ikn} \leq c_n^*\right) = \sum_k \psi_{ikn}\Psi_{in}\left(c_n^*\right)^{\theta} = \Psi_{in}\left(c_n^*\right)^{\theta}.$$

Hence,

$$\Pr\left(\arg\min_k C_{ikn} = l \mid \min_k C_{ikn} \leq c_n^*\right) = \psi_{iln};$$

QED.

B.3 Derivations of Equations 9 and 7

Multiplying $(B.2)$ by the measure of firms in i, M_i, and using (3), we get the measure of firms from i that serve market n from location l,

$$M_{iln} = M_i\psi_{iln}\Psi_{in}\left(\frac{\sigma w_n F_n}{X_n}\right)^{-\theta/(\sigma-1)}\frac{P_n^{\theta}}{\widetilde{\sigma}^{\theta}}. \tag{B.3}$$

Since the sales of a firm with cost c in a market n are $\widetilde{\sigma}^{1-\sigma}X_n P_n^{\sigma-1}c^{1-\sigma}$, equation (5) implies that total sales from n to l by firms from i, X_{iln}, are

$$X_{iln} = M_i\psi_{iln}\Psi_{in}\widetilde{\sigma}^{1-\sigma}X_n P_n^{\sigma-1}\int_0^{c_n^*}\theta c^{\theta-\sigma}dc.$$

Solving for the integral, using (3) and simplifying yields

$$X_{iln} = \frac{\widetilde{\sigma}^{-\theta}\theta}{\theta-\sigma+1}M_i\psi_{iln}\Psi_{in}\left(\sigma w_n F_n\right)^{(\theta-\sigma+1)/(1-\sigma)}X_n^{\theta/(\sigma-1)}P_n^{\theta}. \tag{B.4}$$

Combining (B.4) and (B.3) yields (9). In turn, the formula for the price index in (1) together with the pricing rule in (2), the density in (5), and the cut-off in (3) imply that

$$P_n^{-\theta} = \zeta^{\theta} \left(\frac{w_n F_n}{X_n} \right)^{(\theta-\sigma+1)/(1-\sigma)} \sum_k M_k \Psi_{kn}, \tag{B.5}$$

where $\zeta \equiv \left(\frac{\widetilde{\sigma}^{1-\sigma}\theta}{\theta-\sigma+1} \right)^{1/\theta} \left(\frac{\sigma}{\widetilde{\sigma}^{1-\sigma}} \right)^{\frac{\sigma-1-\theta}{\theta(\sigma-1)}}$. Plugging this result into (B.4), we obtain (7) by noting that λ_{in}^E is given by expression (8).

B.4 Proof of Proposition 1

Using the equilibrium conditions for the case of free entry and setting $\tau_{ln} = 1$ and $\gamma_{il} = 1$ for all i, l, n, we get that

$$\Psi_{in} = T_i^e \left[\sum_k \left(T_k^p w_k^{-\theta} \right)^{\frac{1}{1-\rho}} \right]^{1-\rho} \equiv \Psi_i \,, \tag{B.6}$$

and we can write

$$\lambda_{in}^E = \frac{M_i T_i^e}{\sum_k M_k T_k^e} \equiv \lambda_i^E,$$

and

$$\lambda_{ln}^T = \frac{\left(T_l^p w_l^{-\theta} \right)^{1/(1-\rho)}}{\sum_k \left(T_k^p w_k^{-\theta} \right)^{1/(1-\rho)}} \equiv \lambda_l^T.$$

These expressions imply that the labor market clearing and zero-profit conditions (i.e., equations 13 and 15) can be written as

$$\frac{1}{\widetilde{\sigma}} \frac{\left(T_i^p w_i^{-\theta} \right)^{1/(1-\rho)}}{\sum_k \left(T_k^p w_k^{-\theta} \right)^{1/(1-\rho)}} \sum_k w_k L_k + w_i M_i f^e = w_i L_i \left(1 - \frac{\theta-\sigma+1}{\sigma\theta} \right), \tag{B.7}$$

and

$$w_i = \eta \frac{T_i^e / f^e}{\sum_k M_k T_k^e} \sum_k w_k L_k. \tag{B.8}$$

Recalling that $M_i f^e = L_i^e$, and combining $(B.7)$ and $(B.8)$ yields

$$r_i \equiv \frac{M_i f^e}{L_{i.}} = \left(1 - \frac{\theta-\sigma+1}{\sigma\theta} \right) - \frac{1}{\widetilde{\sigma}} \frac{f^e}{\eta} \frac{\left[T_i^p \left(T_i^e \right)^{-\theta} \right]^{1/(1-\rho)}}{\sum_k \left[T_k^p \left(T_k^e \right)^{-\theta} \right]^{1/(1-\rho)}} \frac{\sum_k M_k T_k^e}{T_i^e L_i} \,.$$

