Innovation and Production in the Global Economy

(PRELIMINARY AND INCOMPLETE)*

Costas Arkolakis — Yale and NBER

Natalia Ramondo — Arizona State

Andrés Rodríguez-Clare — UC Berkeley, Penn State, and NBER

Stephen Yeaple — Penn State and NBER

December 12, 2011

Abstract

We develop a monopolistic competition model of trade and multinational production (MP). Firms receive an idiosyncratic vector of productivities for different locations from a multivariate distribution. They also face distance related trade and MP costs. Thus, individual firms face a proximity—versus—comparative advantage trade-off to serve individual locations from close-by or high productivity locations. The model gives simple structural expressions for bilateral trade and MP. We use these expressions to calibrate the model across a set of OECD countries. We quantify the implications of openness to trade and MP on the allocation of employment between production and innovation, as well as the implications for wages, profits and overall welfare.

*We thank Pol Antras, Arnaud Costinot, Jonathan Eaton, Gene Grossman, Samuel Kortum, and Felix Titelnot, as well as seminar participants at Arizona State University, Boston University, the Cowles Foundation, MIT, Penn State, Princeton, Stanford, University of Oregon, and Yale for insightful comments. We also thank Treb Allen for excellent research assistance and the Human Capital Foundation (http://www.hcfoundation.ru) for support. All remaining errors are our own. The statistical analysis of firm-level data on U.S. multinational corporations reported in this study was conducted at the International Investment Division, U.S. Bureau of Economic Analysis, under arrangements that maintained legal confidentiality requirements. Views expressed are those of the authors and do not necessarily reflect those of the Bureau of Economic Analysis. This paper previously circulated under the title: "Proximity versus Comparative Advantage: A Quantitative Theory of Trade and Multinational Production".

1 Introduction

A fundamental feature of "globalization" is the increasing geographic separation between innovation and production. With the rapid growth of multinational production (MP), it is increasingly likely that knowledge developed by a firm in one country will be exploited in production facilities scattered throughout the globe.[1] Together with the growth of MP has come a public uneasiness as to its impact on welfare. Multinational firms are frequently accused of "shipping jobs overseas," thereby benefitting their host countries at the expense of workers in the home country. However, this discussion typically abstracts from the fact that foreign affiliates generate profits abroad and induce greater innovation in the source country.[2]

Understanding the interaction of these different forces and how they together affect welfare requires careful general equilibrium analysis. This paper develops a quantitative multi-country general equilibrium model in which the location of innovation and production is endogenous and geographically separable. We build on the established theory of international trade by relaxing the assumption that production occurs in the same country where the firms and ideas are created.

Formally, we model innovation as the creation of heterogenous firms that sell differentiated goods in monopolistically competitive markets.[3] We depart from Melitz (2003) 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 "proximity-comparative advantage trade-off." On the one hand, firms want to be close to their customers to avoid trade costs; on the other hand, they want to produce in the country where they would achieve the minimum unit cost, i.e., the country that has a comparative advantage in production for this particular firm.

By allowing firms to produce outside of their home country, multinational production leads some countries to specialize in innovation and others to specialize in production.[4] Countries that

[1] As of 2004, over 65 percent of U.S.-based manufacturing firms sales abroad are due not to exports from the United States but to the foreign affiliates of these firms. Further, almost one-fifth of U.S. exports can be attributed to the affiliates of foreign firms operating in the United States.

[2] Indeed, in the case of Ireland, a popular host country for MP, the profits earned by the affiliates of foreign multinationals account for as much as 20 percent of Irish GDP.

[3] This is what is commonly referred to as "entry" in the context of the Melitz (2003) model.

[4] In the absence of multinational production, the share of labor devoted to innovation would be the same in all countries. This is consistent with the version of the Melitz/Chaney model presented in Arkolakis, Demidova,

specialize in innovation have a net inflow of profits that compensates for the cost of innovation and allows them to run a trade deficit. Loosely speaking, these countries export ideas and import goods.

There are two forces that determine the geographic allocation of innovation: first, countries that have a high productivity in innovation relative to production will tend to specialize in innovation, and second, home market effects (HME) imply that country size and location matter for the allocation of production and innovation – in particular, home market effects 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 "comparative advantage in innovation" while the second force is related to proximity to consumers (for production) and workers (for innovation). This is another sense in which our model exhibits a proximity-comparative advantage trade-off – in this case, however, the trade-off takes place at the aggregate rather than the firm level, and operates through general equilibrium forces.

One of the issues that we explore with our model regards the effect of MP on real wages. We first consider a version of our model with exogenous innovation, i.e., the measure of firms in each country is exogenous and the free-entry condition is ignored. In this case, we show that under certain parameter values, MP may actually hurt workers in countries endowed with a high ratio of firms to workers. The reason why MP may hurt workers is intuitive and resonates with the popular discussion about this issue: MP makes it feasible for firms to produce outside of their home country, and this effectively generates competition for home-country workers.

This negative effect of MP on real wages critically depends on the assumption that innovation is exogenous. In the full model with endogenous innovation we find that this negative effect of MP on workers is no longer present. The reason is simple: with endogenous innovation, a decline in MP frictions leads countries with a comparative advantage in innovation to reallocate workers from production to innovation. This reallocation benefits workers by increasing the measure of national firms and by reducing the size of the production sector, thereby improving the country's terms of trade.[5]

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

[5] The result that under exogenous innovation MP can hurt workers in advanced countries while this is no

We calibrate the model to match trade and MP flows and then use the calibrated model to address a series of questions. How does a general decline in trade and MP costs affect the location of innovation, and what are the implications for relative and real wages in different countries? In particular, does MP hurt workers in countries that have a comparative advantage in innovation? How does a unilateral change in trade or MP costs for one country affect not only that country but also its neighbors? How much do different countries gain from trade and MP? At this point, our quantitative results are very preliminary, so we will not describe them here at length – suffice it to say for now that our results suggest that even with exogenous innovation, a decline in MP frictions does not hurt workers in rich countries.

The mechanisms at work in our model have antecedents in the classic work on trade and MP (see Markusen (2002)). This literature highlights at least four key ideas: (1) MP allows innovation (entry) to be geographically separated from production, (2) countries may differ in their relative costs in innovation and production, which leads to a tendency toward some specialization in one of these two activities (3) the non-rivalry of technology within the firm allows multi-plant production, and (4) trade costs encourage while MP costs discourage multi-plant production. The incorporation of these features into general equilibrium trade models dates back to Helpman (1984) and Markusen (1984). Helpman (1984) focuses on motivations for the geographic separation of innovation and production, while Markusen (1984) focuses on the motivation for and welfare implications of multi-plant production.[6] By simplifying comparative advantage to a probabilistic setting and by replacing plant-level fixed costs with marketing fixed costs, we gain the ability to construct a tractable and quantifiable, multi-country model that incorporates the most important mechanisms found in this earlier work.

Our model provides a strict generalization of the Melitz (2003)-Chaney (2008) model of trade. 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, Demidova, Klenow, and

longer possible under endogenous innovation is similar to the results in Rodriguez-Clare (2010) in the context of a model of trade and offshoring.

[6] 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.

Rodríguez-Clare (2008)). Another strict generalization of the Melitz-Chaney model to allow for MP is Helpman, Melitz, and Yeaple (2004) (henceforth HMY). Our approach has significant differences with HMY. First, we allow for comparative advantage in innovation so that some countries specialize in innovation and exhibit net outward MP while others specialize in production and exhibit net inward MP. Second, by considering a general equilibrium model we can study the role of home market effects on the geography of innovation and production. Third, our model easily accommodates the possibility that multinational affiliates may use some production locations as export platforms to other countries, while this possibility leads to severe computational problems in HMY.[7]

One potential drawback of our approach relative to HMY (and the quantitative application of their framework by Irarrazabal, Moxnes, and Opromolla (2009)) 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 buys us the tractability to handle export platforms. It is important to note that some of the key implications of the proximity-concentration trade-off appear in our model through alternative mechanisms. For instance, our model is consistent with large firms having more affiliates – larger firms serve more markets, and this leads them to open more affiliates to avoid trade costs.

A close relative to our model is Ramondo and Rodríguez-Clare (2009), which extends the perfect competition Ricardian framework of Eaton and Kortum (2002) to allow for MP. Whereas both models have similar predictions regarding aggregate trade and MP flows, the counterfactuals are different because our model takes into account the effect of trade and MP costs on profits (under exogenous innovation) or the location of innovation (under endogenous innovation). This leads to important differences in the welfare implications of trade or MP liberalization.

Our model is also related to Prescott and McGrattan (2010) and McGrattan (2011). These papers extend the neoclassical growth model by introducing a non-rival "knowledge capital" that can be used in any location. The use of knowledge capital accumulated in one country to produce in another country is interpreted as MP while trade takes place only as a way to

[7]For example Irarrazabal, Moxnes, and Opromolla (2009) develop the first multi-country calibration of the HMY model and they are forced to ignore export platform MP. They are also forced to introduce bilateral constraints in terms of the wages and trade costs to be able to solve the model. Our approach is more amenable to calibration and this allows us to conduct a series of general equilibrium counterfactual exercises.

transfer the returns to capital. We think of our approaches as complementary: while our model can more easily connect to the trade and MP data, the McGrattan and Prescott approach allows for an analysis of the transition path as countries open up to MP.

Finally, Eaton and Kortum (2007) explore similar issues as those in this paper in the context of a two country model with endogenous innovation and Bertrand competition. In particular, they allow for diffusion of ideas so that innovators in one country can use their ideas for production in the other country. Diffusion in their model affects specialization in innovation in a way analogous to what MP affects specialization in our setup. By allowing for heterogenous productivities for each firm in different locations, we avoid the Ricardian-type discontinuities that lead to a multiplicity of cases in Eaton and Kortum (2007), and in that way extend the model to multiple countries and connect it to trade and MP data.

2 The Model

We now describe the details of our proximity–versus–comparative advantage model. As in Melitz (2003), a continuum of firms produce differentiated goods under monopolistic competition and decide whether to pay a fixed marketing cost to serve each particular market. We extend this model by allowing each firm to produce anywhere in the world, albeit at varying productivity levels. Faced with costs of exporting and with costs of producing outside of its home market, each firm decides which markets to serve and where to locate production to serve those markets. Our choice for the functional form of the distribution from which firms draw their productivity upon entry leads to a parsimonious characterization of firm choices and aggregate trade and MP flows.

2.1 The Environment

We consider a world economy comprising 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 Dixit-Stiglitz preferences with elasticity of substitution $\sigma > 1$. The associated price index is given by

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

where $p_i(\omega)$ is the price of good ω. We adopt the convention that $p_i(\omega) = +\infty$ if good ω is not available in country i.

Each good ω is potentially produced by a single firm under monopolistic competition. 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. A firm from country i can serve country n by (a) producing in i and exporting to country n, by (b) opening an affiliate in country $l \neq i, n$ and exporting from there to country n, or by (c) 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 specific and location specific. In particular, a firm is distinguished by a productivity vector $\boldsymbol{z} = (z_1, z_2, ..., z_N)$. Here z_l determines the firm's productivity if it decides to produce in country l, as explained below.

Firms from l that sell in country n incur a "marketing" fixed cost $w_n F_n$ and an "iceberg" transportation cost of $\tau_{ln} \geq 1$. We assume that $\tau_{nn} = 1$ and that the triangular inequality holds (i.e., $\tau_{il}\tau_{ln} \geq \tau_{in}$). Moreover, we assume the existence of bilateral "iceberg" multinational production (MP) costs $\gamma_{il} \geq 1$ with $\gamma_{ll} = 1$. Letting $\xi_{iln} \equiv \gamma_{il} w_l \tau_{ln}$, these assumptions imply that a firm from i producing in location l in order to serve market n has unit cost $C_{iln} \equiv \xi_{iln}/z_l$ and a fixed cost of $w_n F_n$. Note that all heterogeneity across firms is associated with differences in the productivity vector $\mathbf{z}$, while the trade and MP costs $\{\tau_{ln}\}$ and $\{\gamma_{il}\}$ as well as wages (and hence ξ_{iln}) is common across firms.

