NBER WORKING PAPER SERIES

INNOVATION AND PRODUCTION IN THE GLOBAL ECONOMY

Costas Arkolakis
Natalia Ramondo
Andrés Rodríguez-Clare
Stephen Yeaple

Working Paper 18972
http://www.nber.org/papers/w18972

NATIONAL BUREAU OF ECONOMIC RESEARCH
1050 Massachusetts Avenue
Cambridge, MA 02138
April 2013

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. 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. The views expressed herein are those of the authors and do not necessarily reflect the views of the National Bureau of Economic Research.

Innovation and Production in the Global Economy
Costas Arkolakis, Natalia Ramondo, Andrés Rodríguez-Clare, and Stephen Yeaple
NBER Working Paper No. 18972
April 2013
JEL No. F1,F23,F6

ABSTRACT

One feature of globalization is that countries are increasingly specialized in either innovation or in production. To understand the forces behind this specialization and its welfare consequences, we develop a monopolistic competition model of trade and multinational production (MP) in which firms face a tradeoff between producing close to customers and producing in the least-cost location. At the country level, the location of innovation and production is determined by comparative advantage and home market effects that arise from the interaction of trade and MP costs with increasing returns to scale. The model yields simple structural expressions for bilateral trade and MP that we use to calibrate the model across a set of OECD countries. Our counterfactual exercises shed light on the effect of falling MP costs, and the entry of China into world markets, on welfare between and within countries.

Costas Arkolakis
Department of Economics
Yale University, 28 Hillhouse Avenue
P.O. Box 208268
New Haven, CT 06520-8268
and NBER
costas.arkolakis@yale.edu

Natalia Ramondo
Arizona State University
Department of Economics
POB 879801
Tempe, AZ 85287-9801
Natalia.Ramondo@asu.edu

Andrés Rodríguez-Clare
University of California at Berkeley
Department of Economics
Berkeley, CA 94720-3880
and NBER
andres1000@gmail.com

Stephen Yeaple
Department of Economics
The Pennsylvania State University
520 Kern Building
University Park, PA 16802-3306
and NBER
sry3@psu.edu

1 Introduction

One consequence of globalization is that goods are increasingly being produced far from where ideas are created. This international specialization in innovation or 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 also become more pronounced. 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%.[1]

This phenomenon raises a set of important questions. How does the increasing ability to locate production abroad affect the geography of innovation and national welfare? Do some countries gain more than others? Could some countries be made worse off? Are workers in some countries harmed in the process? To tackle these questions, we develop a quantitative, multi-country general equilibrium model that builds on the established theory of international trade, but that allows firms to separate innovation and production geographically.

Following Melitz (2003), we model innovation as the creation of heterogeneous firms that sell differentiated goods in monopolistically competitive markets that are 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 with the productivity levels across locations drawn from a multivariate distribution. In deciding where to produce to serve a particular market, firms face a trade off between being close to their customers to avoid trade costs and producing in the country where they would achieve the minimum unit cost. By allowing firms to produce outside of their home country, multinational production (MP) allows some countries to specialize in innovation and others to specialize in production.[2] Countries that specialize in innovation have a net inflow of profits that compensates for the cost of innovation; loosely speaking, these countries export ideas and import goods.

There are two forces that determine the allocation of innovation across countries: First, countries that have a high productivity in innovation relative to production tend to specialize

[1]Among the fastest growing destinations for the foreign affiliates of U.S. multinationals is China, which now accounts for one in eight employees of the foreign affiliates of U.S. firms.

[2]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 and Rodríguez-Clare (2008), where entry is endogenous, but not affected by trade costs. An equivalent result is derived by Eaton and Kortum (2001) in a setting with Bertrand competition.

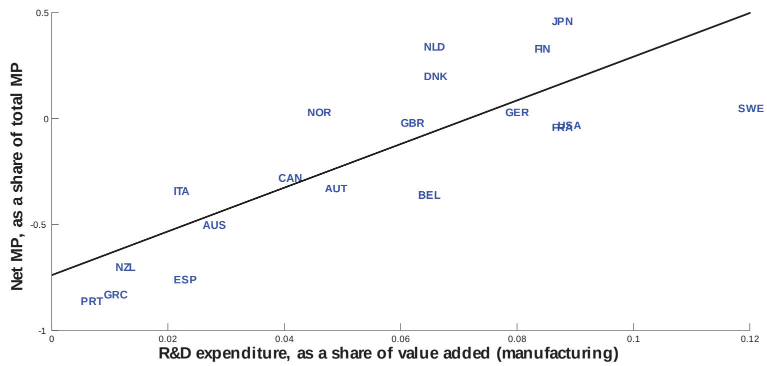

Figure 1: R&D and Net MP

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. Data construction for MP is described in the appendix.

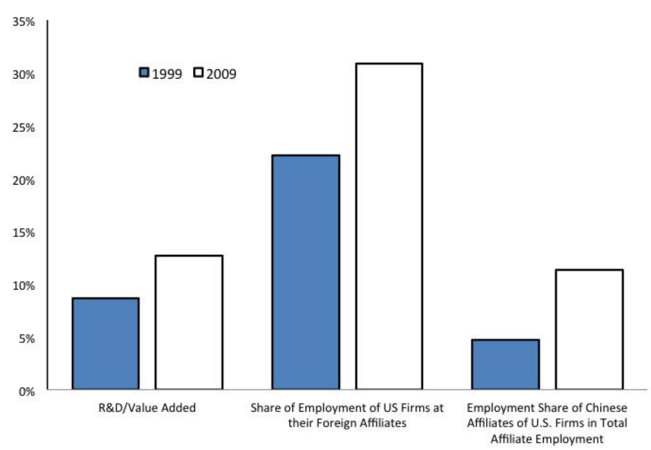

Figure 2: R&D and Employment of U.S. manufacturing firms and multinationals

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.

in innovation; and second, home market effects (HME) imply that country size and location affect the allocation of production and innovation. HMEs lead production to concentrate in countries with large "market potential," while they draw innovation towards countries with large "production potential," i.e., countries that have a large labor force or that are well connected to such countries. We can think of the first of these forces as a "comparative advantage in innovation" while the second force is related to the proximity to consumers (for production) and workers (for innovation).

The model provides a natural environment in which to explore the implications of MP on real wages. This turns out to depend critically on the assumptions regarding the ability of workers to move between production and innovation sectors. In the theory section we first consider a "specific-factors" version of the model in which there is no labor mobility between sectors. In this case we find a result that resonates with the popular view on the effect of MP on wages: by moving production abroad, multinationals can lower the real wage of production workers in countries that have a comparative advantage in innovation. In contrast, under perfect labor mobility between innovation and production and with frictionless trade, we show that, starting from a scenario with no MP, eliminating all barriers to MP necessarily leads to an increase in real wages everywhere. This strong result depends, however, on the assumption that trade is frictionless and all barriers to MP are eliminated. For example, a unilateral reduction in inward MP costs can make a country worse off, because of a resulting contraction of innovation and expansion of production that worsen the country's terms of trade.

To arrive at more precise conclusions regarding the impact of MP on the geography of innovation and welfare in our multi-country general equilibrium setting, we calibrate the model to data for 18 OECD countries and then perform a series of counterfactual exercises. We start by deriving a novel implication of our model, namely that trade flows by 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 a gravity equation that partially identifies key structural parameters of our model. We then re-estimate the same equation on overall trade flows, obtaining an additional target for our calibration and confirming the model's implication regarding relative trade cost sensitivities. The rest of the calibration is very parsimonious: we avoid adding extraneous elements to the theory and simply find the trade and MP costs that make the model perfectly fit the bilateral trade and MP flows in the data.

As mentioned above, the effect of MP on the location of innovation and on welfare depends on the assumptions regarding labor mobility between innovation and production. For the quantitative analysis, we focus on a case that is easy to calibrate and yet combines features of the two extreme cases considered in the theoretical analysis. In particular, we assume that, as in Roy (1951), there is perfect labor-mobility across sectors but workers are heterogeneous in their abilities in innovation and production.

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 results in greater specialization across countries in innovation and production and real incomes rise on average by about 2%. Perhaps surprisingly, production workers gain everywhere, and it is workers employed in innovation who experience losses in some countries. Countries that are initially less innovative, such as the southern European countries bordering the Mediterranean Sea, are those suffering a decline in innovation and in the real income of workers employed in innovation. Second, we explore the implications of the integration of China to the world economy. In this exercise, developed countries are induced to follow the "Apple model:" they specialize in innovation and China becomes a key manufacturing center for the whole world. Contrary to popular fears, production workers in all countries gain, although workers initially employed in innovation benefit by more than those employed in production.

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 in innovation and production, which leads to a tendency toward 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 model that incorporates the most important mechanisms found in this earlier work.

Our model provides a strict generalization of the Melitz (2003) model of trade as specified by Chaney (2008). In particular, when MP costs go to infinity, our model collapses to a general equilibrium version of that model with endogenous entry (as in Arkolakis et al.