Noting that $1 - \frac{\theta-\sigma+1}{\sigma\theta} = \eta + 1/\widetilde{\sigma}$, and letting $A_i \equiv \left(T_i^p \right)^{1/(1-\rho)} / L_i$ gives

$$r_i = \frac{1}{\widetilde{\sigma}} \left(1 - \frac{f^e}{\eta} \frac{A_i / \left(T_i^e \right)^{\theta/(1-\rho)}}{\sum_k A_k L_k / \left(T_j^e \right)^{\theta/(1-\rho)}} \frac{\sum_j M_j T_j^e}{T_i^e} \right) + \eta. \tag{B.9}$$

Finally, notice that by the definition of r_i, we have $M_i = r_i L_i / f^e$, which can be substituted in $(B.9)$ to construct the term

$$\sum_k M_k T_k^e = \eta \frac{\sum_k L_k T_k^e}{f^e} \,. \tag{B.10}$$

Replacing back in $(B.9)$ and defining $\delta_i \equiv L_i T_i^e / \sum_k L_k T_k^e$, we finally obtain (20) and the necessary and sufficient condition for this expression to hold, as indicated in the Proposition.

B.5 Proof of Proposition 2

Part (i) First, as a preliminary result, we establish that if $L_1 > L_2$, then $\omega \equiv w_1/w_2 > 1$. The absence of trade costs implies that $\lambda_{in}^E \equiv \lambda_i^E$ for any i, n (For future reference, note that this implies that $\lambda_1^E + \lambda_2^E = \lambda_{11}^E + \lambda_{21}^E = 1$) and that $\Psi_{in} \equiv \Psi_i$ for any i, n. The zero-profit condition, Equation (15), implies

$$L_1^e = \eta \lambda_1^E (L_1 + L_2/\omega), \tag{B.11}$$

$$L_2^e/\omega = \eta \lambda_2^E (L_1 + L_2/\omega). \tag{B.12}$$

Using these equations together with the definition of λ_i^E, which implies that

$$\lambda_i^E = \frac{M_i \Psi_{in}}{\sum_k M_k \Psi_{kn}} = \frac{M_i \Psi_i}{\sum_k M_k \Psi_k}, \tag{B.13}$$

and $M_i f^e = L_i^e$, we have $\omega = \Psi_1/\Psi_2$.

Using the definition of Ψ_{in} and the assumption of $A_1 = A_2$, we can obtain after some derivations

$$\frac{L_1}{L_2} = \omega^{\frac{\theta}{1-\rho}} \frac{\omega^{1/(1-\rho)} - \gamma^{-\theta/(1-\rho)}}{1 - \gamma^{-\theta/(1-\rho)} \omega^{1/(1-\rho)}}. \tag{B.14}$$

The right hand side (RHS) of this equation is increasing in ω which implies that ω is increasing in L_1/L_2. Since $L_1/L_2 = 1$ implies $\omega = 1$, then $L_1/L_2 > 1$ implies $\omega > 1$, which proves the preliminary result.

Second, using the previous result we can prove that if $L_1 > L_2$ then $r_1 > r_2$. The proof is by contradiction. Suppose that $r_1 < r_2$. From the labor market clearing condition, Equation (13), and from (14) and $\lambda_{in}^T = \lambda_i^T \equiv \lambda_{ii}^T$, we have

$$w_i L_i^e = w_i L_i \left(1 - \frac{\theta - \sigma + 1}{\sigma \theta}\right) - \frac{1}{\widetilde{\sigma}} \lambda_i^T \sum_k w_k L_k \implies$$

$$r_i = \eta + 1 - 1/\sigma - \frac{1}{\widetilde{\sigma}} \frac{\lambda_i^T \sum_k w_k L_k}{w_i L_{i.}}$$