The productivity vector of firms in country i is randomly assigned according to the multivariate distribution given by

$$\Pr(Z_1 \leq z_1, ..., Z_N \leq z_N) = G_i(z_1, ..., z_N) = 1 - \left(\sum_{l=1}^{N} \left[T_{il} z_l^{-\theta} \right]^{\frac{1}{1-\rho}} \right)^{1-\rho} \tag{2}$$

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 $\theta > \sigma - 1$.[8] Several comments are in order. First, the marginals have conditional distributions that are Pareto. In particular,

$$\Pr(Z_l \leq z_l) = \lim_{x \to \infty} G_i(x, ..., z_l, ..., x) = 1 - T_{il} z_l^{-\theta},$$

so 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, in the limit as $\rho \to 1$ we have $G_i(z_1, ..., z_N) = 1 - \max_l T_{il} z_l^{-\theta}$.[9] In this case, the elements of $\boldsymbol{z}$ are perfectly correlated. Finally, if $\rho = 0$, then 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 $k \neq l) = \left(T_{il}/\widetilde{T}_i\right)\left(1 - \widetilde{T}_i z_l^{-\theta}\right)$.[10] This case is equivalent to simply having the production location l chosen randomly with probabilities $T_{il}/\widetilde{T}_i$ among all possible locations $i = 1, ..., N$, and then the productivity Z_i chosen from the Pareto distribution $1 - \widetilde{T}_i z_l^{-\theta}$ with $z_l \geq \widetilde{T}_i^{1/\theta}$. Figures 1 and 2 illustrate how the distribution depends on the value of ρ.

In the rest of the paper we make the following assumption:

Assumption 1 $T_{il} = T_i^e T_l^p$ and $\sum_l \left(T_l^p\right)^{1/(1-\rho)} = 1$.

This assumption implies that $\widetilde{T}_i = \left[\sum_l T_{il}^{1/(1-\rho)}\right]^{1-\rho} = T_i^e$, so 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

---

[8] This distribution can be seen as a reformulation of an Archimedean copula of Pareto distributions. Consider the copula $C(x_1, x_2) \equiv \max\left\{1 - \left[(1-x_1)^{\frac{1}{1-\rho}} + (1-x_2)^{\frac{1}{1-\rho}}\right]^{1-\rho}\right\}$. This is copula 4.2.2 in Nielsen (2006). If z_1 and z_2 are distributed Pareto with $z_l \sim 1 - T_l z_l^{-\theta}$, then the previous copula leads to distribution $G(z_1, z_2) = \max\left\{1 - \left[(T_1 z_1^{-\theta})^{\frac{1}{1-\rho}} + (T_2 z_2^{-\theta})^{\frac{1}{1-\rho}}\right]^{1-\rho}, 0\right\}$. The support of this distribution is 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 Nelsen, 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.

[9] Let $x \equiv \max_l T_{il} z_l^{-\theta}$ and note that $G_i(z_1, ..., z_N) = 1 - x\left(\sum_{l=1}^N \left[\frac{T_{il} z_l^{-\theta}}{x}\right]^{\frac{1}{1-\rho}}\right)^{1-\rho}$. As $\rho \to 1$ then $\left[\frac{T_{il} z_l^{-\theta}}{x}\right]^{\frac{1}{1-\rho}} \to 0$ for all l except $v \equiv \arg\max_l T_{il} z_l^{-\theta}$, for which $\left[\frac{T_{iv} z_v^{-\theta}}{x}\right]^{\frac{1}{1-\rho}} = 1$ for all ρ, so $\sum_{l=1}^N \left[\frac{T_{il} z_l^{-\theta}}{x}\right]^{\frac{1}{1-\rho}} \to 1$ and hence $G_i(z_1, ..., z_N) \to 1 - \max_l T_{il} z_l^{-\theta}$.

[10] To see this, 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 $k \neq l) = 1 - \left[\sum_{k \neq l}^N T_{ik}/\widetilde{T}_i + T_{il} z_l^{-\theta}\right]$, and this is equal to $\left(T_{il}/\widetilde{T}_i\right)\left(1 - \widetilde{T}_i z_l^{-\theta}\right)$.

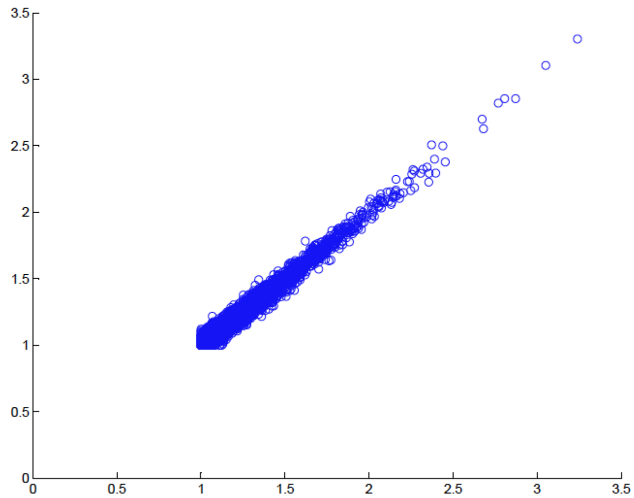

Figure 1: Simulation for 10,000 draws from $1-\left(\left(T_1 z_1^{-\theta}\right)^{\frac{1}{1-\rho}}+\left(T_2 z_2^{-\theta}\right)^{\frac{1}{1-\rho}}\right)^{1-\rho}$ with support $z_l \geq \left(T_1^{1/(1-\rho)}+T_2^{1/(1-\rho)}\right)^{(1-\rho)/\theta}$, for $T_1=T_2=2^{\rho-1}$, $\theta=7.2$, $\rho=0.9$.

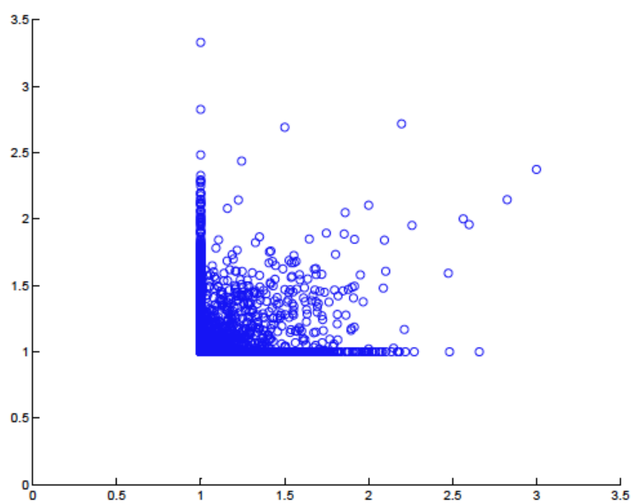

Figure 2: Simulation for 10,000 draws from $1-\left(\left(T_1 z_1^{-\theta}\right)^{\frac{1}{1-\rho}}+\left(T_2 z_2^{-\theta}\right)^{\frac{1}{1-\rho}}\right)^{1-\rho}$ with support $z_l \geq \left(T_1^{1/(1-\rho)}+T_2^{1/(1-\rho)}\right)^{(1-\rho)/\theta}$, for $T_1=T_2=2^{\rho-1}$, $\theta=7.2$, $\rho=0.1$.

country l's *productivity in production.* We will continue to write T_{il} rather than $T_i^e T_l^p$ for notational convenience. In calibrating the model we will capture forces that have specific effects on bilateral MP flows through the country-pair specific MP costs, γ_{il} (imposing $\gamma_{ii} = 1$), and leave T_i^e and T_l^p to capture productivity parameters for innovation and production that affect overall trade and MP patterns at the country level.

We will consider two cases regarding firm entry. One case entails exogenous entry, which implies that the measure of entrants in each market i is exogenous. For this case we disregard entry costs and simply assume that all workers are engaged in production. The other case entails endogenous entry. In particular, the measure of entrants in each market i is determined so that the expected profits are equal to the cost of entry, $w_i f_i^e$.

2.2 Firm's Problem

We can think of the firm's problem as follows. First, for each market n a firm decides what is the cheapest location from which to serve that market, the solution of $\arg\min_l C_{iln}$. Second, the firm decides what price to charge. Given our assumption for preferences, this choice leads to the a mark-up of $\widetilde{\sigma} \equiv \sigma / (\sigma - 1)$ over marginal cost, so the price is

$$p_{in} = \widetilde{\sigma} \min_l C_{iln} \ . \tag{3}$$

Third, the firm calculates the associated profits and if these profits are higher than the fixed marketing cost $w_n F_n$ then the firm chooses to serve market n. Therefore, a firm from i will serve market n if and only if $\min_l C_{iln} \leq c_n^*$, where c_n^* is the maximum unit cost under which gross profits in market n are enough to cover $w_n F_n$, and is defined by

$$c_n^* = \left(\frac{\sigma w_n F_n}{X_n} \right)^{1/(1-\sigma)} \frac{P_n}{\widetilde{\sigma}}, \tag{4}$$

where X_i is total expenditure in country i.

We assume that for all pairs $\{i, n\}$ there are firms from i that will decide not to serve market n. Since all productivity draws in country i are higher than or equal to $\widetilde{T}_i^{1/\theta}$, this is guaranteed by the following condition, which we maintain throughout the rest of the paper:

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

The following result is a key ingredient in the analysis that follows:

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 proofs of all the results in the paper are provided in the Appendix. We now turn to the model's implications for aggregate trade and MP flows.

2.3 Aggregate implications

Let M_i denote the measure of firms in country i, let M_{iln} denote the measure of firms from i that serve market n from location l, and let X_{iln} denote the total value of the associated sales. Using the pricing rule in (3) and the cut-off rule in (4) together with the results of Lemma 1 we can show (see Appendix) that

$$M_{iln} = \frac{\theta - \sigma + 1}{\sigma\theta}\frac{X_{iln}}{w_n F_n} \tag{7}$$

and

$$X_{iln} = \psi_{iln}\lambda_{in}^E X_n, \tag{8}$$

where

$$\lambda_{in}^E \equiv \frac{\sum_l X_{iln}}{X_n} = \frac{M_i \Psi_{in}}{\sum_k M_k \Psi_{kn}}, \tag{9}$$

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).

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 (ignoring the origin of firms), $\lambda_{ln}^T \equiv \sum_i X_{iln}/\sum_{i,k} X_{ikn}$, while MP shares are given by production shares across firms from different origins (ignoring the destination of that production), $\lambda_{il}^M \equiv \sum_n X_{iln} / \sum_{j,n} X_{jln}$. Letting $Y_l \equiv \sum_{i,n} X_{iln}$ denote the value of all goods produced in country l (output) and recalling that $X_n \equiv \sum_{i,l} X_{iln}$ is total expenditure by consumers in country n, trade and MP shares can be written more succintly as $\lambda_{ln}^T = \sum_i X_{iln}/X_n$ and $\lambda_{il}^M = \sum_n X_{iln}/Y_l$. Using expression (8) we immediately obtain

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

and

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

Let Π_{iln} denote aggregate profits net of fixed marketing costs but gross of entry costs associated with sales X_{iln}. 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 (7), 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. $\sum_{i,n} \Pi_{iln} = \eta Y_l$.

Marketing wages paid in n are $\sum_{i,l} M_{iln} w_n F_n$ while wages in production in n are $Y_n/\tilde{\sigma}$. So total wages paid to workers in production and marketing in country n are $\sum_{i,l} M_{iln} w_n F_n + Y_n/\widetilde{\sigma}$. Letting L_n^p denote the amount of labor devoted to production and marketing in country n, and using (7), we can then write the labor market clearing condition for workers in production and marketing in country n as

$$\frac{\theta - \sigma + 1}{\sigma\theta} (X_n - Y_n) + (1 - \eta) Y_n = w_n L_n^p. \tag{13}$$

For future reference, note also that, using (11) and (12) and adding over l and n, total profits made by firms from country i are given by

$$\sum_{l,n} \Pi_{iln} = \eta \sum_n \lambda_{in}^E X_n \ . \tag{14}$$

2.4 Equilibrium

The set of equilibrium conditions depends on whether we have exogenous or endogenous entry. The current account balance and the labor market clearing conditions are used to determine equilibrium in both cases. However, under exogenous entry all workers are engaged in production and marketing whereas under endogenous entry labor can also be used for innovation. In the case of endogenous entry, a zero-profit condition is also used to solve for the equilibrium since the measure of entrants (M_i) is endogenous.