[3]Examples of work that most closely resembles our own are Markusen and Venables (1998) and Markusen and 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 and 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.

(2008)). Our approach has significant differences with another strict generalization of the Melitz model that allows for MP, namely Helpman, Melitz and Yeaple (2004) (henceforth HMY). First, our general equilibrium approach allows us to think about how comparative advantage and home market effects lead some countries to specialize in innovation and exhibit net outward MP, while others specialize in production and exhibit net inward MP. Second, our probabilistic approach to the modeling of productivity across multiple production locations makes it easy to capture how multinational affiliates use certain countries as export platforms, while this would lead to severe computational problems in HMY. Finally, we use our calibrated model to perform counterfactual analysis and provide quantitative answers to the questions raised above regarding the effect of MP on welfare and the location of innovation.

One potential drawback of our approach relative to HMY (and the quantitative application of their framework by Irarrazabal, Moxnes and Opromolla (2012)) is that we do not allow for fixed costs of running foreign affiliates. Thus, our model does not have a proximity-concentration trade-off. This simplification allows us to avoid a very complex discrete-choice problem and gives us the necessary tractability to handle export platforms and international specialization in innovation and production. A recent paper by Tintelnot (2012) develops a model that allows for export platforms with fixed costs of production. His paper, however, abstracts from entry and so is not equipped to analyze international specialization in innovation and production, which is the focus of our paper.

A close relative to our model is Ramondo and Rodríguez-Clare (2013), which extends the perfect competition Ricardian framework of Eaton and Kortum (2002) to allow for MP (and "intra-firm" trade). Whereas both models have similar expressions for aggregate trade and MP flows, the perfectly competitive framework of Ramondo and Rodríguez-Clare (2013) does not allow for profits and innovation, which play critical roles in our analysis.[4] Also related are Eaton and Kortum (2007) and Rodríguez-Clare (2010), which develop theories to show how technology diffusion or offshoring may lead countries to specialize in innovation or production.

Burstein and Monge-Naranjo (2009), McGrattan and Prescott (2010), and McGrattan (2011) also study the impact of MP on welfare. These papers extend the neoclassical growth model by introducing a rival "managerial capital" or a non-rival "knowledge capital" that can be used in any location. The movement of managerial or knowledge capital from one country to another is interpreted as MP, while trade takes place only as a way to transfer the returns to capital. We think of our approaches as complementary: While our model can more eas-

[4]Another relevant contribution is Ramondo (2012), who adapts the Eaton and Kortum (2002) framework to model MP (without trade). This model also has perfect competition, so there are no profits and no innovation.

ily connect to the trade and MP data, the simplification of the trade dimension in the cited 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.

Finally, our model is well suited to analyze the impact of globalization (and MP in particular) on within-country income distribution. 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 and Hanson (1999); and Grossman and Rossi-Hansberg (2008)). 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 depressing effect of Chinese manufacturing exports on the employment and wages of developed country manufacturing workers (e.g. Autor, Dorn and Hanson (2012)).

Before proceeding, one conceptual issue is worth noting. In this paper we focus on multinational production as the vehicle through which international specialization takes place, but there are alternative arrangements, such as the licensing of technology **and** other contractual relationships that do not involve ownership (outsourcing). Our model is consistent with all of these mechanisms, but because there is little data on arm's length offshoring we can only measure the offshoring done within multinationals. There is a possibility that relying on only MP data might create a distorted view of international specialization, but our focus on a sample of similarly developed countries reduces this concern (see e.g. Antras (2003)).[5]

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

[5]When we introduce China to our quantitative exercises, we do not calibrate to MP flows.

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.[6]

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 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 $\boldsymbol{z}$, 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

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

is perfectly inelastic.[7] 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 $\boldsymbol{z}$, 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 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).

[7] 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.

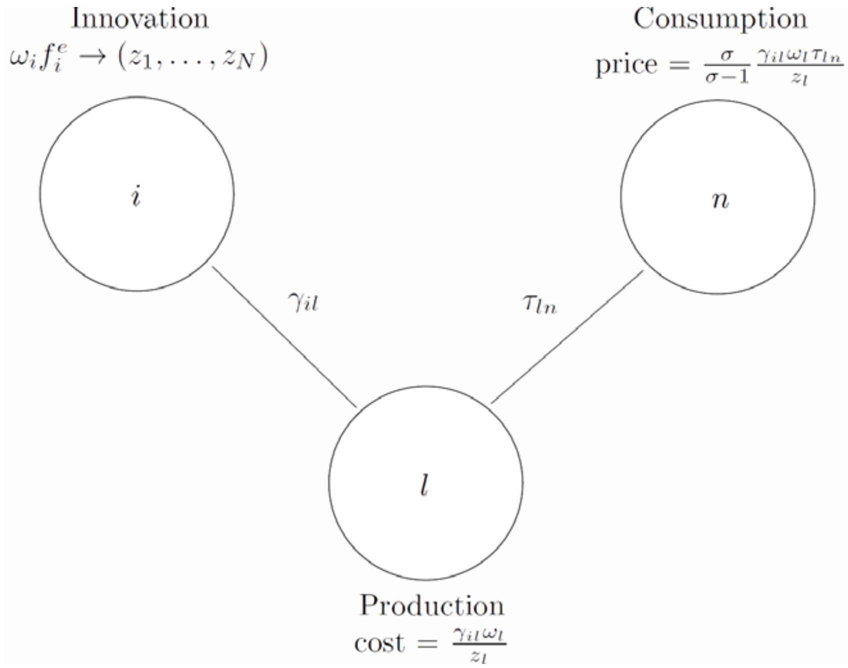

Figure 3: Firm Pricing
Note: Innovation is done in country i at cost $w_i f_i^e$, production is done in country l at unit cost $\gamma_{il} w_l / z_l$, and consumption is done in country n at price $\tilde{\sigma}\gamma_{il} w_l \tau_{ln} / z_l$.

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 $\theta > \max(1, \sigma - 1)$.[8] 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

[8] 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.

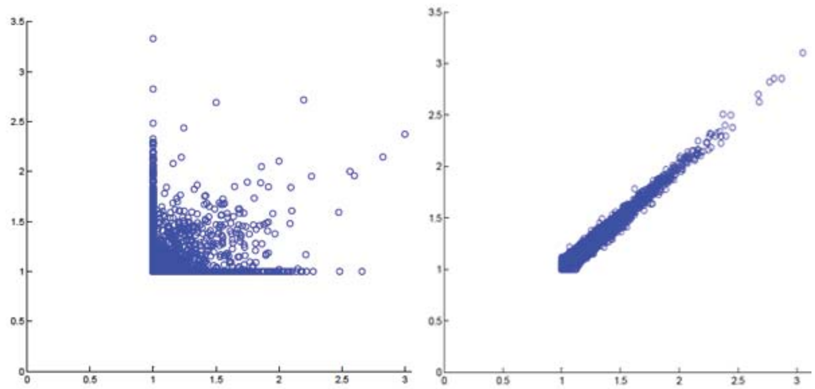

Figure 4: Simulated draws from multivariate Pareto with $\rho = 0.1$ (left panel) $\rho = 0.9$ (right panel)
Note: Simulation for 10,000 draws with distribution specified with $N = 2$, $T_1 = T_2 = 2^{\rho-1}$, $\theta = 7.2$.

$1 - \widetilde{T}_i z_l^{-\theta}$ 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*.[9] We will continue to write T_{il} rather 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.

[9]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).

Lemma 1 *Let* $\Psi_{in} \equiv \left[\sum_k \left(T_{ik}\xi_{ikn}^{-\theta}\right)^{\frac{1}{1-\rho}}\right]^{1-\rho}$ *and* $\psi_{iln} \equiv \left(T_{il}\xi_{iln}^{-\theta}/\Psi_{in}\right)^{\frac{1}{1-\rho}}$. *The (unconditional) probability that a firm from i will serve market n from l at cost c, for* $c \leq c_n^*$, *is*

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

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 in country i. The result in (9) implies that a share $\frac{\theta-\sigma+1}{\sigma\theta}$ of output produced in any country is used to pay for the marketing of that output, hence a share $1-\eta-\frac{\theta-\sigma+1}{\sigma\theta}$ of output is left to pay for production labor. We then have

$$L_i^p = \left(1-\eta-\frac{\theta-\sigma+1}{\sigma\theta}\right)\frac{Y_i}{w_i} + \sum_{j,l} M_{jli} F_i.$$

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

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

Next, we characterize the current account balance condition.[10] 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 con-

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

dition 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

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

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 $w^e_i L^e_i$, where w^e_i is the wage paid to workers in the innovation sector. Using (12) and noting that $L^e_i = M_i f^e$, the zero-profit condition is

$$\eta \sum_n \lambda^E_{in} X_n = M_i w^e_i f^e. \tag{15}$$

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 $M_i = L^e_i / f^e$ (where L^e_i is exogenous). The wage in the innovation sector, w^e_i, is recovered from (15). Under endogenous entry, perfect labor mobility between production and innovation implies that w^e_i must be equal to w_i.[11] Equations (13) , (14) and (15) with w^e_i 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 sector (or simply the innovation share), $r_i \equiv w^e_i L^e_i / \left(w^e_i L^e_i + w_i L^p_i \right)$, which is also equal to $r_i = \sum_{l,n} \Pi_{iln} / X_i$. The equilibrium conditions above imply that

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

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

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

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 consistent with the results in the version of the Melitz model presented in Arkolakis et al. (2008), where entry is endogenous but not affected by trade costs.[12]

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,*

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

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

The share of labor devoted to innovation in country i is

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

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.[13] 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 of an 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.