If $r_1 < r_2$ then labor market clearing in the two countries requires

$$\frac{\lambda_1^T}{w_1 L_1} > \frac{\lambda_2^T}{w_2 L_2} \tag{B.15}$$

Using the definition for λ_l^T, the result $\lambda_{in}^E = \lambda_i^E$, and $(B.11)$ and $(B.12)$, after some derivations expression (B.15) implies

$$L_2 r_2 \omega^{\frac{\rho}{1-\rho}} \left(\gamma^{-\theta/(1-\rho)} - \omega^{\theta/(1-\rho)+1}\right) > L_1 r_1 \left(\omega^{\theta/(1-\rho)+1} \gamma^{-\theta/(1-\rho)} - 1\right),$$

which will finally allow us to prove the result by contradiction. Note that when $L_1 > L_2$ we have $\omega > 1$, so that the term in parentheses on the left-hand-side of this inequality is negative. If $\omega^{\theta/(1-\rho)+1} \gamma^{-\theta/(1-\rho)} \geq 1$, then the inequality is violated and the desired contradiction is shown. Alternatively, if $\omega^{\theta/(1-\rho)+1} \gamma^{-\theta/(1-\rho)} < 1$ we can substitute out L_2/L_1 from

the inequality using $(B.14)$ to arrive at an expression that given the assumption that $\theta > 1$ contradicts the initial assertion that $r_1 < r_2$. Thus, since this assertion leads to a contradiction in all cases, we conclude that $r_1 > r_2$, which completes the proof of part i).

Part (ii) To simplify the notation, without loss of generality, we assume that $T_1^e = T_2^e = 1$ and use T_i as shorthand for T_i^p. Frictionless MP implies that $\Psi_{in} = \Psi_n$ for any i, j, n and $\lambda_{ij}^E = M_i/(M_1 + M_2)$ for any i, j. The labor market clearing in production and marketing is given by (13), which in this case implies

$$w_1 L_1^p = \frac{\theta - \sigma + 1}{\theta\sigma} X_1 + \frac{1}{\widetilde{\sigma}} \left[\lambda_{11}^T w_1 L_1 + \lambda_{12}^T w_2 L_2 \right].$$

But given the absence of MP costs and the implication that $\Psi_{in} = \Psi_n$ for any i, then this equation can be rewritten as

$$w_1 L_1^p = \frac{\theta - \sigma + 1}{\theta\sigma} X_1 + \frac{1}{\widetilde{\sigma}} \left(T_1 w_1^{-\theta} \right)^{\frac{1}{1-\rho}} \left[\Psi_1^{-1/(1-\rho)} w_1 L_1 + \tau^{-\theta/(1-\rho)} \Psi_2^{-1/(1-\rho)} w_2 L_2 \right].$$

In an interior solution, when both countries innovate, we must have $w_1 = w_2$, or else only the lower wage country would innovate. We normalize this wage to one. Using the definition of Ψ_{in}, and given the assumption that $A_i \equiv T_i^{1/(1-\rho)}/L_i$ we have $\Psi_{in} = \left[\sum_k A_k L_k \tau_{kn}^{-\theta/(1-\rho)} \right]^{1-\rho}$. Finally, using symmetry $A_1 = A_2 = A$ and letting $t \equiv \tau^{-\theta/(1-\rho)}$ and $l_1 \equiv L_1/(L_1 + L_2)$ we get the above expression to be

$$\frac{L_1^p}{L_1} = \frac{\theta - \sigma + 1}{\theta\sigma} + \frac{1}{\widetilde{\sigma}} \left[\frac{l_1}{t + l_1(1-t)} + \frac{t(1-l_1)}{1 - l_1(1-t)} \right].$$

and similarly for the second country. Noting that $r_i = 1 - L_i^p/L_i$, we have[1]

$$\begin{aligned} r_1 &= \eta \left[1 + \theta \frac{t(1-t)(1-l_1)(1-2l_1)}{(l_1 + (1-l_1)\,t)(l_1 t + (1-l_1))} \right] \\ r_2 &= \eta \left[1 + \theta \frac{t(1-t) l_1 (2l_1 - 1)}{(l_1 + (1-l_1)\,t)(l_1 t + (1-l_1))} \right] \end{aligned}$$

It is clear from these expressions that for any $l_1 \in (1/2, 1)$ that $r_1 < \eta$ and $r_2 > \eta$, **QED**.