We start by characterizing the current account balance.[11] 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 country $i's$ goods, $\sum_{j,n} M_{jin} w_n F_n$; (ii) wages paid to workers engaged in marketing for sales in country i, $\sum_{j,l} M_{jli} w_i F_i$; (iii) net profits, which are equal to profits made by domestic firms, $\sum_{l,n} \Pi_{iln}$, minus profits made domestically by foreign firms, $\sum_{j,n} \Pi_{jin}$. Thus, we can write the current account balance condition (i.e., total expenditure equals total income) as

$$X_i = Y_i - \sum_{j,n} M_{jin} w_n F_n + \sum_{j,l} M_{jli} w_i F_i + \sum_{l,n} \Pi_{iln} - \sum_{j,n} \Pi_{jin}. \tag{15}$$

Using (12), (8), and (13) we can rewrite this condition as

$$X_i = w_i L_i^p + \eta \sum_n \lambda_{in}^E X_n. \tag{16}$$

Next, consider the labor market clearing condition. Total output in country i is $Y_i = \sum_n \lambda_{in}^T X_n$, hence we can rewrite (13) as

$$\frac{\theta - \sigma + 1}{\sigma\theta} X_i + \frac{1}{\tilde{\sigma}} \sum_n \lambda_{in}^T X_n = w_i L_i^p. \tag{17}$$

Notice also that λ_{in}^E and λ_{in}^T are functions of wages, $\boldsymbol{w}$, and entry levels, $\boldsymbol{M}$ (where variables in bold are used to denote vectors)

Under exogenous all workers are engaged in production/marketing, $L_i^p = L_i$. Thus, equations

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

(16) and (17) 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).

Under endogenous entry, labor used in production/marketing and entry must add up to the total labor supply so that the labor market clearing condition is

$$L_i^p + M_i f_i^e = L_i. \tag{18}$$

Together with (17) we get the labor market clearing condition under endogenous entry,

$$\frac{\theta - \sigma + 1}{\sigma\theta} X_i + \frac{1}{\tilde{\sigma}} \sum_n \lambda_{in}^T X_n + w_i M_i f_i^e = w_i L_{i.} \tag{19}$$

Equilibrium entry, M_i, is determined by the zero-profit condition, namely $\sum_l \Pi_{il} = M_i w_i f_i^e$ (we assume throughout the paper that the equilibrium is interior, so that $L_i^p < L_i$ and $M_i > 0$). Using (14), this condition can be written as

$$\eta \sum_n \lambda_{in}^E X_n = M_i w_i f_i^e. \tag{20}$$

Equations (16), (19) and (20) constitute a system of $3N$ equations to solve for equilibrium levels of $\boldsymbol{X}$, $\boldsymbol{M}$, $\boldsymbol{w}$ (up to a constant determined by the numeraire).

For future reference, note that (18) together with equations (16) and (20) imply that under endogenous entry all income takes place through wages so that

$$X_i = w_i L_i \ . \tag{21}$$

Moreover, letting r_i denote the share of labor devoted to innovation in country i, $r_i \equiv 1 - L_i^p / L_i$, then (19) together with (21) implies that

$$r_i - \eta = \frac{1}{\tilde{\sigma}} \left(\frac{X_i - Y_i}{X_i} \right) \ . \tag{22}$$

2.5 Special Cases

We now turn to present a number of special cases of the model that we can characterize analytically. These cases will illustrate the basic forces behind the results of our quantitative analysis in a later section. Focusing on these cases will allow us to (1) establish a benchmark against which to compare the impact of MP, (2) show the role of comparative advantage in innovation vs. production, and (3) explore the role of home market effects (HME). In all the special cases we consider here we assume there is endogenous entry.

2.5.1 Infinite MP costs - the no MP benchmark

It is instructive to consider the extreme 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 $\Psi_{in} = T_{ii} \left(w_i \tau_{in}\right)^{-\theta}$. Given Assumption 1, and using (10) we obtain the expression for bilateral trade shares,

$$\lambda_{in}^T = \frac{M_i T_i^e T_i^p \left(w_i \tau_{in}\right)^{-\theta}}{\sum_k M_k T_k^e T_k^p \left(w_k \tau_{kn}\right)^{-\theta}}. \tag{23}$$

These shares are just like in Eaton and Kortum (2002), but instead of the Frechet technology parameter we now have $M_i T_i^e T_i^p$.

With infinite MP costs, the equilibrium conditions under endogenous entry imply that entry is equal to

$$\widetilde{M}_i = \eta L_i / f_i^e. \tag{24}$$

This result is important because it shows that, with no MP, trade has no effect on the share of labor devoted to innovation in any country. More precisely, for all i we have $M_i f_i^e / L_i = \eta$, so the share of labor devoted to innovation is independent of trade costs $\{\tau_{ln}\}$ and also independent of entry costs $\{f_i^e\}$. This is reminiscent of the results of Eaton and Kortum (2001) and is consistent with the results of Arkolakis, Costinot, and Rodríguez-Clare (2010), who show that trade has no impact on entry in the general-equilibrium endogenous-entry version of the Melitz/Chaney model.

2.5.2 A frictionless world - the role of comparative advantage

We now discuss the role of comparative advantage in innovation vs. 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 for the equilibrium shares of labor devoted to innovation:

Proposition 1 *Consider a frictionless world under endogenous entry with $f_i^e = f^e$ for all i. Assume that for all i we have*

$$1 - (1-\eta)\,\widetilde{\sigma} < \frac{A_i / \left(T_l^e\right)^{\theta/(1-\rho)+1}}{\sum_k \delta_k A_k / \left(T_k^e\right)^{\theta/(1-\rho)+1}} < 1 + \eta\widetilde{\sigma}. \tag{25}$$

The share of labor devoted to innovation in country i is

$$r_i \equiv \frac{L_i^e}{L_i} = \frac{1}{\widetilde{\sigma}} \left(1 - \frac{A_i / \left(T_l^e\right)^{\theta/(1-\rho)+1}}{\sum_k \delta_k A_k / \left(T_k^e\right)^{\theta/(1-\rho)+1}} \right) + \eta. \tag{26}$$

The condition in (25) guarantees that innovation shares in (26) satisfy $0 < r_i < 1$. If (25) is not satisfied, then at least one country would be completely specialized in innovation or production, i.e., $L_i^e = 0$ or $L_i^e = L_i$ for some i.

Proposition 1 summarizes how the different parameters 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., $r_i > \eta$. The countries that have comparative advantage in innovation will also have a trade deficit (i.e., $X_i > Y_i$) as can be seen in equation (22).

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

Under endogenous entry but with positive trade and MP costs, our model exhibits home market effects that affect the allocation of production and innovation across large and small countries. To illustrate this home market effects in the simplest way, consider a world with two countries that are symmetric except for size. We can obtain some analytical results for two extreme cases, one with frictionless trade and the other with frictionless MP.

Proposition 2 *Consider a two-world country under endogenous entry. Assume that* $A_1 = A_2$, $T_1^e = T_2^e = T^e$, $f_1^e = f_2^e = f^e$, *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 we have* $r_1 > r_2$.
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 we have* $r_1 < r_2$.

The first part of the proposition shows the existence of a home market effect 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 in the case with frictionless MP. In that case since MP is frictionless but trade is costly, it makes sense to have the large country specialize in production.

2.6 Welfare Implications

In this section we will illustrate that the model gives simple and intuitive expressions for the gains from trade and multinational activity of firms. These expressions will be used to study the welfare implications of openness to trade and multinational production and are derived in appendix (B.5).

2.6.1 Gains from Openness

By comparing the real expenditure per capita to the one in isolation we can compute a formula for the gains from openess (i.e., the change in welfare as we move from isolation to the actual equilibrium),

$$GO_n = \left(\frac{X_{nnn}}{X_n}\right)^{-\frac{1-\rho}{\theta}} \left(\frac{\sum_l X_{nln}}{X_n}\right)^{-\frac{\rho}{\theta}} \left[\boldsymbol{\chi}\left(\frac{\frac{1}{\tilde{\sigma}} + \frac{1}{\sigma} - \eta}{\frac{1}{\tilde{\sigma}}\frac{Y_n}{X_n} + \frac{1}{\sigma} - \eta}\right)^{1+\frac{1}{\theta}\frac{\theta-\sigma+1}{(\sigma-1)}} + (1-\boldsymbol{\chi})\left(\frac{M_n}{\widetilde{M}_n}\right)^{1/\theta}\right], \tag{27}$$

where $\boldsymbol{\chi}$ is an indicator variable that takes the value of 1 under exogenous entry and 0 under endogenous entry. This expression relates the gains from openness to observable variables together with parameters ρ and θ. In fact, when there is no MP we can show that the gains from openness (i.e., the gains of 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 (2010).[12]

If $\rho = 0$ then the first two terms of the RHS of (27) collapse to $(X_{nnn}/X_n)^{-1/\theta}$. The term X_{nnn}/X_n is an inverse measure for the degree of openness of country n. As one would expect, this measure implies that a country is more open with MP than without it, since $X_{nnn}/X_n < \lambda_{nn}^T = \sum_i X_{inn}/X_n$. To understand what happens 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 RHS 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. Taking as given the equilibrium flows X_{iln}, $\rho > 0$ leads to lower gains than $\rho = 0$. The reason is that, if productivity draws are correlated, the gains associated with the second term are not as important.

Compared to the result for gains from openness in the perfect competition setup of Ramondo and Rodríguez-Clare (2009), we now have the extra term

$$\chi\left(\frac{\frac{1}{\tilde{\tilde{\sigma}}}+\frac{1}{\sigma}-\eta}{\frac{1}{\tilde{\tilde{\sigma}}}\frac{Y_n}{X_n}+\frac{1}{\sigma}-\eta}\right)^{1+\frac{1}{\theta}\frac{\theta-\sigma+1}{(\sigma-1)}} + (1-\chi)\left(\frac{M_n}{\widetilde{M}_n}\right)^{1/\theta}.$$

The first term captures the gains associated with the profit channel under exogenous entry: countries with a net profit inflow due to a net outflow of MP have $X_n/Y_n > 1$ and this increases real expenditure per capita directly and indirectly through its effect on domestic variety. The second term captures the effect of MP on entry. To express this effect in terms of observable variables, note that under endogenous entry we have $X_n = w_n L_n$. We can use the labor market

[12] As we showed above, even if M_n is endogenous, in the absence of MP, M_n is not affected by trade, so M_n is the same in isolation and in the equilibrium with trade but no MP.

clearing condition, the definition of r_n and L_n^e, and equation (22) to obtain

$$\frac{M_n}{\widetilde{M}_n} = \frac{L_n^e/f_n^e}{\eta L_n/f_n^e} = \frac{r_n}{\eta} = \theta\left(\frac{X_n - Y_n}{X_n}\right) + 1 \tag{28}$$

This expression implies that countries with net outward MP flows (e.g. the United States) will have $X_n > Y_n$ and will experience an increase in entry as a result of openness. For these countries, given equation (27), our monopolistic competition setup implies larger gains from openness than the perfect competition model of Ramondo and Rodríguez-Clare (2009) while the opposite conclusion is true for countries with $X_n < Y_n$ (e.g., Ireland).

2.6.2 Multinational Production and Real Wages

As mentioned in the Introduction, there is widespread concern that the globalization of production by U.S. firms may have a detrimental effect on domestic workers. In this subsection we explore whether this effect is possible in our model. 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 high-productivity firms (under exogenous entry) or a comparative advantage in innovation (under endogenous entry). To make the analysis more illustrative, we consider the cases of exogenous and endogenous entry separately and focus on the comparative statics of a move from a situation with frictionless trade but no MP to a situation with both frictionless trade and frictionless MP.

For 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 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 $m_i \equiv M_i T_i^e / L_i$. We will assume that countries differ only in m_i, so we impose that $A_i \equiv \left(T_i^p\right)^{1/(1-\rho)} / L_i = A$ for all i. This assumption implies that productivity in production is the same across countries. Thus, if $m_i = m$ for all i, then wages would also be the same across countries.

Proposition 3 *Consider the case with exogenous entry and assume without loss of generality that* $T_i^e = A_i = 1$ *and* $\rho \to 1$, $m_i \leq \frac{L_i}{\sum_k L_k} \frac{(\theta+1)(\theta\sigma-\sigma+1)}{(\theta-\sigma+1)}$ *for all* i. *Assume 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 wages iff $\sigma < \bar{\theta} \equiv \frac{(1+\theta)^2}{1+\theta+\theta^2}$,

ii) increases real profits and real expenditure, for any value of σ.