[13]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 and 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).

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 and Kortum (2002), and Arkolakis, Costinot and 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 $\left(X_{nnn}/X_n\right)^{-1/\theta}$. 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 and 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).[14] 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 and 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. As in Ramondo and Rodríguez-Clare (2013), 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 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

[14]In Ramondo and 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.

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?

Unfortunately, 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.[15] 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, $GMP_i < 1$.[16]

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 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 and Rodríguez-Clare (2013)) which may not be enough to compensate for the loses arising from the fall in innovation.

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]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 $\gamma_{21} = 4$.

[16]A numerical example as in the previous footnote reveals that these results extend to unilateral and multilateral trade liberalization, so that we can have $GT_i < 1$ for some countries if home-market effects push r_i below η.

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

$$\ln X_{iln} = \alpha^r_{il} + \mu^r_{in} - \frac{\theta}{1-\rho} \ln \tau_{ln}, \tag{22}$$

where α^r_{il} and μ^r_{in} are fixed effects.[17] 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 $-\theta/\left(1-\rho\right)$.

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.[18] 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

[17]Given i, the fixed effect captured by α^r_{il} varies by location of production and corresponds (in the model) to $\alpha^r_{il} = \ln\left(M_i\left[T^e_i T^p_l (w_l \gamma_{il})^{-\theta}\right]^{\frac{1}{1-\rho}}\right)$, while the fixed effect captured by μ^r_{in} varies by country of destination and corresponds (in the model) to $\mu^r_{in} = \ln\left(X_n \Psi_{in}^{\frac{-\rho}{1-\rho}} / \sum_k M_k \Psi_{kn}\right)$.

[18]This feature of the gravity model of trade led Eaton and Kortum (2002), Donaldson (2012), and Simonovska and Waugh (2009) 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 Data Appendix.

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).[19]

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

[19] 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.[20] 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. 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 observations. The trade elasticity in the restricted regression of 10.9 is our estimate of $\theta/(1 - \rho)$.

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

	Tariff	Distance Dummies D2	D3	D4	D5	D6	Other Gravity Controls Self	Border	Lang.	R-sq.
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.

	Tariff	Distance Dummies D2	D3	D4	D5	D6	Other Gravity Controls 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 2: Restricted and Unrestricted Gravity: Poisson PML Estimator

Notes: Robust standard errors in parenthesis.

The trade elasticity in the unrestricted regression is significantly smaller at 4.3, as expected.

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 and 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 and Tenreyro (2006) and report the coefficient estimates in Table 2. 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.

The coefficient estimates of 4.3 and 5.4 obtained from the unrestricted gravity estimation are in the ballpark of estimates of the trade elasticity obtained by the trade literature. Anderson and 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 and Parro (2010), fall within the same range. Simonovska and Waugh (2009), using price differentials, and a refinement of the methodology of Eaton and Kortum (2002), bring these estimates closer to 4.

4 Calibration

We restrict our analysis to 18 OECD countries for which we have good data for both trade and MP.[21] We use STAN data on manufacturing trade flows from country l to country n as the empirical counterpart of bilateral trade in the model, and STAN production data minus aggregate exports to compute home sales. 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 UNCTAD data on the gross value of production for multinational affiliates from country i in country l to construct the empirical counterpart of bilateral MP flows and obtain an $N \times N$ matrix of production shares, λ_{il}^{M}.[22] Finally, we measure the $N \times 1$ vector of labor endowments, L_i, as equipped labor, from Klenow and Rodríguez-Clare (2005), adjusted 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 estimated above in Table 1. To determine the exact levels of θ and ρ, we use the predictions of the model regarding the unrestricted gravity regression coefficient. Computing that same elasticity in the model, however, requires to calculate the model's equilibrium and generate a data set of trade flows and trade costs, a procedure that we describe below. For $\theta = 4$ and $\rho = 0.63$, which satisfy $\theta / (1 - \rho) = 10.9$, the model predicts an OLS unrestricted trade elasticity of -4.3. As an alternative calibration, we match the Poisson PML estimators: the coefficients of 8.3 and 5.4, for restricted and unrestricted gravity, respectively. This yields $\theta = 4.95$ and $\rho = 0.40$.

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 and Pilat (1996) are around 20 percent, while estimates by Domowitz, Hubbard and 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.[23]

[21]The sample of countries is the same as in Eaton and Kortum (2002) and Ramondo and Rodríguez-Clare (2013). It includes Australia, Austria, Benelux (Belgium, Luxembourg, and Netherlands aggregated as one country), Canada, Denmark, Spain, Finland, France, United Kingdom, Germany, Greece, Italy, Japan, Norway, New Zealand, Portugal, Sweden, and the United States.

[22]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.

[23]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 calibration of the rest of the parameters proceeds in three steps. The first step of the calibration algorithm uses our constructed trade and MP shares as well as data on L_i and X_i together with the equilibrium conditions of the model with endogenous entry to construct output, Y_i, wages, w_i, innovation shares, r_i, and implied current account deficits, which we label as Δ_i. We need to allow for current account deficits in order for the model to match the data. Table 3 presents the data used and the results in this step of the calibration.

Given the aggregate variables calculated in the first step and the values for σ and $\theta/(1-\rho)$, the second step of the algorithm searches over two matrices of parameters, $\widetilde{\gamma}_{il} \equiv T_{il}^{-1/\theta}\gamma_{il}$ and τ_{ln}, to exactly match the data on trade and MP shares. The fact that $T_{il}^{-1/\theta}$ and γ_{il} matter through their product implies that our procedure is not able to separately identify technology parameters from MP costs, hence we cannot determine whether it is comparative advantage or home market effects that determine the innovation patterns across countries. Intuitively, a high observed innovation rate r_i could be explained by the model either as a consequence of a high T_i^e relative to $T_i^p/L_i^{1-\rho}$ (as in Proposition 1), or favorable size and trade/MP costs (as in Propositions 2 and 3). To proceed, we impose $T_i^e = 1$ and $T_i^p = (L_i)^{1-\rho}$ for all i. Together with $\gamma_{ll} = \tau_{ll} = 1$ for all l, this assumption implies that $\widetilde{\gamma}_{ll} = L_l^{-(1-\rho)/\theta}$, which affects the level of the calibrated $\widetilde{\gamma}_{il}$ and τ_{ln}. However, the implied trilateral flows (X_{iln}) and the results of the counterfactual exercises performed with the calibrated model for a percentage change in parameters γ_{il} and τ_{ln}, are not affected by this assumption, as we formally discuss in Appendix B.10.[24]

The final step of the algorithm is to run an unrestricted gravity regression in which the dependent variable are the simulated trade shares from l to n and the regressors are the trade cost from l to n calibrated in the previous step, and two sets of country fixed effects.

It is useful to study the implications of the model for average outward and inward trade and MP costs, weighted by the size of the partner countries. Average outward MP costs are informative about a country's ability to take its ideas abroad while inward MP costs are informative about a country's ability to adopt foreign ideas. Innovation in equilibrium is determined by these costs together with inward and outward trade costs, as shown in Propositions 2 and 3. Figure 5 plots average inward and outward MP and trade costs for our

to 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 and Sichel (2009). Their estimates, however, include a two percent that corresponds to payments for advertising, branding, etc. These costs are closer, in our model, to payments for marketing fixed costs.

[24]The level of real income in our baseline calibration depends on the normalization, but its changes, for given percentage changes in γ_{il} and τ_{ln}, do not. Related to our approach, Burstein and Vogel (2012) use numerical methods to estimate the level of trade frictions using a trade model where gravity relationships do not take a simple analytical form. They also discuss how different restrictions on parameters affect the level of the estimated parameters and counterfactuals, finding little effects by imposing, for example, symmetry on trade costs.