B.6 Proof of Proposition 3

The proof of both parts comprises of a number of straightforward algebraic steps for which we simply sketch the intuition. In both cases, we solve for r_1 using the labor market clearing (13) and free entry conditions (15). Then, imposing alternatively frictionless trade or MP results in expressions in terms of the ψ_{iln}, which determine the size of r_1 relative to η. For part (i), frictionless trade but costly MP, showing that $r_1 > \eta$ requires showing that $\theta w_1/w_2 > \psi_{211}/\psi_{121}$. This inequality can be shown to hold using the definition of ψ_{iln} and by proving that $\gamma_{12} < \gamma_{21} \implies w_1/w_2 > 1$. For part (ii), frictionless MP but costly trade, we need to show $r_1 < \eta$. Here, the derivations imply that $r_1 = \eta + \frac{\sigma-1}{\sigma}\left(\lambda_{21}^T - \lambda_{12}^T\right)$, and the inequality can be shown to be true as long as $\psi_{112} > \psi_{121} \implies \lambda_{21}^T < \lambda_{12}^T$. This last condition follows from $\tau_{21} > \tau_{12}$. In both parts, it is easy to show that, $r_1 > \eta \implies r_2 < \eta$.

B.7 Real Wage in Terms of Flows

We prove two lemmas that characterize the real wage and real expenditure under exogenous and endogenous entry.

Lemma 2 *Under exogenous entry, real wages are given by*

$$\frac{w_n}{P_n} = \kappa_n \left(T_n^e T_n^p M_n\right)^{\frac{1}{\theta}} \left(\frac{X_{nnn}}{X_n}\right)^{-\frac{1-\rho}{\theta}} \left(\frac{\sum_l X_{nln}}{X_n}\right)^{-\frac{\rho}{\theta}} \left(\frac{1}{\widetilde{\sigma}}\frac{Y_n}{X_n} + \frac{\theta-\sigma+1}{\theta\sigma}\right)^{\frac{\sigma-1-\theta}{\theta(\sigma-1)}}, \tag{B.16}$$

and real expenditure is given by

$$\frac{X_n}{P_n}\frac{1}{L_n} = \kappa_n \left(T_n^e T_n^p M_n\right)^{\frac{1}{\theta}} \left(\frac{X_{nnn}}{X_n}\right)^{-\frac{1-\rho}{\theta}} \left(\frac{\sum_l X_{nln}}{X_n}\right)^{-\frac{\rho}{\theta}} \left(\frac{1}{\widetilde{\sigma}}\frac{Y_n}{X_n} + \frac{\theta-\sigma+1}{\theta\sigma}\right)^{\frac{\sigma-1-\theta}{\theta(\sigma-1)}-1},$$

where $\kappa_n \equiv \zeta \left(F_n/L_n\right)^{\frac{\sigma-1-\theta}{\theta(\sigma-1)}}$.

Proof: We start with the definition of λ_{ln}^T, Equation (10). Using also the definitions of ψ_{iln} and ξ_{iln}, setting $l = n$ and solving for w_n, we have

$$w_n = \left[\frac{\lambda_{nn}^T}{\sum_k \left(T_{kn}\gamma_{kn}^{-\theta}/\Psi_{kn}\right)^{1/(1-\rho)} \lambda_{kn}^E}\right]^{-(1-\rho)/\theta}.$$

Using the result for the Dixit-Stiglitz price index in $(B.5)$, $\zeta \equiv \left(\frac{\widetilde{\sigma}^{1-\sigma}\theta}{\theta-\sigma+1}\right)^{1/\theta}\left(\frac{\sigma}{\widetilde{\sigma}^{1-\sigma}}\right)^{\frac{\sigma-1-\theta}{\theta(\sigma-1)}}$, and noting that the definition of λ_{in}^E implies that $\sum_k M_k \Psi_{kn} = M_n \Psi_{nn}/\lambda_{nn}^E$, we can write

$$P_n = \zeta^{-1}\left[\left(\frac{w_n F_n}{X_n}\right)^{1-\theta/(\sigma-1)} \frac{M_n \Psi_{nn}}{\lambda_{nn}^E}\right]^{-1/\theta}.$$

Combining the two previous expressions and using $T_{in} = T_i^e T_n^p$, we get

$$\frac{w_n}{P_n} = \zeta \left(T_n^e T_n^p M_n\right)^{1/\theta} \left(\lambda_{nn}^T\right)^{\frac{\rho-1}{\theta}} \left(\lambda_{nn}^E\right)^{-\frac{1}{\theta}} \left[\sum_k \left(T_k^e \gamma_{kn}^{-\theta}\frac{\Psi_{nn}}{\Psi_{kn}}\right)^{\frac{1}{1-\rho}} \lambda_{kn}^E\right]^{\frac{1-\rho}{\theta}} \left(\frac{w_n F_n}{X_n}\right)^{\frac{\sigma-1-\theta}{\theta(\sigma-1)}}. \tag{B.17}$$