Opening multinational production implies a downward pressure to the nominal wages of the "idea abundant" countries since firms now have the ability to locate where cheaper labor exists. In fact, under the condition of m_i versus l_i all countries devote labor to production and wages equalize under free trade. The resulting higher nominal profits imply that more varieties are potentially consumed in the country, which decreases the country's price index. Real wages will increase if the elasticity of substitution is low enough so that the price index declines compensate for the decrease in real wages. The positive effect on real profits is so strong that also implies that overall real expenditure is always increasing.

A key assumption of the exogenous entry setup is that ideas can enter the market without cost. In the free entry case new ideas require the use of labor for the production of the entry cost. In this situation we can prove a stronger result for the beneficial role of MP to real wages and expenditures.

Proposition 4 *Consider the case of endogenous entry and assume that condition* (25) *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 expenditures).*

Comparing Propositions 3 and 4 reveals that the results of a decline in MP costs critically depend on whether entry is exogenous or endogenous. The possibility of a negative effect of MP on wages in countries with a high $M_i T_i^e / L_i$ ratio arises because the same number of workers in L_i now have fewer goods that are produced there – but if entry is endogenous, then a natural outcome is that workers engaged in production will move to innovation. By decreasing the supply of labor to the production sector, this leads to an improvement in the country's terms of trade, and this mechanism is what allows the country to avoid a decline in real wages under endogenous entry.

3 Calibration

We fit our model with endogenous entry to international expenditure, production, trade, and multinational sales data. We first discuss how can we obtain information about the level of technological parameters θ and ρ by looking at the elasticity of expenditure by country n on goods produced in country l with respect to trade friction between n and l. Given this information we describe a methodology to estimate the technology parameters, θ and ρ, the $N \times 1$ vectors $\boldsymbol{T}^e$ and $\boldsymbol{T}^p$, and the $N \times (N-1)$ trade and MP frictions, τ and γ, using data on endowments (the $N \times 1$ vector of equipped labor), the $N \times (N-1)$ matrix of trade shares and MP shares λ^M.

3.1 Gravity and Trade Elasticities

Loosely speaking, the value of θ governs the substitutability across products of heterogeneous firms from a given origin and the value of ρ governs the substitutability across different production locations for a given firm. To infer the value of these parameters, we will consider the trade elasticity estimated from two distinct gravity equations.

The first gravity equation, which is unique to our model of trade and MP, 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 a sample restricted to firms that originate in a particular i (here, the United States), we refer to this equation as "restricted gravity." The second gravity equation is defined over $X_{ln} \equiv \sum_i X_{iln}$, the sales of *all* firms that are operating in country l and selling in country n. Because this gravity equation is defined over firms from all countries, we refer to this equation as "unrestricted gravity."

3.1.1 Restricted Gravity

To estimate the restricted gravity equation, we use expression (8 –see also equation (35) in the Appendix–) and take logarithms to obtain

$$\ln X_{iln} = \alpha^r_{il} + \mu^r_{in} - \frac{\theta}{1-\rho} \ln \tau_{ln}, \tag{29}$$

where α_{il}^r is a location of production fixed effect that corresponds (in the model) to

$$\alpha_{il} = \ln\left(M_i\left[T_i^e T_l^p (w_l \gamma_{il})^{-\theta}\right]^{\frac{1}{1-\rho}}\right)$$

and μ_{in}^r is a country of destination fixed effect that corresponds (in the model) to

$$\mu_{in}^r = \ln\left(\frac{X_n \Psi_{in}^{\frac{-\rho}{1-\rho}}}{\sum_k M_k \Psi_{kn}}\right).$$

Equation (29) 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}. To estimate $\theta/(1-\rho)$ we exploit the fact that it affects the relationship between X_{iln} and τ_{ln}.

A difficulty of operationalizing (29) is that we must have an accurate measure of the relative size of trade frictions between countries l and n. The standard practice in the gravity literature is to use a proxy for τ_{ln} such as distance or shared language. However, such practice does not reveal the structural parameters of interest as the coefficient estimate on the relevant trade friction conflates the variation of τ_{ln} with the proxy and the trade elasticity of interest.[13]

We rely on a measure of the size of trade costs that is directly related to a critical component in τ_{ln}, which is the asymmetric treatment across locations of production in the tariffs applied to goods. Specifically, we operationalize equation (29) by parameterizing trade costs as

$$\ln \tau_{ln} = \ln(1 + t_{ln}) + \sum_k \delta_k [1|d_{ln} \in d_k] + \Theta H_{ln} + e_{iln},$$

where t_{ln} is the simple average tariff applied by n on goods from l, $[1|d_{ln} \in d_k]$ indicator variables for a given distance between n and l whose marginal effect on trade cost is given by δ_j, and H_{ln} is a vector of standard gravity controls, including a shared language, shared colonial history, a shared border, and a "border effect" indicator variable, called *self*, that is equal to one if $l = n$.

[13] This fact led Eaton and Kortum (2002), Donaldson (2008), 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 variations for the same purpose so we need to resort to different ways of measuring τ_{ln}.

This yields the "restricted" gravity equation,

$$\ln X_{iln} = \alpha_l^r + \mu_n^r + \beta^r \ln(1 + t_{ln}) + \sum_k \widetilde{\delta}_k^r [1 | d_{ln} \in d_k] + \widetilde{\Theta}^r H_{ln} + \widetilde{e}_{iln}, \tag{30}$$

that we estimate. To the extent that constructed measures of t_{ln} accurately capture variation in asymmetric trade frictions between countries, the coefficient β^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 the effect of the variable on the size of trade cost and $\theta/(1-\rho)$. Because in our data there will be multiple observations for each production location l and for each destination country n, we can estimate equation (30) via least squares with dummy variables.

To estimate equation (30) 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, a composite of fourteen European Union countries, Japan, and the United Kingdom. We also observe the domestic sales of U.S. firms in the United States (netting out 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 in the data 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 firms from l pay country n's MFN applied tariff rates).[14]

There are several concerns that arise in using tariff data to estimate the trade elasticity. First, there is the problem of endogeneity: country pairs for which there is a natural affinity for trade are more likely to agree to preferential trading arrangements. For this reason, it is important that we include standard gravity controls in (30) which proxy for this affinity. To the

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

extent that there are other determinants of preferential trading agreements that are excluded from (30), there may be an upward bias in the size of the trade elasticity. 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 tend to be either ad hoc or have an element of endogeneity to them. We plan to explore alternative aggregation schemes in future versions of the paper. To the extent that the level of policy induced trade frictions between countries is seriously mismeasured our estimate of β^r will be biased downward. Finally, we include *self* to control for the variation in t_{ln} that is due to unmeasured border effects, such as administrative and information costs, that local production avoids.

3.1.2 Unrestricted Regression

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

$$\ln X_{ln} = \alpha_l^u + \mu_n^u + \beta^u \ln(1 + t_{ln}) + \sum_k \widetilde{\delta}_k^u [1|d_{ln} \in d_k] + \widetilde{\Theta}^u H_{ln} + v_{iln}. \tag{31}$$

The coefficient estimate $\hat{\beta}^u$ from the "unrestricted" gravity equation does not have a structural interpretation, but it can still provide information on the relative magnitudes of θ and ρ. To see this, recall that if MP were not possible then all exports would be done by local firms and the coefficient on tariffs would be equal to θ. This is because, without MP, the correlation in productivity across production locations (determined by ρ) is irrelevant, so the trade elasticity is given by θ, as in the standard Melitz/Chaney model. Now, since in the data most exports are done by domestic firms, then X_{ln} disproportionately contains information on the operations of domestic firms, and this suggests that $\hat{\beta}^u$ will be closer to θ than is the case for $\hat{\beta}^r$, which should be higher and equal to $\theta/(1-\rho)$. In other words, we expect that $0 > -\theta > \hat{\beta}^u > \hat{\beta}^r = -\theta/(1-\rho)$. Another benefit to estimating the unrestricted gravity equation is because doing so will allow us to compare the trade elasticity implied by our tariff-based methodology to elasticities obtained by the price-gap methodologies as mentioned above.

We estimate (31) using data on trade volumes of manufacturing industries and domestic absorption. We restrict the sample so that the coverage of the restricted and unresticted samples (l, n) is the same.

3.1.3 Results

The coefficient estimates for the two regressions are reported in Table 1.

		Distance Dummies					Other Gravity Controls				
	Tariff	D2	D3	D4	D5	D6	Self	Bord	Lang	Col	R-sq.
Restricted	-10.8	-0.4	-2.5	-3.2	-2.5	-3.2	2.2	0.3	0.3	0.3	0.84
	(-3.1)	(-0.8)	(-4.8)	(-6.3)	(-4.9)	(-5.5)	(3.4)	(0.6)	(1.0)	(0.8)	
Unrestricted	-4.1	-0.7	-1.3	-2.0	-1.6	-1.5	3.6	1.1	-0.3	0.1	0.89
	(-2.1)	(-2.7)	(-4.5)	(-7.2)	(-5.9)	(-4.6)	(10.0)	(3.7)	(-1.7)	(0.4)	

Table 1: Restricted and Unrestricted Gravity

In Table 1 each column corresponds to a dependent variable, while the first and second rows correspond to the restricted and unrestricted specifications, respectively. T-statistics are shown in parentheses under their respective coefficient estimate. 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 has 317 observations. The trade elasticity in the restricted regression of 10.8 is our estimate of $\theta/(1-\rho)$. Note that the trade elasticity in the unrestricted regression is significantly smaller at 4.1. This smaller value implies that the exports of multinational firms are far more sensitive to trade costs than those of domestic firms as expected. This greater sensitivity of the sales of multinational affiliates to trade costs than for domestic firms also appears in the coefficient estimates on the distance categories. For each distance class, the coefficient is more negative in the restricted regression than in the unrestricted regression.

The coefficient on tariffs of 4.1 obtained from the unrestricted gravity estimation is of a remarkably similar magnitude to the trade elasticity obtained from the price gap literature. Eaton and Kortum (2002) find that the elasticity varies between 2 and 12, while using a refined methodology Simonovska and Waugh (2009) find that the trade elasticity is in the neighborhood of 4. The similarity between our estimate from the unrestricted gravity regression to that obtained from price-gap methodology gives us some confidence in the restricted entry coefficient of 10.8.

3.2 The Estimation Algorithm

We restrict our analysis to the set of nineteen OECD countries considered by Eaton and Kortum (2002): Australia, Austria, Belgium, Canada, Denmark, Spain, Finland, France, United Kingdom, Germany, Greece, Italy, Japan, Netherlands, Norway, New Zealand, Portugal, Sweden, and the United States. We use STAN data on manufacturing trade flows from country l to country n as the empirical counterpart for trade in the model, X_{ln} for $n \neq l$, and STAN production data (less aggregate exports) for X_{nn}. We use these data to calculate the $N \times N$ matrix of trade shares λ_{ln}^{T} and the $N \times 1$ vector of aggregate expenditure by country X_n. Further, we calculate the $N \times 1$ vector of aggregate outputs Y_l from $Y_l = \sum_n \lambda_{ln}^{T} X_n$. We use UNCTAD data on the gross value of production for multinational affiliates from i in l as the empirical counterpart of bilateral MP and use this data to calculate the $N \times N$ matrix of ownership shares λ_{in}^{M}. Finally, we measure the $N \times 1$ vector of endowments L_i as equipped labor.[15]

We need to set values for θ, ρ, and σ. For our model to match the restricted elasticity estimated above we impose $\theta / (1 - \rho) = 10.8$. We then set the values of σ and θ so as to do well along three dimensions: the implied mark-up for each firm, $\sigma / (\sigma - 1)$, the implied profit share, $\eta \equiv \frac{1}{\tilde{\sigma}\theta}$, and the unrestricted elasticity implied by the calibrated model. The estimates of Martins, Scarpetta, and Pilat (1996) for the average mark-up across OECD countries are in the range of 13% to 26%, implying a markup of around 20%, while in the United States a 15% of total income corresponds to intangible capital (see Corrado, Hulten, and Sichel (2009)). These estimates imply values $\sigma = 6$ and $\eta = 0.15$. Given $\eta = 1/(\tilde{\sigma}\theta)$, these values imply $\theta = 5.56$. But this level of θ implies an unrestricted elasticity that is too high. We compromise and choose the parameter values of $\sigma = 4$ and $\theta = 4.3$. The implication is that both the mark-up and the profit share will be a bit high ($\widetilde{\sigma} = 1.33$ and $\eta = 17.4$), but the benefit is that the calibrated model will imply an unrestricted elasticity closer to 4.