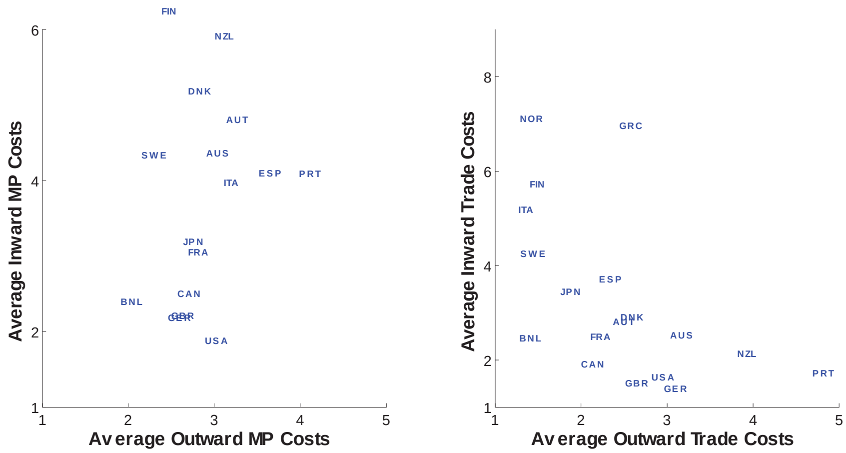

Figure 5: Average Calibrated MP and Trade Costs.
Note: Letting $l_i \equiv L_i / \sum_k L_k$ and $\widetilde{l}_l \equiv l_l / \sum_{n \neq l} l_n$, we compute $\overline{\tau}_l^{OUT} \equiv \sum_{n \neq l} \widetilde{l}_n \tau_{ln}$, $\overline{\tau}_n^{IN} \equiv \sum_{l \neq n} \widetilde{l}_l \tau_{nl}$, $\overline{\gamma}_i^{OUT} \equiv \sum_{l \neq i} \widetilde{l}_l \gamma_{il}$, $\overline{\gamma}_l^{IN} \equiv \sum_{i \neq l} \widetilde{l}_i \gamma_{il}$.

sample of OECD countries.

Highly developed economies, in particular United States, Canada, Germany, Great Britain, France, and Benelux, appear very open in both dimensions, whereas Greece comes out as a very closed economy. Other countries exhibit a more mixed set of costs. For example, Portugal has high outward and low inward MP costs, while Norway has high inward and low outward costs for both trade and MP.

Using the calibrated model, in the next Section we perform several counterfactual exercises to understand the impact of trade and MP on innovation and welfare.[25] Before doing that, in the rest of this section we discuss the forces leading to high innovation rates in a few small countries in the calibrated model, and assess some additional implications of the calibrated model.

4.1 Understanding high innovation rates in small countries

In the data, Benelux, Denmark, Finland and Norway have large outward MP flows relative to their inward MP flows, which goes hand-in-hand with high innovation rates in our model (see Equation 16). Given the strong home market effects present in our model, such high innovation rates in small countries may seem paradoxical. Indeed, if we eliminate differences

[25] In order not to have to take a stand on how current account imbalances change as we move from one equilibrium to the other, we recompute the model's equilibrium imposing current account balance and then perform the counterfactuals. Thus, for example, the gains from openness do not confound the gains coming directly from trade and MP and the gains coming indirectly from the effect of transfers on varieties available for consumption.

	X	L	Y	Δ/X	w	r
Australia	0.04	0.05	0.04	0.14	0.66	0.17
Austria	0.02	0.03	0.02	0.19	0.71	0.14
Benelux	0.08	0.07	0.08	-0.11	1.22	0.25
Canada	0.08	0.10	0.09	0.00	0.88	0.14
Denmark	0.02	0.02	0.01	0.02	0.66	0.21
Spain	0.09	0.08	0.08	0.11	0.97	0.15
Finland	0.02	0.02	0.02	-0.11	1.06	0.21
France	0.17	0.18	0.17	0.01	0.93	0.18
Great Britain	0.16	0.18	0.14	0.12	0.75	0.14
Germany	0.26	0.38	0.28	-0.08	0.74	0.20
Greece	0.02	0.01	0.01	0.24	0.91	0.17
Italy	0.17	0.12	0.17	0.01	1.40	0.17
Japan	0.59	0.52	0.62	-0.06	1.20	0.20
Norway	0.02	0.01	0.01	0.10	1.13	0.22
New Zealand	0.01	0.01	0.01	0.19	0.43	0.07
Portugal	0.02	0.03	0.02	0.21	0.45	0.09
Sweden	0.03	0.03	0.03	-0.08	1.09	0.18
United States	1.00	1.00	0.98	0.02	0.98	0.19

Table 3: Aggregate Variables: Model's Calibration.

Note: The variables X and L refer to absorption and equipped labor in manufacturing, respectively, in the data (relative to U.S.). The variables Y, Δ/X, w, and r, refer to output, current account deficit (as a share of X), wages, and the share of labor in innovation, respectively, as implied by the model's equilibrium.

in trade and MP costs across countries (i.e., set $\tau_{ln} = \tau$ for all $l \neq n$ and $\gamma_{il} = \gamma$ for all $i \neq l$), and set $T_i^e = 1$ and $T_i^p = L_i^{1-\rho}$ for all i, so that the only difference across countries is size (L_i), then innovation rates r_i would be strongly increasing in size L_i thanks to the home market effects discussed in Section 2.4. This point is illustrated in Figure 6, which shows r_i under (i) the calibrated parameters but setting all current account imbalances to zero (from here onwards we refer to this as the "baseline calibration"), and (ii) as in (i) but setting trade and MP costs at the average values in the calibrated model.[26] The correlation between r_i and L_i is 0.10 in the baseline calibration but raises to 0.86 when there are no differences in MP and trade costs across country pairs. Thus, when all country pairs face the same MP frictions, small and centrally located countries, like Benelux and Denmark, lose their advantage as places for innovation, and become attractive places for production.

How does the calibrated model generate the high innovation rates for Benelux (and other small nordic countries) implied by our data? The key force behind Benelux's high r in the calibrated model is the country's low outward MP costs. This cost is on average $\overline{\gamma}_{BEN}^{OUT} = 1.57$, the lowest in our sample of countries – the second lowest average outward MP costs is for

[26]The average of γ_{il} for all $i \neq l$ is 2.5, while the average of τ_{ln} for all $l \neq n$ is 2.7.

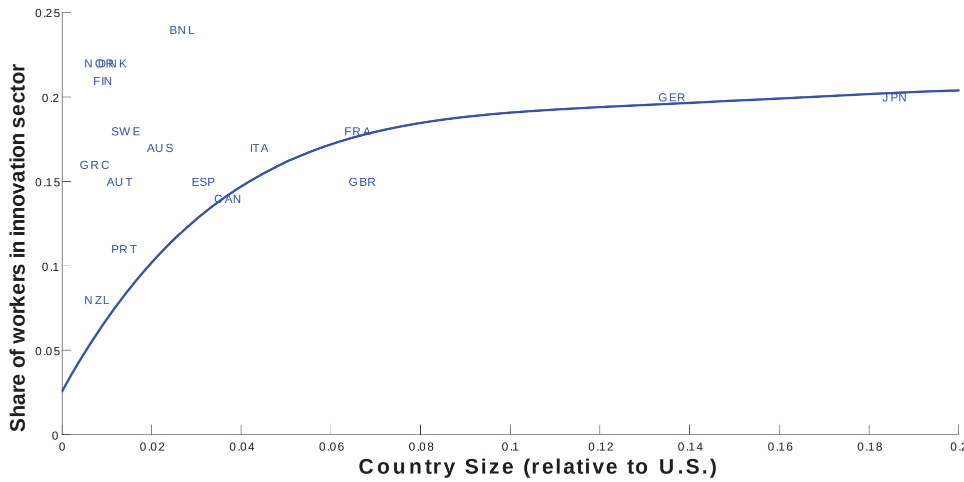

Figure 6: Comparative Advantage and Home Market Effects

Note: r according to baseline calibration (country labels) and according to counterfactual equilibrium with all trade and MP costs set at average values (solid curve).

Finland, with $\overline{\gamma}_{FIN}^{OUT} = 1.98$. The low outward MP cost makes Benelux an attractive location for innovation.[27]

4.2 Additional Implications

The calibrated model has additional implications, other than the ones explicitly used in the algorithm above, that we can contrast with the data.

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 **3,** 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 correlation between the two variables, in spite of the fact that R&D data were not used in the calibration of the model. Whereas we model innovation in a simple, and rather restrictive way, our calibration indicates that the model captures well the linkage between trade, MP and innovation.[28]

[27] Benelux also has a low inward MP cost, which by Proposition 3 would imply a low r, but this effect is dominated by the effect of the low outward MP cost. To verify this claim, we conducted an experiment in which we started with all trade and MP costs at their average values (as in the solid curve in Figure 6) and then lowered the outward and inward MP costs of Benelux by 40% (i.e., $\gamma_{il} = 0.6\overline{\gamma}$ for $i = BEN$ and $l \neq BEN$ or $i \neq BEN$ and $l = BEN$). This resulted in an increase in Benelux's r from 10.6% to 40.8%.