Notice that using (7), the definition of ψ_{iln}, and simplifying, we get

$$\sum_k \left(T_k^e \gamma_{kn}^{-\theta}\frac{\Psi_{nn}}{\Psi_{kn}}\right)^{\frac{1}{1-\rho}} \lambda_{kn}^E = \frac{\left(T_n^e\right)^{\frac{1}{1-\rho}} \lambda_{nn}^E}{X_{nnn}/\sum_i X_{inn}}.$$

Plugging this into $(B.17)$, and using the definitions of λ_{nn}^T, and λ_{nn}^E, we get that the real wage is given by

$$\frac{w_n}{P_n} = \zeta \left(T_n^e M_n\right)^{1/\theta} \left(\frac{X_{nnn}}{X_n}\right)^{-\frac{1-\rho}{\theta}} \left(\frac{\sum_l X_{nln}}{X_n}\right)^{-\frac{\rho}{\theta}} \left(\frac{w_n F_n}{X_n}\right)^{\frac{\sigma-1-\theta}{\theta(\sigma-1)}}. \tag{B.18}$$

Under restricted entry, the labor market clearing condition is given by (13), which implies

$$w_n L_n = \frac{1}{\widetilde{\sigma}} Y_n + \frac{\theta-\sigma+1}{\theta\sigma} X_n.$$

Solving this equation for w_n/X_n and substituting in (B.18) leads to equation (B.16) for real wages. To derive $X_n/P_n = (X_n/w_n) \times (w_n/P_n)$, we simply use the expression for real wages

and we obtain the ratio X_n/w_n using the labor market clearing condition above. This last step completes the proof; **QED**.

Lemma 3 *Under endogenous entry, real wage and real expenditure are given by*

$$\frac{w_n}{P_n} = \frac{X_n}{P_n}\frac{1}{L_n} = \kappa_n \left(T_n^e M_n\right)^{1/\theta} \left(\frac{X_{nnn}}{X_n}\right)^{-\frac{1-\rho}{\theta}} \left(\frac{\sum_l X_{nln}}{X_n}\right)^{-\frac{\rho}{\theta}} . \tag{B.19}$$

Proof: This follows immediately from imposing $X_n = w_n L_n$ in (B.18), **QED**.

B.8 Proof of Proposition 4

The assumption $\left(T_i^p\right)^{1/(1-\rho)}/L_i^p = \tilde{A}$ implies that $T_{ii} = T_i^e T_i^p = T^e \left(L_i^p\right)^{1-\rho}$. Let $m \equiv \sum_k M_k / \sum L_k^p$, let W_i be the real wage in country i under frictionless trade and no MP, and let W_i^* be the real wage in country i under frictionless trade and MP and define $l_i \equiv L_i^p / \sum_k L_k^p$ and $m_i \equiv T_i^e L_i^e / L_i^p$. We first characterize the expressions for welfare under restricted entry in the following Lemma, which we prove in the online appendix.

Lemma 4 *Under restricted entry, consider a world where $\widetilde{A}_i = \widetilde{A}$, $T_i^e = T^e$ for all i, and where $\frac{m_i}{m} < \frac{(\theta+1)(\theta\sigma-\sigma+1)}{(\theta-\sigma+1)}$ for all i, and assume $\rho \to 1$. The ratio of the real wage under frictionless trade and MP to the real wage under free trade and no MP, $\mathcal{W}_i \equiv W_i^*/W_i$, is given by the expression*

$$\left(\mathcal{W}_i\right)^\theta = \frac{\left[(1-\eta)\, m + \eta m_i\right]^\upsilon m^{1-\upsilon}}{m_i^{\theta/(1+\theta)} \sum_k m_k^{1/(1+\theta)} l_k} , \tag{B.20}$$

where $\upsilon \equiv \theta/(\sigma-1) - 1$.

With the help of this Lemma, we can now proceed to prove the two parts of the proposition. Notice that around symmetry $m_i \simeq m$, the restriction in the lemma is always satisfied, so that we can make use of the lemma for proving Proposition 4.

Part (i) We first show that real wages increase iff $\sigma < \bar{\theta} \equiv \frac{(1+\theta)^2}{1+\theta+\theta^2}$. To do that, we will use the above lemma. Taking logs of (B.20), differentiating with respect to the size of one country m_i, and evaluating at symmetry, $m_i = m$ for all i, we get that the sign of this derivative is determined by

$$\upsilon\left[(1-\eta)\, l_i + \eta\right] + (1-\upsilon)\, l_i - \frac{\theta}{1+\theta} - \frac{1}{1+\theta} l_i ,$$

or equivalently, by the sign of $\upsilon\eta - \frac{\theta}{1+\theta}$. The condition $\upsilon\eta > \frac{\theta}{1+\theta}$ is equivalent to $\sigma < \frac{(1+\theta)^2}{1+\theta+\theta^2} \equiv \bar{\theta}$, which proves part i).