We set $f_i^e = 1$ for all i. This second assumption is a necessary normalization since f^e and T^e are not separately identified. We now turn to the estimation algorithm used to identify the remaining parameters of the model.

Preliminary calculations

[15] See Ramondo and Rodríguez-Clare (2009) for an explanation of how these equipped labor levels are constructed.

The algorithm is based on the fact that we have enough trade and MP cost parameters to exactly match bilateral trade and MP shares. Given that, we can proceed to infer a number of variables in the model using data for trade and MP shares even before calibrating the trade and MP cost parameters:[16]

1. Under endogenous entry, total wages (or national income) are given by

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

   where this equation is obtained by combining (19) and (20) with equation (11) summed over all l's.

2. In our model we assumed current account balance, which implies $X_i = w_i L_i$. But this relationship is not satisfied in the data. To proceed, 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, $w_i L_i$,

$$\Delta_i = X_i - w_i L_i.$$

3. We infer the share of labor that is devoted to innovation as implied by equation (22).

4. 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 $f_i^e = 1$ has been imposed.

5. Finally, the $N \times 1$ vector of T_l^p are chosen so that productivity in production is effectively the same across countries: $A_i = A$ for all i (recall that $A_i \equiv (T_i^p)^{1/(1-\rho)} / L_i$). Without loss of generality we set $A = 1$, so $T_l^p = L_l^{1-\rho}$.

The Estimation of Parameters

[16] In practice, the algorithm matches all these magnitudes very well, but not exactly as in the data.

To estimate the matrices of trade and MP frictions and $N \times 1$ vector of technology parameters $\boldsymbol{T}^e$, we use trade and MP shares, and a two-step iterative procedure explained below.

Step 1: Compute Trade and MP frictions

Given the above pre-determined variables and parameters and a guess of technology parameters $\boldsymbol{T}^e$, we can compute τ's and γ's. To do this, we note that the model delivers structural equations for trade and MP shares. We simply solve a system of $2N(N-1)$ equations, equations (10) and (11), in $2N(N-1)$ unknowns, the γ_{il} and the τ_{ln}. We impose $\gamma_{jj} = \tau_{jj} = 1$ for all j, but the system is not over-identified because $\sum_l \lambda_{ln}^T = \sum_i \lambda_{il}^M = 1$.

Step 2: Update the technology parameters T^e

We update our guess of technology parameters $\boldsymbol{T}^e$ by solving the system of equations from the zero-profit condition, equation (20). If convergence in $\boldsymbol{T}^e$ has not been achieved update $\boldsymbol{T}^e$ using a mixture of the initial guess and the new values and return to **Step 1**. Otherwise the procedure ends.[17]

A natural question is about the source of identification of $\boldsymbol{T}^e$'s. Indeed, one could imagine that a change in $\boldsymbol{T}^e$'s could be perfectly compensated by changing $\boldsymbol{\gamma}$'s in such a way that MP shares are not affected. This intuition can be verified in the version of our model with no trade, where it is true that the $\boldsymbol{T}^e$'s cannot be separately identified from the $\boldsymbol{\gamma}$'s. Similarly, the $\boldsymbol{T}^e$'s cannot be separately identified from the $\boldsymbol{\tau}$'s in a model with no MP. These observations imply that having both MP and trade is necessary for identification of $\boldsymbol{T}^e$'s separately from $\boldsymbol{\gamma}$'s and $\boldsymbol{\tau}$'s. To understand the identification in the case of both MP and trade, note first that for any level of $\boldsymbol{T}^e$'s, the above estimation algorithm finds the $\boldsymbol{\gamma}$'s and $\boldsymbol{\tau}$'s that make the model match the MP and trade shares. In addition, in the model with both MP and trade, the zero-profit conditions depend on λ_{in}^E's, which are different from either λ_{il}^T's or λ_{il}^M's. Thus, in a sense, the $\boldsymbol{\gamma}$'s are identified by the MP shares, the $\boldsymbol{\tau}$'s are identified by the trade shares, and the $\boldsymbol{T}^e$'s are identified by the zero profit conditions through their impact on λ_{in}^E's.

Once convergence has been achieved, we can compute an artificial data set of bilateral trade volumes, X_{ln} and trade costs τ_{ln}. We then regress the logarithm of these trade volumes on

[17] Since we are calibrating the model to data on shares, we need to specify the scale of $\boldsymbol{T}^e$ – to do so we set the value of T^e for the US to one.

location and source fixed effects and the implied trade costs τ_{ln}. The coefficient on the $\ln\tau_{ln}$ is the calibrated model's implied unrestricted trade elasticity that we will compare with the value of -4.1 that we got from the data in the unrestricted gravity regression.

4 Results

We now discuss the results of our calibration exercise. We begin by reporting summary measures of the key model parameters which include the measures of T_i^e, which capture comparative advantage in innovation, and the average inward and outward trade and MP frictions facing firms, τ and γ. To aid in understanding the workings of the model we consider a number of comparative static exercises.

By construction, the model fits the data to which it is calibrated nearly perfectly. The R-squared between the fitted and actual trade shares is 0.97 and 0.99 between the fitted and actual MP shares. The small discrepancies between the fitted and actual data arise because the estimated τ's and γ's are constrained to be no less than one in the iterative algorithm and this constraint binds for a small fraction of elements in the respective matrices.[18]

The model implies an unrestricted trade elasticity of -4.73 which is a bit higher than the one we estimated in the data ($\hat{\beta}^u = -4.1$). We now discuss two relevant dimensions to evaluate the performance of the model. First, recall that we infer the value of the innovation share r_i using the model together with data on trade and MP shares. How does this relate to R&D intensity in the data? In Figure 3 we plot a scatter of r_i/r_{US} along the horizontal axis versus the share of labor devoted to R&D in the data (relative to the US). There is a remarkable positive association between the two variables in spite of the fact that we did not use any R&D data to estimate the innovation intensity in the model.

Second, we compare the importance of exports by foreign affiliates in the model and in the data. We focus on the share of production by firms from i in country l that is sold outside of l, relative to total production by firms from i in country l. We refer to this as the BMP share of i in l, and denote it by bmp_{il}, so $bmp_{il} \equiv \sum_{n\neq l} X_{iln}/\sum_n X_{iln}$. Using BEA data we can

[18] The constraint binds in 2% and 1.5% of possible cases for trade and MP, respectively. The Netherlands stand out as a somewhat problematic case which may reflect the fact that it is a major entrepot center.

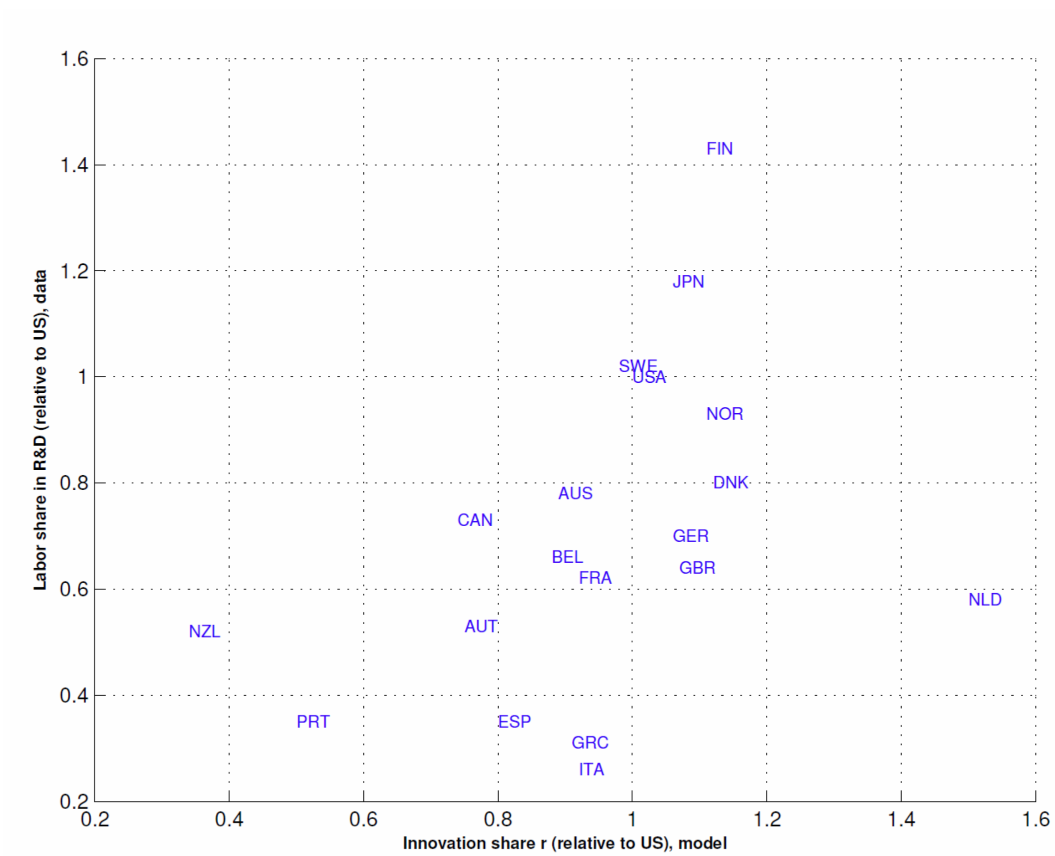

Figure 3: Innovation share in the model versus R&D intensity in the data (both as labor shares normalized to the US).

measure bmp_{il} for $i = US$. Doing that and averaging across l we obtain an overall BMP share for foreign affiliates of US multinationals, bmp_{US}, in the manufacturing sector. This comes out to be 39%. This says that, on average, foreign affiliates of US multinationals export 39% of their output. Our model comes up short in this dimension: when we compute bmp_{US} in the model we get 11%. We believe that the model's failure in this respect is due to the fact that, having no non-tradable goods, the implied trade costs are too high – preliminary calculations with a model that includes non-tradable goods do significantly better in this regard.

4.1 The Parameter Estimates

Table 2 contains summary measures of the parameters identified by the estimation algorithm. For each country in the sample, the country's innovation productivity parameter is shown in the first column. We present it as $(T_i^e)^{1/\theta}$ since this has a more natural interpretation as productivity in innovation than T_i^e. In the second and third columns, weighted average outward and inward γ's are shown, where the weights are the shares of global labor. In the fifth and sixth column are the weighted average outward and inward estimates of τ's.

It is important here to point out how the parameters are identified by the data. Figure 4 shows a scatter plot with the calibrated levels of T_i^e on the horizontal axis and the innovation shares r_i from the data on the vertical axis. The figure suggests that high levels of r_i in the data lead to a high levels of T_i^e in the calibration. Figure 5 shows scatter plots with a measure of bilateral trade or MP flows on the horizontal axis and a measure of bilateral trade or MP costs on the vertical axis. This figure suggests that trade and MP shares pin down trade and MP costs in the calibration, respectively.[19] In sum, the parameters in the model are identified in a clear and intuitive manner from the data: innovation productivity levels T_i^e are pinned down by innovation intensities in the data, r_i, and trade and MP costs, $\{\tau_{ln}\}$ and $\{\gamma_{il}\}$, are identified by trade and MP shares in the data, $\left\{\lambda_{ln}^T\right\}$ and $\left\{\lambda_{il}^M\right\}$.

As an additional check on our results, Figure 6 shows a scatter plot of the the trade costs

[19] The measure of bilateral trade flows is the log of the product of the normalized trade flows for each country pair, i.e., $\frac{\lambda_{ln}^T}{\lambda_{ll}^T}\frac{\lambda_{nl}^T}{\lambda_{nn}^T}$. An analogous measure is used for bilateral MP flows. For trade costs we use the log of the product of trade costs for each country pair, i.e., $\tau_{ln}\tau_{nl}$, and analogously for MP costs. We choose these measures because in a model without MP we would have $\frac{\lambda_{ln}^T}{\lambda_{ll}^T}\frac{\lambda_{nl}^T}{\lambda_{nn}^T} = (\tau_{ln}\tau_{nl})^{-\theta}$ so $\frac{\lambda_{ln}^T}{\lambda_{ll}^T}\frac{\lambda_{nl}^T}{\lambda_{nn}^T}$ seems to be a natural measure of bilateral trade.