[28] 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 percent. There are two explanations for this. First, the R&D data is 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 percent in the whole economy and 8.7 percent in manufacturing. Second, R&D in the data captures only a small part of what

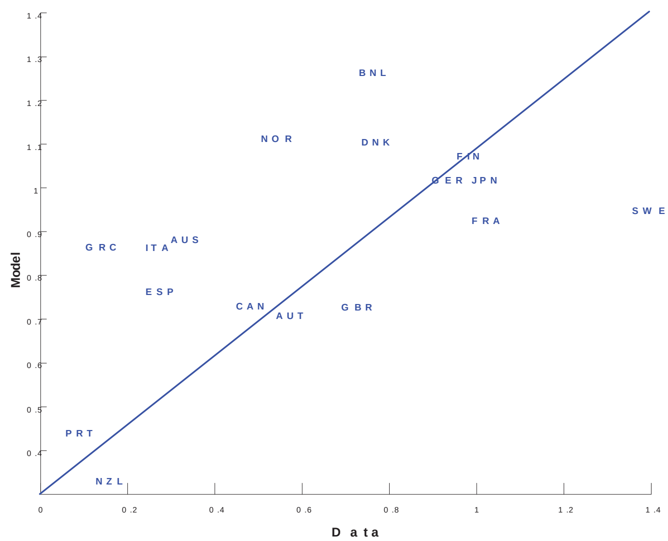

Figure 7: Employment Shares in R&D: Model and Data.
Note: Employment Shares in R&D are: in the model, the (equilibrium) variable r; in the data, employment in R&D, as a share of total employment, from the World Development Indicators, an average over the nineties.

Trade and MP costs. 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 estimated coefficients are increasing in distance and decreasing in sharing a border and language, and all the estimated coefficients are significant at the one-percent level. Figure 8 illustrates the strong positive relationship of calibrated trade and MP costs with geographical distance.

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}. 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. By construction, the model matches total MP flows, so $\sum_n X_{iln} = \sum_n \widehat{X}_{iln}$, where the X_{iln} refers to the data and $\widehat{X}_{iln}$ refers to the model. But we can compare the composition of MP flows across destinations between the model and the data. We think of this as an "out of sample" check on the model.[29]

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.

[29]As explained in Section 3.1, our data for trilateral flows comes from the BEA and is restricted to X_{iln} for $i = US$ and the top five markets for US sales: Canada, European Union, Great Britain, Japan, and the United States. It is important to emphasize here that these data X_{iln} were used in the calibration of the model only to

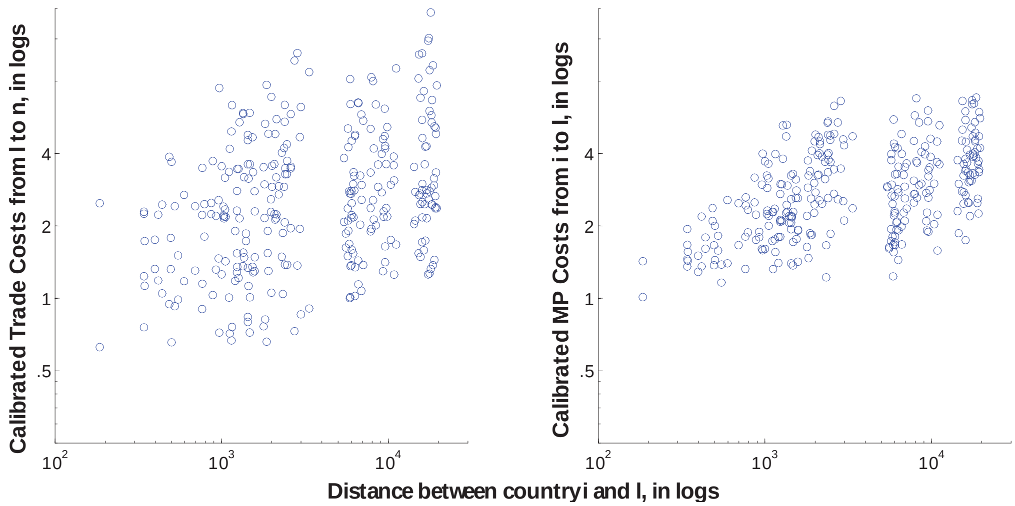

Figure 8: Calibrated Trade and MP Costs and Geographical Distance.
Note: Calibrated trade costs refer to τ_{ln}. Calibrated MP costs refer to $\widetilde{\gamma}_{il}$.

Following Ramondo and 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. We also refer to the ratio of total BMP flows over total MP flows as the BMP share, which we measure as $BMP_{il} \equiv \sum_{n \neq l} X_{iln} / \sum_n X_{iln}$. 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 0.08 while in the data this is 0.39. This behavior of the model is caused by the high value of the parameter ρ, which leads multinationals to serve foreign markets through exports or MP, but not BMP. Indeed, when the parameter ρ is lower, BMP shares go up. For $\rho = 0.40$—as calibrated using the Poisson PML coefficients in the unrestricted gravity regressions—, the average BMP share is 0.13 rather than 0.08.[30]

Apart from this failing, however, the model does a good job in matching the composition of BMP flows across destinations. The correlation between model and data BMP flows in logs (i.e., between log X_{iln} and log $\widehat{X}_{iln}$ for $n \neq l$) is 0.78 –in levels, the correlation is 0.95–. Of course, part of that is mere gravity: all flows in the model and the data are much higher when the destination market is large than when it is small. To proceed, we compare the trilateral flows implied by the model with a simple, yet intuitive, benchmark according to which the export behavior of foreign affiliates and domestic firms is the same,

pin one parameter, ρ, but they were not used in the calibration of trade and MP costs. This implies that although the model perfectly matches the bilateral trade and MP data, it does not do so for trilateral flows.

[30]In fact, the share of BMP does not change with the parameter θ, but only with ρ.

$$X_{iln} = \frac{\sum_n X_{iln}}{\sum_{i,n} X_{iln}} \cdot \sum_i X_{iln}.$$

Our calibrated mode implies systematic deviations from this benchmark because the affiliates of U.S. firms have different location opportunities for producing for a country n than other firms operating in country l.[31]

To test that the model has predictive power in excess of the simple benchmark, we estimate the following econometric model via OLS:

$$\frac{X_{iln} \cdot \sum_{i,n} X_{iln}}{\sum_i X_{iln} \cdot \sum_n X_{iln}} = b \frac{\widehat{X}_{iln} \cdot \sum_{i,n} \widehat{X}_{iln}}{\sum_i \widehat{X}_{iln} \cdot \sum_n \widehat{X}_{iln}} + \varepsilon_{iln}, \tag{25}$$

where $l \neq n \in \{\text{Canada, the E.U., Japan, the U.K., the U.S.}\}$ and ε_{iln} represents factors outside the model. Our choice of destination countries n is limited by the availability of data. The model has explanatory power in excess of the simple benchmark if the coefficient b on the model generated data is positive and statistically significant.

Estimating (25) on the 82 observations for which we have the necessary data yields a coefficient estimate of $\widehat{b}$ of 0.69 with a heteroscedasticity-corrected standard error of 0.18 and an R-squared value of 0.71. These results strongly suggest that the model does a good job in matching the pattern of trilateral flows observed in the data and far exceeds the predictive power of the simple benchmark model.[32]

5 Counterfactual Experiments

Armed with our calibrated model, we proceed to address the questions that motivate our study by performing a series of counterfactual experiments. 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 (under endogenous entry), real income, and real wages of workers in the innovation and production sectors. Finally, we analyze the effects on these same variables of the integration to the world economy of a low-productivity country, an experiment that is motivated by the emergence of China as a key location for manufacturing production over the last two decades.[33]

[31] The benchmark model in fact corresponds to the actual model analyzed in this paper when $\rho = 0$.

[32] When fixed effects by destination country (n) and production country (l) are included in the regression, the coefficient on predicted trilateral flows changes only moderately to 0.64 with a standard error of 0.21 and an R-squared of 0.74.

[33] The reader should keep in mind that our model was calibrated to manufacturing, hence the welfare implications correspond to real variables deflated by the price index in manufacturing only.

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 the Appendix for details). This version of our model allows us to address our 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 relative productivities in the two activities.[34] 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 and Waugh (2013), and Hsieh, Hurst, Jones and 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. In all the counterfactual exercises below, we first remove the current-account imbalances by computing the equilibrium with $\Delta_i = 0$ for all i. The resulting equilibrium is referred to as the "baseline equilibrium" below.

5.1 The Gains from Openness

Table 4 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 2.5.[35]. 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 a country with high net inward MP flows like New Zealand, for which the gains from openness are also large, but with 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 Great Britain are examples of countries that, like New Zealand, are specialized in production ($r < \eta$) and

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

[35]The results in Table 4 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		
Australia	0.167	1.156	1.190	0.972	1.004	1.028
Austria	0.149	1.443	1.528	0.945	1.068	1.029
Benelux	0.240	1.819	1.709	1.064	1.126	1.128
Canada	0.142	1.566	1.679	0.933	1.056	1.056
Denmark	0.215	1.394	1.346	1.035	1.118	1.038
Spain	0.148	1.128	1.197	0.943	0.993	0.9998
Finland	0.208	1.287	1.254	1.026	1.061	1.058
France	0.179	1.182	1.195	0.989	1.035	1.018
Great Britain	0.147	1.321	1.404	0.941	1.024	1.049
Germany	0.195	1.225	1.213	1.010	1.048	1.048
Greece	0.164	1.075	1.112	0.967	0.998	0.995
Italy	0.168	1.094	1.125	0.972	1.006	1.001
Japan	0.200	1.052	1.035	1.016	1.025	1.019
Norway	0.216	1.230	1.187	1.036	1.091	1.033
New Zealand	0.082	1.432	1.761	0.813	0.993	1.038
Portugal	0.109	1.489	1.706	0.873	1.018	1.036
Sweden	0.183	1.431	1.440	0.993	1.062	1.068
United States	0.192	1.083	1.077	1.006	1.016	1.023
Average	0.171	1.300	1.342	0.974	1.041	1.037

Table 4: 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).

have a negative indirect effect.