Part (ii) Now consider real expenditures. Defining $W_i \equiv w_i/P_i$, with no MP we have

$$X_i/P_i = \frac{W_i}{1-\eta},$$

whereas with frictionless trade and MP we have

$$X_i^*/P_i^* = \frac{X_i}{Y_i}\frac{W_i^*}{1-\eta}.$$

Consider the ratio, $\mathcal{X}_i \equiv \left(X_i^*/P_i^*\right)/\left(X_i/P_i\right)$, and solve for X_i/Y_i (see Online Appendix),

$$\mathcal{X}_i = \left(1-\eta+\eta\frac{m_i}{m}\right)\frac{W_i^*}{W_i},$$

and hence, using (B.20),

$$\mathcal{X}_i = \left(\frac{\left[(1-\eta)\,m+\eta m_i\right]^{\upsilon+\theta} m^{1-\upsilon+\theta}}{m_i^{\theta/(1+\theta)}\sum_k m_k^{1/(1+\theta)} l_k}\right)^{1/\theta}.$$

This expression is similar to what we had above for real wages, only that instead of υ we now have $\upsilon+\theta$. Thus, the condition for real income to increase is that $(\upsilon+\theta)\,\eta > \frac{\theta}{1+\theta}$. Notice that this condition is equivalent to $\theta > \sigma-1$, which we always require for the various integrals to have a finite mean. Thus, real expenditure must increase with MP.

In a similar manner we can show that real profits increase under frictionless trade and MP. Profits in country i are $\Pi_i = X_i - w_i L_i^p$, so that real profits per person of country i are $\pi_i = x_i - W_i$. With frictionless trade and no MP we have $x_i = W_i/(1-\eta)$, so that $\pi_i = W_i\eta/(1-\eta)$, while with frictionless trade and MP we have $\pi_i^* = W_i^*\frac{m_i}{m}\eta/(1-\eta)$. This implies that

$$\frac{\pi_i^*}{\pi_i} = \frac{m_i}{m}\frac{W_i^*}{W_i},$$

and hence,

$$\frac{\pi_i^*}{\pi_i} = \frac{\left[(1-\eta)\,m+\eta m_i\right]^{\upsilon} m^{-\upsilon}}{m_i^{-\frac{1}{1+\theta}}\sum_k m_k^{1/(1+\theta)} l_k}.$$

Taking logs, differentiating, and evaluating at symmetry, we obtain

$$\frac{1}{m}(1-l_i)\left[\upsilon\eta+\frac{1}{1+\theta}\right],$$

which is always positive, implying that real profits are higher with free MP versus no MP; **QED**.

B.9 Proof of Proposition 5

To prove Proposition 5 we first compute the real wage under two scenarios: (i) frictionless trade and frictionless MP; and (ii) frictionless trade but no MP. Then we will compare the two.

(i) Frictionless trade and frictionless MP. From $(B.8)$ and $w_N = 1$ we get

$$w_n = T_n^e/T_N^e\,. \tag{B.21}$$

Using (B.10), which holds in the case of frictionless trade and MP, together with (B.6), (14) and (B.21), to replace into the price index (equation $B.5$), we obtain the real wage in country

n under frictionless trade and MP,

$$\frac{w_n}{P_n} = \zeta \eta^{1/\theta} T_n^e / f^e \left[\left(\frac{F_n}{L_n} \right)^{(\theta-\sigma+1)/(1-\sigma)} \left\{ \sum_k \left[T_k^p \left(T_k^e / f^e \right)^{-\theta} \right]^{1/(1-\rho)} \right\}^{1-\rho} \left(\sum_k \frac{L_k T_k^e}{f^e} \right) \right]^{1/\theta}. \tag{B.22}$$

(ii) Frictionless trade but no MP. Given that there is no MP, trade is balanced so that $X_n = Y_n$ and $L_n^e = \eta L_n$ for all n. Therefore the current account balance in (14) together with the fact that all income is accrued to labor, $X_n = w_n L_n$, and $L_n^e = \eta L_n$ imply $w_n L_n = \sum_k \lambda_{nk}^E X_k$. But since there is frictionless trade but no MP, then by replacing for the definition of λ_{in}^E, the current account balance can be written as

$$w_n L_n = \frac{M_n T_n^e T_n^p w_n^{-\theta}}{\sum_k M_k T_k^e T_k^p w_k^{-\theta}} \sum_k X_k.$$