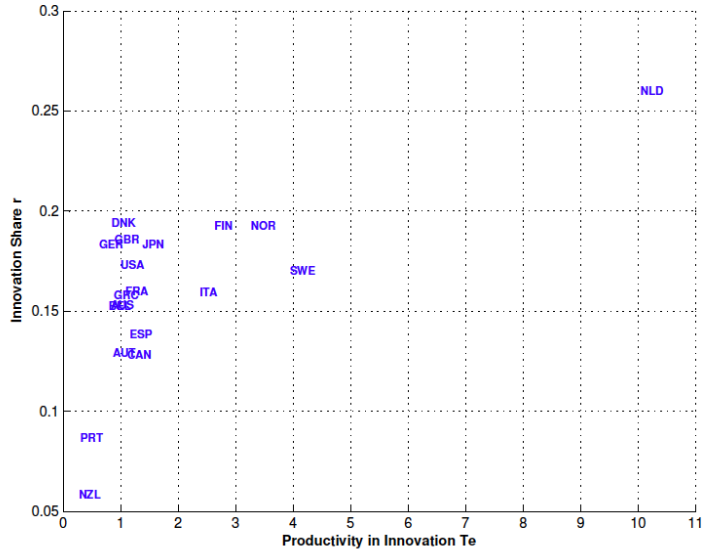

Figure 4: Data r_i against calibrated T_i^e

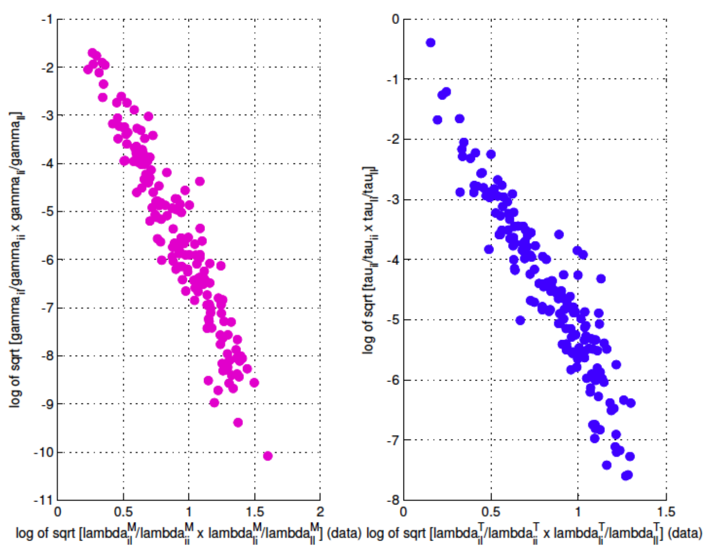

Figure 5: Normalized bilateral MP shares (left) and trade shares (right) from the data (horizontal axis) against bilateral MP and trade costs (in logs).

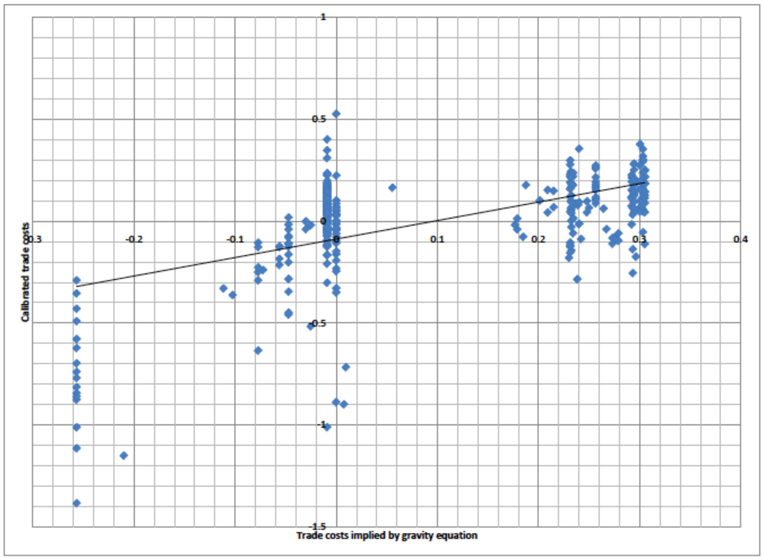

Figure 6: Trade costs implied by gravity regression versus calibrated trade costs (demeaned).

implied by the (restricted) gravity equation and our calibrated trade costs.[20] The figure suggests that the calibrated trade costs have the expected correlation with the typical gravity variables such as distance and border effects.

The seventh and eight columns of Table 2 contain the share of labor employed in innovation as implied by the data and the counterfactual share of labor employed in innovation that obtains from simulating the model holding fixed the estimated trade and MP frictions while constraining each country's innovation productivity parameter to unity.

In interpreting the estimated parameters it is useful to recall that the model interprets variation in production and ownership patterns strictly through the lens of proximity versus comparative advantage. A country can be an attractive location for innovation either because it is very capable at generating high productivity firms (T_i^e) or because it affords excellent access to a large labor force (relative to its attractiveness as a production location).

The estimates of $(T_i^e)^{1/\theta}$ vary considerably across countries ranging from a low of 0.74 for New Zealand to 1.71 for the Netherlands. The variation in these parameters gives us some sense about who has a comparative advantage in innovation versus production: the Netherlands and the Scandinavian countries stand out as having a strong comparative advantage in innovation.

[20] The trade costs implied by the restricted gravity equation are the fitted values of X_{iln} divided by the trade elasticity estimate of 10.8. Note that because the trade costs implied by the restricted gravity equation have been estimated using importer and exporter fixed effects, and so have been demeaned, the calibrated trade costs have been demeaned to make them comparable.

	$(T_i^e)^{1/\theta}$	$\boldsymbol{\gamma}$ Outward	$\boldsymbol{\gamma}$ Inward	$\boldsymbol{\tau}$ Outward	$\boldsymbol{\tau}$ Inward	r_i Calibrated	r_i With $T_i^e = 1$
Australia	0.96	2.57	3.58	3.22	2.95	0.164	0.123
Austria	0.97	3.02	3.87	2.45	3.38	0.145	0.071
Belgium	0.95	1.86	2.22	2.18	2.16	0.162	0.536
Canada	1.03	2.46	2.16	2.28	2.22	0.131	0.056
Denmark	0.96	2.69	4.10	2.68	3.49	0.196	0.246
Spain	1.04	3.16	3.31	2.33	4.08	0.147	0.088
Finland	1.25	2.94	3.72	2.00	5.30	0.187	0.092
France	1.02	2.50	2.52	2.29	2.67	0.161	0.116
Gr. Britain	0.98	2.26	2.32	2.62	2.09	0.190	0.178
Germany	0.90	1.88	2.27	2.69	1.75	0.179	0.309
Greece	0.97	3.13	6.21	2.32	7.73	0.178	0.152
Italy	1.22	3.30	2.97	1.72	5.03	0.160	0.080
Japan	1.08	2.25	2.39	1.83	3.03	0.180	0.167
Netherlands	1.71	3.07	1.38	1.78	2.65	0.254	0.001
Norway	1.32	3.06	4.30	2.05	5.93	0.201	0.005
New Zealand	0.74	2.64	5.64	3.16	3.33	0.086	0.403
Portugal	0.76	2.76	4.55	3.38	2.87	0.110	0.253
Sweden	1.38	2.86	2.57	2.10	3.73	0.166	0.001
United States	1	1.52	1.23	1.88	1.12	0.175	0.173

Table 2: Summary Statistics for Benchmark Calibration

Regarding the results for trade and MP frictions, two comments are in order. First, in general, the level of trade frictions appears to be very large. In interpreting the magnitude of these estimates, note that the parameters used in the estimating algorithm were chosen to match a relatively low trade elasticity of between 4 and 5. The lower the trade elasticity, the higher frictions between countries need to be to fit the data. Second, because of the weighting scheme, large countries will generally be more open to MP and trade because they have frictionless access to their own markets ($\gamma_{ll} = \tau_{nn} = 1$). Japan and the United States, the two largest countries, have well below average levels of average MP and trade frictions.

4.2 Comparative Advantage and Home Market Effects

The most effective way to interpret the relative importance of comparative advantage in innovation versus production and the advantages conveyed by geography is to shut down the comparative advantage in innovation by setting all $T_i^e = 1$ while holding fixed the trade and

MP frictions and then to compare across the two equilibria the share of labor that is employed in innovation. This comparison involves the seventh and eight columns of Table 2. We see that shutting down differences in T_i^e can have enormous effects on the share of a country's resources that are devoted to production. For instance, in the case of the Netherlands, the share of labor used in innovation falls from 0.25 to 0.001, which indicates that the only reason that the huge foreign network of Dutch multinationals can be rationalized given its location is by having a strong comparative advantage in innovation. In the absence of its strong comparative advantage in innovation, foreign firms would dominate Dutch production. The same outcome obtains for the other countries with initially high T_i^e, Finland, Norway, and Sweden, where geography is evidently not favorable to entry. By contrast, eliminating the implied comparative disadvantage of Belgium, Germany, and New Zealand are associated with much higher levels of innovative activity.

The comparison of the equilibrium r across countries in the last column of Table 2 reveals the importance of home market effects (HMEs), since that is the only determinant of innovation intensities once we set $T_i^e = 1$ for all i. For example, innovation in Sweden basically shuts down while it increases to 54% in Belgium. The explanation is Sweden's geography is relatively attractive as a place for production while Belgium's geography is relatively attractive as a place for innovation. New Zealand's innovation intensity also becomes quite high when we set all T_i^e to one, the reason being that given its isolation New Zealand is a bad location for production, so its resources are better deployed in innovation. Exactly the opposite happens for Canada: a low estimated cost for US firms to do business in Canada and a low cost of exporting from Canada to the United States makes Canada a good location for production, as reflected in a low r of 5.6%.

4.3 Gains from Openness

Equation (27) and (28) imply that the gains from openness can be computed from flows,

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

The gains from openness can be decomposed into a direct effect operating through the trade and MP flows and an indirect effect that captures the effect of openness on profits (in the case of exogenous entry) or innovation (in the case of endogenous entry).

One problem in computing GO_n with our calibrated model is that we have allowed for current account imbalances, and it is not clear how to think about the gains from openness in the presence of such imbalances. To proceed, we recomputed the equilibrium imposing current account balance, and then calculated the gains from openness and its decomposition into the direct and indirect effects operating through profits and innovation (see Table 3). A small and open economy like Belgium gets enormous gains, 121%, coming primarily from the direct effect. The Netherlands also enjoys large gains from openness, but being specialized in innovation, a significant part of these gains come from a positive indirect effect. The contrary occurs in New Zealand, where the gains are also large, but a negative indirect effect of -17% decreases those gains significantly.

	r	GO	Direct Effect	Profits	Innovation
Australia	0.148	1.154	1.200	0.961	0.972
Austria	0.127	1.403	1.482	0.947	0.956
Belgium	0.195	2.210	2.121	1.042	0.989
Canada	0.128	1.498	1.612	0.929	0.943
Denmark	0.189	1.353	1.325	1.021	1.027
Spain	0.138	1.123	1.184	0.948	0.955
Finland	0.191	1.261	1.235	1.021	1.021
France	0.158	1.181	1.208	0.977	0.980
Great Britain	0.183	1.246	1.230	1.013	1.017
Germany	0.177	1.189	1.188	1.001	1.005
Greece	0.158	1.078	1.107	0.974	0.980
Italy	0.160	1.079	1.100	0.981	0.980
Japan	0.183	1.049	1.034	1.014	1.014
Netherlands	0.313	1.833	1.617	1.133	1.115
Norway	0.187	1.230	1.207	1.019	1.031
New Zealand	0.065	1.390	1.670	0.832	0.887
Portugal	0.089	1.371	1.533	0.894	0.915
Sweden	0.173	1.377	1.381	0.997	0.993
United States	0.172	1.074	1.077	0.997	0.999

Table 3: Gains from Openness

4.4 Effects of Globalization

As mentioned in the Introduction, we want to explore the effects of an increase in MP by firms in rich countries. As these firms move operations to low wage locations, how does this affect workers in rich countries? How do the results depend on whether entry is exogenous or endogenous? Under endogenous entry, what is the effect on innovation in countries with comparative advantage in innovation or production? One problem with our current calibration is that it only includes OECD countries. To proceed, we replace New Zealand by a country with a size of China. In particular, we compute the equilibrium for a world that is like the calibrated model above except that now New Zealand has China's population. From now on we refer to this country as China. (We also set all current account deficits to zero, but this doesn't affect the results we report below).