In Table 4 we also present the gains from trade and the gains from MP. These welfare changes tend to be low, 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 and Rodríguez-Clare (2013)). Can trade or MP lead to welfare losses? Our model yields an affirmative answer, although the losses are always small. Spain, Greece and New Zealand lose from trade, while Spain and Greece lose from MP. The explanation for these effects is contained in Section 2.5. If calibrated inward MP costs are low relative to calibrated outward MP costs, then lowering MP costs from infinity to their calibrated values (while leaving trade costs at their calibrated levels) leads to a reallocation of resources from innovation to production, triggering a deterioration in the country's terms of trade. The same thing happens when lowering trade costs from infinity to their calibrated values (while leaving MP costs at their calibrated levels). As shown in the first column of Table 4, 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 for the three versions of the model – results are presented in Table 5. Columns 1-2, 3-4 and 7-8 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 77% in the linear model, 15.5% in the Roy model and 5.8% 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 γ_{il} is low has a stronger effect than the same percentage decline in γ_{il} when γ_{il} is high. The effect is much smaller according to 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 Greece experiences overall loses, and only according to the linear model. As explained in Section 2.5, aggregate loses arise because of a reallocation of resources from innovation to production and the associated terms of trade deterioration. Since this reallocation is smaller, or non-existent, in the Roy and exogenous-entry models, loses are less likely in these models.

What happens to the real wages of workers that remain in the production sector? Proposition 4 shows that, under exogenous entry, real production wages could fall with MP liberalization in a country with a mix of high innovation productivity and a large endowment of innovation workers. The results in Table 5 show that this does not happen in our calibrated economy: real production wages increase with MP liberalization in all the countries in our sample.

Something very different happens with real innovation wages, which tend to fall with MP liberalization in the countries that are net recipients of MP – for example, real innovation wages decrease by 2% to 4% in Spain, New Zealand, Portugal and Greece 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 (e.g., Benelux, Denmark and Norway), whereas the opposite happens in countries exhibiting net inflows of MP. Surprisingly, real production

	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
Australia	-1.6	1.5	-2.5	1.5	1.7	0.6	-1.9	1.6	2.0	-0.4
Austria	-10.1	1.9	-7.7	1.9	2.4	-0.7	-3.8	1.9	2.6	-1.9
Benelux	76.9	6.9	15.5	5.2	3.4	10.3	5.8	4.6	2.7	10.6
Canada	-3.1	2.9	-2.5	3.0	3.1	2.1	-0.8	3.1	3.2	2.3
Denmark	5.3	1.8	3.9	1.8	1.5	3.2	2.5	1.7	1.0	4.2
Spain	-10.9	0.4	-8.0	0.6	1.1	-2.2	-5.7	0.9	1.9	-4.9
Finland	2.3	2.4	2.9	2.4	2.2	3.4	1.9	2.3	1.8	4.3
France	-5.7	0.9	-2.3	1.1	1.2	0.3	-1.3	1.1	1.4	-0.2
Great Britain	-11.5	2.3	-3.7	2.5	2.7	1.2	-1.0	2.6	2.8	1.6
Germany	-8.8	2.4	1.2	2.1	2.0	2.5	0.9	1.9	1.7	2.9
Greece	-6.6	-0.1	-5.4	0.0	0.4	-1.8	-4.3	0.2	1.1	-4.1
Italy	-6.0	0.2	-4.5	0.3	0.6	-1.2	-3.2	0.5	1.2	-2.8
Japan	0.9	0.6	1.0	0.6	0.5	0.9	0.8	0.6	0.4	1.4
Norway	3.5	1.4	2.9	1.4	1.1	2.4	2.0	1.3	0.7	3.4
New Zealand	-40.4	2.2	-19.0	3.0	3.5	-3.9	-7.7	3.3	4.0	-4.7
Portugal	-22.4	2.5	-15.3	2.5	3.1	-3.0	-8.3	2.4	3.5	-6.1
Sweden	-2.5	3.3	1.8	3.1	2.9	3.7	1.4	2.9	2.6	4.3
United States	-2.9	0.8	-0.7	0.8	0.9	0.6	-0.3	0.9	0.9	0.5
average	-2.4	1.9	-2.4	1.9	1.9	1.0	-1.3	1.9	2.0	0.6

Table 5: 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.

wages increase by more than real innovation wages in the United States with MP liberalization, implying a decrease in the share of innovation wages in total income.

We explored how the results presented in Tables 4 and 5 change under our alternative calibration with a higher θ. It is not surprising that the overall gains from openness are lower with this parameter given the welfare expression (21): The average country gains 20% versus 30% in our baseline calibration and as expected the difference is almost entirely due to the direct effect. Moreover, the gains from MP for all countries are larger than one; that is, no country loses from opening up to MP, even though a country like Greece has virtually zero gains. Regarding the MP liberalization exercise, the alternative calibration delivers very similar results to those in Table 5.

5.2 The Rise of the East

Perhaps the single most important 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, but no negative effect on real wages. We now use the calibrated model to explore these questions at a quantitative level.

To conduct this exercise, we think of our model calibrated with data for the late 1990s as corresponding to a world in which China had not yet integrated with the rest of the world. In the counterfactual, China is integrated with the countries in our sample with trade and MP costs given by their (weighted) average in our calibrated model.[36] Moreover, as we did for all other countries, we set the size of China to equal its endowment of equipped labor in manufacturing. The only parameters left to calibrate are T^e_{CH} and T^p_{CH}. We set these parameters so that the resulting equilibrium wage for China relative to the U.S. is equal to that observed in the year 2010, i.e. $w_{CH}/w_{US} = 0.16$, and the equilibrium innovation share is equal to that of Portugal, which has the lowest innovation share in our sample, which implies a target of $r_{CH} = 0.11$.[37]

The results for the three versions of the model are presented in Table 6. Moving from the baseline equilibrium to this counterfactual leads to an increase in r for all the countries in our sample, with the notable exception of Benelux in the linear model. The increase in r is particularly high in New Zealand and Portugal – this is because these countries were specialized in production according to the late 1990s data used for the calibration, and production moves East with the integration of China into the world economy. This is reflected in the dramatic decline in r in China as it becomes a key production site for firms from all over the world.[38]

Remarkably, however, real wages for workers in production do not fall in any of the countries in our sample – see Table 6. In the linear model the negative effect on workers is muted by the reallocation of labor from production to innovation. This was the main lesson from Proposition 5 in Section 2.5. Something similar happens in the Roy model, although worker heterogeneity implies that reallocation will not be as effective in helping production

[36] This implies that $\gamma_{i,CH} = 2.9$, $\gamma_{CH,l} = 2.3$, $\tau_{l,CH} = 4.1$ and $\tau_{CH,n} = 2.8$, for all i, l, n.

[37] The relative wage w_{CH}/w_{US} is obtained as (Y_{CH}/L_{CH}) / (Y_{US}/L_{US}), where L_i is equipped labor computed as in Klenow and Rodriguez-Clare (2005) and Y_i is GDP in US dollars from the WDI.

[38] This implication of the model may seem inconsistent with the observed expansion of R&D taking place in China in recent years. But one must keep in mind that our model is not designed to explain R&D variation across countries or across time – we are only focusing on how trade and MP affect innovation while leaving everything else out of the model. We could easily add a tax that increases the cost of innovation or decreases its net returns as a way to generate such R&D variation, but this is outside the scope of this paper.