Choosing country N labor as the numeraire, and using $M_n = r_n L_n / f^e$ with $r_n = \eta$, the above expression implies that wages can be expressed as

$$w_n = \left(\frac{T_n^e T_n^p}{T_N^e T_N^p} \right)^{\frac{1}{1+\theta}}. \tag{B.23}$$

Also, note that by using $(B.23)$ and the fact that in this case $\Psi_{in} = T_i^e T_i^p w_i^{-\theta}$ we have that

$$\sum_k M_k \Psi_{kn} = \eta \left(T_N^e T_N^p / f^e \right)^{\frac{\theta}{1+\theta}} \sum_k L_k \left(T_k^e T_k^p / f^e \right)^{\frac{1}{1+\theta}}.$$

Finally, we get the real wage by substituting the above relationship and $X_n = w_n L_n$ into the price index (equation $B.5$), and using (B.23),

$$\frac{w_n}{P_n} = \zeta \eta^{1/\theta} \left[\left(\frac{F_n}{L_n} \right)^{\frac{\theta-\sigma+1}{1-\sigma}} \sum_k L_k \left(T_k^e T_k^p / f^e \right)^{\frac{1}{1+\theta}} \right]^{1/\theta} \left(T_n^e T_n^p / f^e \right)^{\frac{1}{1+\theta}} \tag{B.24}$$

Comparison. To prove our result we simply need to show that (B.22) is larger than (B.24), or equivalently,

$$\left\{ \sum_k \left[T_k^p \left(\frac{T_k^e}{f^e} \right)^{-\theta} \right]^{1/(1-\rho)} \right\}^{1-\rho} \geq \left[T_n^p \left(\frac{T_n^e}{f^e} \right)^{-\theta} \right]^{\frac{\theta}{1+\theta}} \sum_j \frac{L_j T_j^e}{\sum_k L_k T_k^e} \left[T_j^p \left(\frac{T_j^e}{f^e} \right)^{-\theta} \right]^{\frac{1}{1+\theta}}. \tag{B.25}$$

Note that the right-hand side of this expression is less than or equal to $\max_k T_k^p \left(\frac{T_k^e}{f^e} \right)^{-\theta}$. We can then write the inequality as,

$$\sum_k \left[T_k^p \left(\frac{T_k^e}{f^e} \right)^{-\theta} \right]^{1/(1-\rho)} \geq \left[\max_k T_k^p \left(\frac{T_k^e}{f^e} \right)^{-\theta} \right]^{1/(1-\rho)},$$

which is always true. **QED.**

B.10 Algorithm for Calibration and Simulations

The algorithm for calibration is divided in three steps explained below.

Step 1. Given a value for θ and ρ we first force the model to perfect match the trade and MP shares, λ_{ln}^T and λ_{il}^M, by using equations (10), (8), and picking $\widetilde{\gamma}_{il} \equiv T_{il} \gamma_{il}^{-\theta}$ and $\widetilde{\tau}_{ln} \equiv$

$\tau_{ln}^{-\theta}$.[36] By fitting these data we can construct a unique set of trilateral flows in the model as (after some algebra)

$$X_{iln} = \frac{\frac{X_l \lambda_{il}^T}{\sum_n X_n \lambda_{ln}^T} \hat{\tau}_{ln}}{\sum_k \left[\sum_j \left(\frac{X_j \lambda_{kj}^T}{\sum_n X_n \lambda_{jn}^T} \hat{\tau}_{jn} \right)^{1/(1-\rho)} \right]^{1-\rho}} X_n$$

where we normalized $\sum_{j'} \hat{\gamma}_{j'j} = 1$ and $\hat{\tau}_{ln}$ solves

$$\frac{\hat{\tau}_{ln}/\hat{\Psi}_n}{\sum_j \hat{\tau}_{lj}/\hat{\Psi}_j} = \frac{X_n \lambda_{ln}^T}{\sum_i X_l \lambda_{il}^T}.$$

where

$$\hat{\Psi}_n = X_n^{-1} \sum_i \left(\sum_l \left(\frac{X_l \lambda_{il}^T}{\sum_n X_n \lambda_{ln}^T} \hat{\tau}_{ln} \right)^{1/(1-\rho)} \right)^{1-\rho}.$$

Using the solution for X_{iln}, equation (B.4), we relate trilateral flows to trade and MP costs:

$$\frac{X_{iln} X_{inl}}{X_{inn} X_{ill}} = \left(\frac{\tau_{ln}}{\tau_{nn}} \frac{\tau_{nl}}{\tau_{ll}} \right)^{-\frac{\theta}{1-\rho}}.$$