		End. Entry, $\gamma_{US,CH} = 1$		Ex. Entry, $\gamma_{US,CH} = 1$	
	r (in %)	r (in %)	% change $\frac{w}{P}$	% change $\frac{w}{P}$	% change $\frac{X}{P}$
Australia	15.1	17.9	6.4	7.3	3.2
Austria	16.0	15.3	4.9	1.0	2.0
Belgium	15.5	0.5	-0.6	0.3	0.9
Canada	1.4	0.1	15.3	-0.5	-0.4
Denmark	19.2	18.5	3.7	0.3	1.3
Spain	13.5	12.7	2.5	0.3	0.8
Finland	17.9	19.4	2.7	0.3	0.0
France	15.9	13.5	2.9	0.5	0.7
Gr. Britain	13.4	14.1	2.6	0.7	1.1
Germany	4.9	16.1	0.4	0.3	0.5
Greece	13.0	16.0	5.2	1.3	0.0
Italy	16.1	15.6	2.3	0.5	0.8
Japan	17.8	16.5	2.3	0.7	1.0
Netherlands	27.8	2.4	3.8	0.2	1.2
Norway	18.5	19.3	1.2	0.3	0.0
'China'	18.8	0.1	29.8	16.1	6.5
Portugal	11.7	13.0	5.2	0.9	2.2
Sweden	18.2	11.4	0.8	-0.3	-0.6
United States	17.8	54.8	33.5	4.4	29.2

Table 4: Counterfactual Experiment

We first compute the equilibrium for this hypothetical world. The first column of Table 4 shows the results for r. China's equilibrium r is a bit higher than η, indicating that it specializes to some degree in innovation. This compares with a low r for New Zealand in the benchmark calibration (see Table 2). This illustrates the importance of HMEs in our model. The average inward gamma that we estimated for New Zealand is high relative to the country's trade costs, so as country size expands, this HMEs lead to an increase in innovation. The second column shows what happens to r as we eliminate frictions for the United States in doing MP in China, i.e. we set $\gamma_{US,CH} = 1$. Since the U.S. has a comparative advantage in innovation relative to China $((T^e_{US})^{1/\theta} = 1 > (T^e_{CH})^{1/\theta} = 0.74)$ then this leads to an increase in r in the U.S. to 0.56 and a collapse in innovation in China, where r basically becomes zero. The third column shows the effect of the decline in $\gamma_{US,CH}$ on the real wage (recall that under endogenous entry this is equal to real expenditure). As suggested by Proposition 4, the decline in MP costs for the United States in China leads to an *increase* in the U.S. real wage of 33.5%. The real wage in China increases by 29.8%. Most of the other countries experience significant increases in real wages thanks to the fact that they can now benefit from the increases in worldwide productivity arising from the possibility of using U.S. technologies to produce in China. Belgium experiences a decline in the real wage, presumably because many U.S. firms that were doing MP in Belgium now move to China. Canada experiences a particularly large increase in the real wage (15.3%), presumably because the low trade and MP costs between the United States and Canada imply that Canada benefits as the United States gets richer.

The model suggests that the negative effect on U.S. workers may indeed take place under exogenous entry. The fourth and fifth columns of Table 4 show the effects of the decline in $\gamma_{US,CH}$ on real wages and real expenditure for all the countries in our sample. The real wage actually *increases* in the United States by 4.4% while real expenditure increases much more (29.2%). The opposite happens in China: the real wage increases by more than real expenditures (16.1% vs 6.5%). The implication is that the possibility outlined in Proposition 3 does not apply in our calibrated model. We explored this further by imagining a world just like the one above but with frictionless trade. In this case it turns out that the decline in $\gamma_{US,CH}$ decreases real wages in the United States by 6% whereas real expenditure increases by 18%.

5 Conclusion

[TBD]

References

ARKOLAKIS, C., A. COSTINOT, AND A. RODRÍGUEZ-CLARE (2010): "New Trade Models, Same Old Gains?," *forthcoming, American Economic Review.*

ARKOLAKIS, C., S. DEMIDOVA, P. J. KLENOW, AND A. RODRÍGUEZ-CLARE (2008): "Endogenous Variety and the Gains from Trade," *American Economic Review, Papers and Proceedings*, 98(4), 444–450.

CHANEY, T. (2008): "Distorted Gravity: The Intensive and Extensive Margins of International Trade," *The American Economic Review*, 98(4), 1707–1721.

CORRADO, C., C. HULTEN, AND D. SICHEL (2009): "Intangible Capital and U.S. Economic Growth," *The Review of Income and Wealth*, 55(3), 661–685.

DEKLE, R., J. EATON, AND S. KORTUM (2008): "Global Rebalancing with Gravity: Measuring the Burden of Adjustment," *IMF Staff Papers*, 55(3), 511–540.

DONALDSON, D. (2008): "Railroads of the Raj: Estimating the Economic Impact of Transportation Infrastructure," Manuscript, London School of Economics.

EATON, J., AND S. KORTUM (2001): "Technology, Trade, and Growth: A Unified Framework," *European Economic Review*, 45(4-6), 1741–1780.

EATON, J., AND S. KORTUM (2002): "Technology, Geography and Trade," *Econometrica*, 70(5), 1741–1779.

——— (2007): "Innovation, Diffusion, and Trade," in *Entrepreneurship, Innovation, and the Growth Mechanism of the Free-Enterprise Economies*, ed. by E. Sheshinski, R. J. Strom, and W. J. Baumol, pp. 276–299. Princeton University Press.

FEENSTRA, R. C., J. ROMALIS, AND P. K. SCHOTT (2002): "US Imports, Exports, and Tariff Data, 1989-2001," *NBER Working Paper*, 9387.

GROSSMAN, G., AND E. HELPMAN (1991): *Innovation and Growth in the Global Economy.* MIT press, Cambridge, Massachussets.

Helpman, E. (1984): "A Simple Theory of International Trade with Multinational Corporations," *Journal of Political Economy*, 92(3), 451–471.

Helpman, E., M. J. Melitz, and S. R. Yeaple (2004): "Export Versus FDI with Heterogeneous Firms," *The American Economic Review*, 94(1), 300–316.

Irarrazabal, A., A. Moxnes, and L. D. Opromolla (2009): "The Margins of Multinational Production and the Role of Intra-Firm Trade," *CEPR Discussion Papers*, 7145.

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

——— (2002): *Multinational Firms and The Theory of International Trade.* MIT Press, Massachussetts.

Markusen, J. R., and A. Venables (1998): "Multinational Firms and the New Trade Theory," *Journal of International Economics*, 46(2), 183–204.

——— (2000): "The Theory of Endowment, Intra-Industry, and Multinational Trade," *Journal of International Economics*, 52(2), 52.

Martins, J. O., S. Scarpetta, and D. Pilat (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.

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.

Prescott, E. C., and E. McGrattan (2010): "Technology Capital and the US Current Account," *American Economic Review*, 100(4), 1493–1522.

Ramondo, N., and A. Rodríguez-Clare (2009): "Trade, Multinational Production, and the Gains from Openness," *NBER working paper*, 15604.

Rodriguez-Clare, A. (2010): "Offshoring in a Ricardian World," *American Economic Journal: Macroeconomics*, 2(2), 227–258.

Simonovska, I., and M. Waugh (2009): "The Elasticity of Trade: Estimates and Evidence," *Manuscript, University of California, Davis and New York University.*

A Data Appendix

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 Bureau of Economic Analysis (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[21] ($l \neq U.S., n$).[22] 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 Feenstra, Romalis, and Schott (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

[21] These countries are Austria, Belgium, Denmark, France, Finland, Germany, Greece, Luxembourg, Ireland, Italy, Netherlands, Portugal, Spain, and Sweden.

[22] 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.

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 tariff levels (t_{ln}) are the simple average of tariff line data where the tariffs are those applied to various country groupings. The raw tariff data was obtained from either the WTO or from the WITS web-site maintained by the World Bank. Tariffs applied by a given country n can differ from their MFN levels across exporting countries l either because no tariff is applied, as when $n = l$ or n and l are both in a free trade agreement or customs union, or because country n extends GSP tariffs to a developing country l. Data for distance (d_{ln}) and for the standard gravity controls (H_{ln}) are from the 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.[23] 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 $l \neq n$.

B Theory Appendix

B.1 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

$$\Pr\left(Z_1 \leq \frac{\xi_{i1n}}{c_{i1n}}, ..., Z_N \leq \frac{\xi_{iNn}}{c_{iNn}}\right) = 1 - \left(\sum_{k=1}^{N}\left[T_{ik}\left(\frac{\xi_{ikn}}{c_{ikn}}\right)^{-\theta}\right]^{\frac{1}{1-\rho}}\right)^{(1-\rho)}. \tag{32}$$

But we know that

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

[23] 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.

hence from (32) we get

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

Noting that

$$\Pr\left(\arg\min_k C_{ikn} = l \ \cap \min_k C_{ikn} = c\right) = \Pr\left(C_{i1n} \geq c, ..., C_{iln} = c, ..., C_{iNn} \geq c\right),$$

and letting $\Psi_{in} \equiv \left[\sum_k \left(T_{ik}\xi_{ikn}^{-\theta}\right)^{\frac{1}{1-\rho}}\right]^{1-\rho}$ we conclude that if $c < \xi_{ikn}\widetilde{T}_i^{-1/\theta}$ for all k then

$$\Pr\left(\arg\min_k C_{ikn} = l \ \cap \min_k C_{ikn} = c\right) = \theta\Psi_{in}^{-\frac{\rho}{1-\rho}} \left(T_{il}\xi_{iln}^{-\theta}\right)^{\frac{1}{1-\rho}} c^{\theta-1} = \psi_{iln}\Psi_{in}\theta c^{\theta-1},$$

where $\psi_{iln} \equiv \left(T_{il}\xi_{iln}^{-\theta}/\Psi_{in}\right)^{\frac{1}{1-\rho}}$. Given Assumption 1 we know that $c_{in}^* < \xi_{ikn}\widetilde{T}_i^{-1/\theta}$ so we can integrate 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

$$\Pr\left(\arg\min_k C_{ikn} = l \ \cap \min_k C_{ikn} \leq c_{in}^*\right) = \psi_{iln}\Psi_{in}\left(c_{in}^*\right)^{\theta}. \tag{33}$$

while

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

hence

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

QED

B.2 Derivations of formulas 7 and 8.

Multiplying (33) by the measure of firms in i, M_i, and using (4), we get the measure of firms from i that serve market n from location l,

$$M_{iln} = M_i\psi_{iln}\Psi_{in}\left(\frac{\sigma w_n F_n}{X_n}\right)^{-\theta/(\sigma-1)} \frac{P_n^{\theta}}{\widetilde{\sigma}^{\theta}}. \tag{34}$$

Since the sales of a firm with cost c in a market n are $\widetilde{\sigma}^{1-\sigma}X_nP_n^{\sigma-1}c^{1-\sigma}$, the result in (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 (4) and simplifying yields

$$X_{iln} = \frac{\widetilde{\sigma}^{-\theta}\theta}{\theta-\sigma+1}M_i\psi_{iln}\Psi_{in}\left(\sigma w_nF_n\right)^{(\theta-\sigma+1)/(1-\sigma)}X_n^{\theta/(\sigma-1)}P_n^{\theta}. \tag{35}$$

$$\begin{aligned}\frac{X_{iln}}{M_{iln}} &= \frac{\frac{\widetilde{\sigma}^{-\theta}\theta}{\theta-\sigma+1}M_i\psi_{iln}\Psi_{in}\left(\sigma w_nF_n\right)^{(\theta-\sigma+1)/(1-\sigma)}X_n^{\theta/(\sigma-1)}P_n^{\theta}}{M_i\psi_{iln}\Psi_{in}\left(\frac{\sigma w_nF_n}{\widetilde{\sigma}^{1-\sigma}X_n}\right)^{-\theta/(\sigma-1)}P_n^{\theta}} \\ &= \frac{\sigma\theta}{\theta-\sigma+1}w_nF_n\end{aligned}$$

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

$$P_n^{-\theta} = \zeta^{\theta} \left(\frac{w_n F_n}{X_n} \right)^{(\theta-\sigma+1)/(1-\sigma)} \sum_k M_k \Psi_{kn}. \tag{36}$$

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 (34) and (35), we get (7) and

$$X_{iln} = \frac{M_i \Psi_{in}}{\sum_k M_k \Psi_{kn}} \psi_{iln} X_n,$$

respectively. Defining $\lambda_{in}^E \equiv \sum_l X_{iln}/X_n$, then this last expression implies that

$$\lambda_{in}^E = \frac{M_i \Psi_{in}}{\sum_k M_k \Psi_{kn}},$$

so we finally establish (8).