	Linear Model		Roy Model				Exogenous Entry			
% change in:	r	X/P	r	X/P	w/P	w^e/P	r	X/P	w/P	w^e/P
Australia	24.2	5.5	21.5	5.2	3.7	12.3	17.8	4.6	0.7	23.3
Austria	18.2	4.8	17.1	4.3	3.3	10.0	15.4	3.7	0.8	19.7
Benelux	-21.8	2.0	4.3	3.0	2.6	4.5	8.6	2.6	0.7	11.4
Canada	19.4	3.7	15.7	3.4	2.5	8.5	12.9	2.8	0.4	16.0
Denmark	19.1	5.0	16.6	5.1	3.5	10.7	13.6	4.8	0.8	19.0
Spain	12.7	4.3	12.7	4.0	3.3	8.3	12.2	3.3	1.1	15.9
Finland	14.0	12.2	12.3	12.0	10.7	16.4	10.8	11.0	7.9	23.1
France	12.8	3.6	12.1	3.5	2.6	7.5	11.2	3.0	0.5	14.5
Great Britain	24.5	4.2	18.3	3.7	2.7	9.7	14.5	3.2	0.4	18.2
Germany	23.2	4.2	16.7	4.2	2.8	9.7	13.2	3.9	0.4	17.6
Greece	24.8	13.2	22.7	12.6	10.9	20.5	19.7	11.4	7.0	33.4
Italy	7.8	5.7	8.2	5.5	4.9	8.3	8.4	4.8	3.1	13.7
Japan	8.0	2.6	8.2	2.7	2.0	5.4	7.9	2.5	0.5	10.7
Norway	8.9	21.9	9.1	21.5	20.5	25.1	9.0	20.1	17.2	31.0
New Zealand	88.9	8.0	46.6	5.5	4.2	19.8	30.9	5.2	0.8	37.7
Portugal	59.3	6.7	38.5	5.3	3.7	17.4	29.6	5.4	0.6	36.6
Sweden	6.7	5.7	8.8	5.7	5.0	8.7	9.6	5.0	2.8	15.1
United States	11.6	3.0	11.2	3.1	2.1	6.8	10.4	2.8	0.2	13.4
China	-46.6	4.6	-44.9	5.8	9.4	-13.3	-46.5	10.4	22.2	-40.9

Table 6: The Rise of China: Linear, Roy, and Exogenous-Entry Model.
Note: w^e is the wage in the innovation sector. For the linear model, changes in X/P are equivalent to changes in w/P.

workers share in the overall gains from specialization in innovation. Indeed, real wages for production workers increase much less than for innovation workers. Not surprisingly, this is even more pronounced when there is no mobility of labor between innovation and production, as in the exogenous-entry model. In that case we find that the share of innovation wages in total income experiences strong increases, ranging from approximately 10% in the United States to 30% in New Zealand and Portugal. Remarkably, this does not mean that production workers are negatively affected by the emergence of China. Returning again to the case of the United States, our finding is that real production wages are basically not affected while real innovation wages increase by 13%. The opposite occurs in China, where real production wages increase by 22% while real innovation wages fall by 41%, resulting in a 40% decline in the share of innovation wages in total income.

Does this mean that there is no basis for the popular fear that the move of manufacturing to low wage countries can hurt production workers? To explore this further, we considered a scenario that comes closer to that considered in Proposition 4. We start with MP costs at their calibrated values and frictionless trade, and then consider a decline in MP costs from the United States to China all the way down to zero costs, i.e., $\gamma_{US,CH} = 1$. This is close to the conditions in Proposition 4, except that we have many countries and only consider frictionless MP from the United States to China, while setting θ, σ and ρ at their calibrated values. In this case production workers in the United States still come out better off, although real production wages increase by only 5.5% whereas real innovation wages increase by 44% and overall real income increases by 15%. The opposite pattern holds in China, where production workers gain 86%, innovation workers lose 90%, and overall welfare increases by 74%.

6 Conclusion

The process of globalization features increasing international specialization in innovation or production. To assess the welfare implications of this process, we develop a quantitative, multi-country 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 tradeoff: trade costs may induce firms to open many plants to be near local customers but this is at the expense of producing in the country with the lowest production cost. 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."

Our quantification of this framework reveals that asymmetric bilateral trade and MP costs play a critical role in determining the structure of global innovation and production. We also demonstrate that falling MP costs generate efficiency gains but can make some countries worse off, particularly when innovation is induced to leave a country, thereby exposing it to a deterioration in its terms of trade. Within countries, workers with specific types of skills can lose even as national welfare rises. Finally, we use our model to study the impact of the integration of China into the global economy. We find that despite having a large impact on the structure of global specialization, with production largely migrating to China and innovation migrating to the developed world, workers in the developed world largely gain.

References

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

Antras, P. (2003), 'Firms, contract, and trade structure', *The Quarterly Journal of Economics* **118**(4), 1375–1418.

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.

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. (2010), 'Estimates of the trade and welfare effects of NAFTA'. Manuscript, University of Chicago and Yale University.

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.

Dekle, R., Eaton, J. & Kortum, S. (2008), 'Global rebalancing with gravity: Measuring the burden of adjustment', *IMF Staff Papers* **55**(3), 511–540.

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. (2012), 'Railroads of the raj: Estimating the economic impact of transportation infrastructure'. Manuscript, MIT.

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.

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.

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. (2012), 'The margins of multinational production and the role of intra-firm trade', *forthcoming, Journal of Political Economy* **7145**.

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', *forthcoming American Economic Review* .

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', *Available at SSRN 1993332* .

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

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. (2009), 'The elasticity of trade: Estimates and evidence', *Manuscript, University of California, Davis and New York University* .

Tintelnot, F. (2012), Global production with export platforms, Technical report.

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

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 ($l \neq U.S., n$).[39] 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 and 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 and Waugh (2009). 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

[39] 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.

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 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.[40] 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 $l \neq n$.

Data used for the calibration procedure. All the data used for the calibration are averages over the period 1996-2001. Bilateral MP data is from UNCTAD, Investment and Enterprise Program, FDI Statistics, FDI Country Profiles, published and unpublished data.[41] The data are on revenues of affiliates from country i in country l, and cover all non-financial majority-owned foreign affiliates.[42] A foreign affiliate is defined as a firm who has more than ten-percent of its shares owned by a foreigner, while a majority-owned foreign affiliate is a firm with more than fifty-percent foreign ownership. The later type of foreign affiliates are the largest part of the total number of foreign affiliates in a host economy.

The UNCTAD measure of MP includes both local sales in l and exports to any other country, including the home country i. Moreover, revenues of affiliates from country i in country n can be captured by the receiving country n (inward), and by the source country i (outward). Thus, potentially, each observation has two data sources. We consider first outward magnitudes as reported by the source country, and complete with inward magnitudes as reported by the receiving country. In this way, we minimize the underestimation arising from some receiving countries reporting only local sales of foreign affiliates rather than total revenues. Additionally, we minimize the problem of receiving countries that do not compile their statistics by reporting the country of the "Ultimate Beneficiary Owner" (i.e. the country of the ultimate investor—by opposition to the country of the immediate investor).

Out of 342 possible country pairs, data are available for 219 country pairs. We impute missing values by running OLS on

$$\log \frac{Y_{in}}{w_n L_n} = \beta_d \log d_{in} + \beta_c b_{in} + \beta_l l_{in} + S_i + D_n + e_{in},$$

where Y_{in} is revenues of affiliates from i in n, $w_n L_n$ is GDP in country l, d_{in} is geographical distance between i and n, b_{in} (l_{in}) is a dummy equal to one if i and n share a border (language), and zero otherwise, and S_i and D_n are two sets of country fixed effects, for source and destination country, respectively. The variable GDP is in current dollars, from the World

[40] 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.

[41] Unpublished data are available upon request at fdistat@unctad.org.

[42] The exception is Portugal that reports affiliates in all sectors including the financial sector.

Development Indicators.

The bilateral trade data and absorption data are from the OECD STAN manufacturing database. Absorption X_n is calculated as gross value of production minus total exports plus imports from OECD (18) countries.

B Theory Appendix

B.1 Properties of the Multivariate Pareto

(i) We show that with $\rho \to 1$ the elements of $\boldsymbol{z}$ 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 $G_i(z_1, ..., z_N) = 1 - x \left(\sum_{l=1}^N \left(T_{il} z_l^{-\theta}/x\right)^{\frac{1}{1-\rho}}\right)^{1-\rho}$. 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 $\widetilde{\sigma}^{1-\sigma}X_nP_n^{\sigma-1}c^{1-\sigma}$, equation (5) implies that total sales from n to l by firms from i, X_{iln}, are

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

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

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

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

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

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

$$
\begin{aligned}
L_1^e &= \eta\lambda_1^E(L_1 + L_2/\omega), && \text{(B.11)}\\
L_2^e/\omega &= \eta\lambda_2^E(L_1 + L_2/\omega). && \text{(B.12)}
\end{aligned}
$$

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

$$
\begin{aligned}
w_i L_i^e &= w_i L_i\left(1 - \frac{\theta-\sigma+1}{\sigma\theta}\right) - \frac{1}{\tilde{\sigma}}\lambda_i^T\sum_k w_k L_k \implies\\
r_i &= \eta + 1 - 1/\sigma - \frac{1}{\tilde{\tilde{\sigma}}}\frac{\lambda_i^T\sum_k w_k L_k}{w_i L_{i.}}
\end{aligned}
$$

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

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

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

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

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 $l_1 \equiv L_1 / \left(L_1 + L_2 \right)$ we get the above expression to be

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

and similarly for the second country. Noting that $r_i = 1 - L_i^p / L_i$, we have$_1$

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

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*

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

and real expenditure is given by

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

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 $(B.5)$, $\zeta \equiv \left(\frac{\tilde{\sigma}^{1-\sigma}\theta}{\theta-\sigma+1}\right)^{1/\theta}\left(\frac{\sigma}{\tilde{\tilde{\sigma}}^{1-\sigma}}\right)^{\frac{\sigma-1-\theta}{\theta(\sigma-1)}}$, 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

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

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*

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

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

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

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

or equivalently, by the sign of $v\eta - \frac{\theta}{1+\theta}$. The condition $v\eta > \frac{\theta}{1+\theta}$ 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),

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

This expression is similar to what we had above for real wages, only that instead of v we now have $v+\theta$. Thus, the condition for real income to increase is that $(v+\theta)\,\eta > \frac{\theta}{1+\theta}$. 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,

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

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

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

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. We infer some of the aggregate equilibrium variables in the model using data on bilateral trade and MP shares, absorption and size of country i. We focus here on the model with endogenous entry and homogenous labor (i.e., the linear model), and later discuss how the calibration of the other two models can be trivially derived from this calibration.