Imposing symmetry and no internal frictions, we have

$$\tau_{ln} = \tau_{nl} = \left(\sqrt{\frac{X_{inn} X_{ill}}{X_{iln} X_{inl}}} \right)^{\frac{1-\rho}{\theta}}.$$

For γ's, imposing symmetry we obtain a similar expression

$$\gamma_{il} = \gamma_{li} = \left(\sqrt{\frac{X_{iin} X_{lln}}{X_{iln} X_{lin}}} \right)^{\frac{1-\rho}{\theta}}.$$

Step 2. We set the parameter T_i^e to match r_i calculated using (16) and data on bilateral trade and absorption (the variable Y is calculated directly from the data using the model condition $Y_l = \sum_n \lambda_{ln}^T X_n$).

Step 3. To finalize the calibration of the model, we iterate on the value of the parameter θ such that we match the "unrestricted" gravity coefficient in the data, by running OLS in the simulated data,

$$\log \lambda_{ln}^T = \beta_u \log \tau_{ln} + D_n + S_l + u_{ln}. \tag{B.26}$$

When we use the Poisson PML estimator of the unrestricted gravity coefficient, we run the corresponding procedure with the simulated data.

The procedure described so far applies to the linear model. For the exogenous entry model we the values for θ, ρ, $\{\tau_{ln}\}$ and $\{\widetilde{\gamma}_{il}\}$ calibrated above and in addition set $L_i^e = r_i L_i$ and $L_i^p = (1 - r_i) L_i$. It is trivial to show that the exogenous entry model with these parameters yields the same equilibrium as the endogenous entry model calibrated above. A similar argument applies to the Roy model (see Appendix B.11 and the online appendix).

[36] We assume $\tau_{ii} = 1$ and $\gamma_{ii} = 1$ for all i for this procedure. Whereas this inverse mapping from data to trade and MP costs is numerically implemented, a proof that the inversion explicitly maps to model fundamentals for a trade model (i.e when MP costs are infinite in our setup) is provided in XXXUniversal GravityXX

B.11 The Roy model

The setup of our Roy model extension is the same as the one in the baseline model. The difference lies in that there is workers are heterogeneous regarding their abilities in innovation and production/marketing. These workers can choose to work in the innovation sector, where the wage per efficiency unit is w_i^e, or the production sector, where the wage per efficiency unit is w_i. To model worker heterogeneity, we assume that workers independently draw a pair of abilities, v^e and v^p, corresponding to the innovation and production sector, from a Fréchet distribution, $\exp\left[-v^{-\kappa}\right]$, with $\kappa > 1$, similar to Lagakos & Waugh (2013) and Hsieh, Hurst, Jones & Klenow (2011). The final productivity of a worker in the production and innovation sectors are $A_i v^p$ and $A_i v^e$, where A_i is a neutral productivity variable that we introduce so that the initial equilibrium is consistent with both the linear model and the Roy model after a suitable choice of A_i as explained in Appendix B.10. A worker with ability pair (v^e, v^p) would work in the innovation sector if and only if $v^e w_i^e \geq v^p w_i$. From the properties of the Fréchet distribution, it is straightforward to show that the supply of efficiency units to the innovation and production/marketing sectors, in country i, are given by

$$L_i^e = \gamma L_i A_i \left[1 + \left(\frac{w_i^e}{w_i}\right)^{-\kappa}\right]^{1/\kappa - 1}, \tag{B.27}$$

and

$$L_i^p = \gamma L_i A_i \left[1 + \left(\frac{w_i^e}{w_i}\right)^{\kappa}\right]^{1/\kappa - 1}, \tag{B.28}$$

respectively, where γ is some positive constant.

The equilibrium levels of $\boldsymbol{X}$, $\boldsymbol{M}$, $\boldsymbol{w}^e$, $\boldsymbol{w}$ are determined solved using (13) by setting $L_i - M_i f^e = L_i^p$, and using $(B.28)$, (14) (15), and by the labor market clearing in innovation, $M_i f^e = L_i^e$, with L_i^e replaced by $(B.27)$.

The calibration of the Roy model is a simple extension of the one done for the linear model. We simply choose A_i so that, together with the values for θ, ρ, $\{\tau_{ln}\}$ and $\{\widetilde{\gamma}_{il}\}$ calibrated in Section 4, the Roy model yields the same equilibrium as the linear model.