1. Output Y_l is calculated using data on trade shares and absorption,

$$Y_l = \sum_n \lambda_{ln}^T X_n.$$

2. National income is given by

$$w_i L_i = \frac{\theta - \sigma + 1}{\sigma \theta} \left(X_l - Y_l \right) + \left(1 - \eta \right) Y_l + \eta \sum_l \lambda_{il}^M Y_l.$$

This equation is obtained by combining (13) and (15) with (11) summed over all l's. Given $w_i L_i$ and our measure of L_i as equipped labor we can then recover the implicit wage, w_i.

3. In our model, current accounts are balanced, which implies $X_i = w_i L_i$. But this relationship is not satisfied in the data. We follow Dekle, Eaton and Kortum (2008) and simply assume that there is an exogenous current account deficit, Δ_i, which allows expenditure, X_i, to be different than total income, $w_i L_i$,
$$\Delta_i = X_i - w_i L_i.$$

4. We infer the share of labor in innovation using (16).

5. The $N \times 1$ vector of entrants $\boldsymbol{M}$ is given by
$$M_i = r_i L_i / f_i^e = r_i L_i,$$
where we have assumed that $f_i^e = 1$.

Step 2. In this step we use the data and the aggregate variables recovered in Step 1 together with the model restrictions to recover trade and MP costs. Define $\widetilde{\gamma}_{il} \equiv T_{il} \gamma_{il}^{-\theta}$. We impose $\widetilde{\gamma}_{ii} = L_i^{1-\rho}$ and $\tau_{ll} = 1$.[43] We also impose $\tau_{ll} = 1$. Given a value for θ and ρ, the aggregate variables computed in Step 1 and an initial guess for the matrices of $\widetilde{\gamma}_{il}$'s and τ_{ln}'s, we use expressions (8), (10), and (11) to compute compute λ_{ln}^T and λ_{il}^M in the model. We iterate on the matrices of $\widetilde{\gamma}_{il}$'s and τ_{ln}'s to precisely match the observed trade and MP shares.[44] For future reference, note that $\widetilde{\gamma}_{il}$ and τ_{ln} together with wages and $\{\theta, \rho\}$ allow us to recover the variables ψ_{iln}, λ_{in}^E, and X_{iln} implied by the model.

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 as calibrated above. A similar argument applies to the Roy model (see Appendix B.11 and the online appendix).

[43] The normalization $\widetilde{\gamma}_{ii} = L_i^{1-\rho}$ can be seen as the result of assuming $T_i^e = 1$, $T_l^p = L_l^{1-\rho}$ and $\gamma_{ii} = 1$.

[44] We do not impose any bound on τ's. Notice also that, for this step, we do not really need θ; we just can redefine $\widetilde{\tau}_{ln} \equiv \tau_{ln}^{-\theta}$.

Calibration and the normalization

We now explain why the normalization that we impose in Step 2 (i.e., $\widetilde{\gamma}_{ii} = L_i^{1-\rho}$) does not affect the results of the counterfactual exercises that we conduct with the calibrated model. In particular, we consider a percentage change of MP and trade costs and show that the resulting change in the equilibrium variables does not depend on the normalization. We focus on the linear endogenous entry model; extending the result to the exogenous entry model and the Roy model is straightforward.

The first step is to show that the calibrated values of θ and ρ do not depend on the normalization. Since the trade elasticity estimated from the restricted gravity regression pins down their ratio, we only need to show that the trade elasticity estimated from the unrestricted gravity regression (for the model) in step 3 is not affected by the normalization. The normalization in step 2 does affect the estimated trade costs, which are used in step 3. But such effects are absorbed by the exporter and importer fixed effects in the regression of step 3, leaving the estimated coefficient on the bilateral trade cost unaffected.

The second step is to show that even if the calibrated values of $\widetilde{\gamma}$ and $\boldsymbol{\tau}$ do depend on the normalization, the proportional changes in the variable of interest given proportional changes in $\widetilde{\gamma}$ and $\boldsymbol{\tau}$ do not depend on the normalization. We use the hat symbol ($\hat{}$) to denote the ratio of the value of a variable in the new equilibrium to its value in the initial equilibrium, $\hat{x} \equiv x'/x$. We consider proportional changes in γ or $\boldsymbol{T}$ (which imply changes in $\widetilde{\gamma}$), and $\boldsymbol{\tau}$. Denote changes in all these parameters as the vector $\hat{\boldsymbol{\Theta}} = \{\hat{\boldsymbol{\gamma}}, \hat{\boldsymbol{T}}, \hat{\boldsymbol{\tau}}\}$. Step 1 of the calibration provides a set of variables $\boldsymbol{\lambda}^T$, $\boldsymbol{\lambda}^M$, $\boldsymbol{X}$, $\boldsymbol{Y}$, and $\boldsymbol{w}$ which do not depend on the normalization, which is only used in step 2. These variables serve as calibration targets in step 2. We now show that changes in wages, $\hat{\boldsymbol{w}}$, total expenditure, $\hat{\boldsymbol{X}}$, and entry, $\hat{\boldsymbol{M}}$, are determined from a system of equations that depends only on $\hat{\boldsymbol{\Theta}}$ and the variables that serve as targets in our calibration algorithm, but not directly on $\widetilde{\gamma}$ and $\boldsymbol{\tau}$. To do so, note that we can use the definition of Ψ_{in} and equations (11) and (10) to write $\hat{\boldsymbol{\Psi}} = h_1\left(\hat{\boldsymbol{\Theta}}, \hat{\boldsymbol{w}}, \hat{\boldsymbol{M}}\right)$, $\hat{\boldsymbol{\lambda}}^M = h_2\left(\hat{\boldsymbol{\Theta}}, \hat{\boldsymbol{w}}, \hat{\boldsymbol{M}}\right)$, $\hat{\boldsymbol{\lambda}}^T = h_3\left(\hat{\boldsymbol{\Theta}}, \hat{\boldsymbol{w}}, \hat{\boldsymbol{M}}\right)$, where the functions h_1, h_2 and h_3 do not depend on the $\boldsymbol{T}$ and γ. Using these definitions and proceeding as in Dekle et al. (2008) we can reformulate our equilibrium system in changes,

$$\frac{\theta - \sigma + 1}{\sigma\theta}\hat{X}_i X_i + \frac{1}{\widetilde{\sigma}}\sum_n \hat{\lambda}_{in}^T \lambda_{in}^T \hat{X}_n X_n = \hat{w}_i w_i \left(L_i - \hat{M}_i M_i f^e\right),$$
$$\eta \sum_n \hat{\lambda}_{il}^M \lambda_{il}^M Y_l = \hat{M}_i M_i \hat{w}_i w_i f^e,$$
$$\hat{X}_i X_i = \hat{w}_i w_i L_i + \Delta_i.$$

The important point here is that this system allows us to solve for $\left(\hat{\boldsymbol{w}}, \hat{\boldsymbol{M}}, \hat{\boldsymbol{X}}\right)$ as a function of target variables $\boldsymbol{\lambda}^T$, $\boldsymbol{\lambda}^M$, $\boldsymbol{X}$, $\boldsymbol{Y}$, and $\boldsymbol{w}$. Having solved this system computing real wage changes is straightforward. We use expression (B.5) and the definition of Ψ_{ii} to write changes in the price index as

$$\widehat{P}_i = \left(\widehat{w}_i / \widehat{X}_i\right)^{\frac{1+\theta-\sigma}{\theta(\sigma-1)}} \left(\widehat{M}_i \widehat{\Psi}_{ii} / \widehat{\lambda}^E_{ii}\right)^{-1/\theta}.$$

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^e_i, 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 and Waugh (2013) and Hsieh, Hurst, Jones and 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 $v^e w^e_i \geq v^p w_i$. 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

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

and

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

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 $L_i - M_i f^e = L^p_i$, and using $(B.28)$, (14) (15), and by the labor market clearing in innovation, $M_i f^e = L^e_i$, with L^e_i 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.