B.3 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

$$\Psi_{in} = T_i^e \left[\sum_k \left(T_k^p w_k^{-\theta} \right)^{\frac{1}{1-\rho}} \right]^{1-\rho} \equiv \Psi_i \tag{37}$$

and $\psi_{iln} = \left(T_i^e T_l^p w_l^{-\theta} / \Psi_{in} \right)^{\frac{1}{1-\rho}}$, hence

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

and

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

These expressions imply that the labor-market clearing and zero-profit conditions (i.e., equations (19) and (20)) can be written as

$$\frac{1}{\widetilde{\sigma}} \frac{\left(T_i^p w_i^{-\theta} \right)^{1/(1-\rho)}}{\sum_k \left(T_k^p w_k^{-\theta} \right)^{1/(1-\rho)}} W + w_i M_i f_i^e = w_i L_{i.} \left(1 - \frac{\theta - \sigma + 1}{\sigma\theta} \right), \tag{38}$$

and

$$\eta \frac{M_i T_i^e}{\sum_k M_k T_k^e} W = M_i w_i f_i^e, \tag{39}$$

respectively, where $W \equiv \sum_j w_j L_j$. This last equation implies that

$$w_i = \eta \frac{T_i^e / f_i^e}{\sum_k M_k T_k^e} W. \tag{40}$$

Assuming that $f_i^e = f^e$ for all i and recalling that $M_i f_i^e = L_i^e$, combining (38), (39), and (40) yields

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

But $1-\frac{\theta-\sigma+1}{\sigma\theta}=\eta+1/\widetilde{\sigma}$, so letting $A_i \equiv \left(T_i^p\right)^{1/(1-\rho)}/L_i$ implies

$$r_i=\frac{1}{\widetilde{\sigma}}\left(1-\frac{f_e}{\eta}\frac{A_i/\left(T_i^e\right)^{\theta/(1-\rho)}}{\sum_k A_k L_k/\left(T_j^e\right)^{\theta/(1-\rho)}}\frac{\sum_j M_j T_j^e}{T_i^e}\right)+\eta. \tag{41}$$

Finally, notice that by the definition of r_i we have $M_i=r_i L_i/f^e$, which can be substituted in (41) to construct the term $\sum_k M_k T_k^e$, which equals to

$$\sum_k M_k T_k^e=\eta\frac{\sum_k L_k T_k^e}{f^e}\ .$$

Replacing back in (41) and defining $\delta_i \equiv L_i T_i^e/\sum_k L_k T_k^e$ we finally obtain (26) and the necessary and sufficient condition for this expression to hold, as indicated in the proposition.

B.4 Proof of Proposition 2

Part i) First, as a preliminary result, we establish that $\bar{w}\equiv w_1/w_2>1$, if $L_1>L_2$. 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 $\lambda_1^E+\lambda_2^E=\lambda_{11}^E+\lambda_{21}^E=1$.) This result also implies that $\Psi_{in}\equiv\Psi_i$ for any i,n. The zero-profit condition, equation (20), implies

$$L_1^e \ = \ \eta\lambda_1^E(L_1+L_2/\bar{w}), \tag{42}$$

$$L_2^e/\bar{w} \ = \ \eta\lambda_2^E(L_1+L_2/\bar{w}). \tag{43}$$

Using these equations together with

$$\lambda_i^E=\frac{M_i\Psi_{in}}{\sum_k M_k\Psi_{kn}}=\frac{M_i\Psi_i}{\sum_k M_k\Psi_k}$$

and $M_1 f^e=L_1^e$, we then have $\bar{w}=\Psi_1/\Psi_2$.

Letting $\phi\equiv\gamma^{-\theta/(1-\rho)}$, and using the definition of Ψ_{in} and the assumption of $A_1=A_2$ we can obtain after some derivations

$$\frac{L_1}{L_2}=\bar{w}^{\frac{\theta}{1-\rho}}\frac{\bar{w}^{1/(1-\rho)}-\phi}{1-\phi\bar{w}^{1/(1-\rho)}}\ . \tag{44}$$

The RHS of this equation is increasing in $\bar{w}$ which implies that $\bar{w}$ is increasing in L_1/L_2. Since $L_1/L_2=1$ implies $\bar{w}=1$, then $L_1/L_2>1$ implies $\bar{w}>1$, which proves the preliminary result.

Second, using this result we can now 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 (19) and equation (21) and $\lambda_{in}^T=\lambda_i^T\equiv\lambda_{ii}^T$ and $W\equiv\sum_k w_k L_k$, we have in country i

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

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

$$\frac{\lambda_1^T}{w_1 L_1} > \frac{\lambda_2^T}{w_2 L_2} \tag{45}$$

Using the expression for λ_l^T, the result $\lambda_{in}^E = \lambda_i^E$, and equations (42) and (43), after some derivations we obtain

$$L_1 r_1 + L_2 r_2 \phi \bar{w}^{\frac{\rho}{1-\rho}} > \bar{w}^{\theta/(1-\rho)} \left(\bar{w} L_1 r_1 \phi + L_2 r_2 \bar{w}^{\frac{1}{1-\rho}} \right).$$

From the definition of the trade shares we have $\lambda_1^T = \psi_{11}\lambda_1^E + \psi_{21}\lambda_2^E$, and $\lambda_2^T = \psi_{12}\lambda_1^E + \psi_{22}\lambda_2^E$ where $\psi_{iln} = \psi_{il} \equiv T_e^{1/(1-\rho)} L_l \left(\gamma_{il} w_l\right)^{-\theta/(1-\rho)} / \left(\Psi_i\right)^{1/(1-\rho)} X$. Using these equation (45) implies

$$\begin{aligned} L_1 w_1^{-\theta/(1-\rho)} \frac{\lambda_1^E \Psi_1^{-1/(1-\rho)} + \phi \lambda_2^E \Psi_2^{-1/(1-\rho)}}{w_1 L_1} &> L_2 w_2^{-\theta/(1-\rho)} \frac{\phi \lambda_1^E \Psi_1^{-1/(1-\rho)} + \lambda_2^E \Psi_2^{-1/(1-\rho)}}{w_2 L_2} \implies \\ w_1^{-\theta/(1-\rho)-1} \left(M_1 \Psi_1^{-\rho/(1-\rho)} + \phi M_2 \Psi_2^{-\rho/(1-\rho)} \right) &> w_2^{-\theta/(1-\rho)-1} \left(\phi M_1 \Psi_1^{-\rho/(1-\rho)} + M_2 \Psi_2^{-\rho/(1-\rho)} \right) \end{aligned}$$

where the last line uses the fact that $\lambda_i^E = M_i \Psi_i / \sum_j M_j \Psi_j$. Dividing by $\Psi_1^{-\rho/(1-\rho)}$ and using the above result that $\bar{w} = \Psi_1/\Psi_2$ we see that this inequality is equivalent to

$$\left(r_1 L_1 + \phi r_2 L_2 \bar{w}^{\rho/(1-\rho)} \right) > \bar{w}^{\theta/(1-\rho)} \left(\phi r_1 L_1 \bar{w} + r_2 L_2 \bar{w}^{1/(1-\rho)} \right)$$

where in the last inequality we used $M_i = r_i L_i / f^e$.

Rearranging this expression, we obtain

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

which will finally allow us to prove the result by contradiction.Note that when $L_1 > L_2$ we have $\bar{w} > 1$, so that the term in parentheses on the left-hand-side of this inequality is negative. If $\bar{w}^{\theta/(1-\rho)+1}\phi \geq 1$, then the inequality must be violated and the desired contradiction is shown. Alternatively, if $\bar{w}^{\theta/(1-\rho)+1}\phi < 1$ we can use (44) and the inequality to arrive to an expression that contradicts the initial assertion that $r_1 < r_2$. Thus, since this assertion leads in 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 of 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 is given by (17), which in this case implies

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

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

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

In an interior solution (defined as a situation in which both countries innovate) we must have

$w_1 = w_2$ or else the lower wage country would be the only to innovate. We normalize this wage to one. Using the definition of Ψ_{in}, and given the assumption $A_i \equiv T_i^{1/(1-\rho)}/L_i$ we have $\Psi_{in} = \left[\sum_k A_k L_k \tau_{kn}^{-\theta/(1-\rho)}\right]^{1-\rho}$ and hence

$$\begin{aligned}
\Psi_1^{1/(1-\rho)} &= AL_1 + AL_2 t \\
\Psi_2^{1/(1-\rho)} &= AL_1 t + AL_2
\end{aligned}$$

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

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

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

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

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.5 Real Wage in Terms of Flows.

We prove the following 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,*

$$\frac{w_n}{P_n} = \kappa_n\left(T_n^e T_n^p M_n\right)^{\frac{1}{\theta}}\left(\frac{X_{nnn}}{X_n}\right)^{-\frac{1-\rho}{\theta}}\left(\frac{\sum_l X_{nln}}{X_n}\right)^{-\frac{\rho}{\theta}}\left(\frac{1}{\tilde{\sigma}}\frac{Y_n}{X_n} + \frac{\theta-\sigma+1}{\theta\sigma}\right)^{\frac{\sigma-1-\theta}{\theta(\sigma-1)}} \tag{46}$$

and real expenditure is given by

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

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

Proof. We start with $\lambda_{ln}^T = \sum_i \psi_{iln}\lambda_{in}^E$ and the definitions of ψ_{iln} and ξ_{iln} to show that

$$\lambda_{ln}^T = \sum_k\left(\frac{T_{kl}\xi_{kln}^{-\theta}}{\Psi_{kn}}\right)^{\frac{1}{1-\rho}}\lambda_{in}^E = \left(w_l\tau_{ln}\right)^{-\frac{\theta}{1-\rho}}\sum_k\left(\frac{T_{kl}\gamma_{kl}^{-\theta}}{\Psi_{kn}}\right)^{\frac{1}{1-\rho}}\lambda_{kn}^E$$

so

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

Using the result for the Dixit-Stiglitz price index in (36), using $\zeta \equiv \left(\frac{\tilde{\sigma}^{1-\sigma}\theta}{\theta-\sigma+1}\right)^{1/\theta}\left(\frac{\sigma}{\tilde{\sigma}^{1-\sigma}}\right)^{\frac{\sigma-1-\theta}{\theta(\sigma-1)}}$ and noting that $\lambda_{in}^E = M_i\Psi_{in}/\sum_k M_k\Psi_{kn}$ implies that $\sum_j M_j\Psi_{jn} = M_n\Psi_{nn}/\lambda_{nn}^E$, 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

$$\frac{w_n}{P_n} = \zeta \left(T_n^e T_n^p M_n\right)^{1/\theta} \left(\lambda_{nn}^T\right)^{\frac{\rho-1}{\theta}} \left(\lambda_{nn}^E\right)^{-\frac{1}{\theta}} \left[\sum_k \left(T_k^e \gamma_{kn}^{-\theta} \frac{\Psi_{nn}}{\Psi_{kn}}\right)^{\frac{1}{1-\rho}} \lambda_{kn}^E\right]^{\frac{1-\rho}{\theta}} \left(\frac{w_n F_n}{X_n}\right)^{\frac{\sigma-1-\theta}{\theta(\sigma-1)}}. \quad (47)$$

Using $X_{iln} = \psi_{iln} \lambda_{in}^E X_n$ and $\psi_{iln} \equiv \left(T_{il} \xi_{iln}^{-\theta} / \Psi_{in}\right)^{\frac{1}{1-\rho}}$ and simplifying we get

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

Plugging this into (47), and using the definitions of λ_{nn}^T, λ_{nn}^E, we get that the real wage is given by

$$\frac{w_n}{P_n} = \zeta \left(T_n^e T_n^p M_n\right)^{1/\theta} \left(\frac{X_{nnn}}{X_n}\right)^{-\frac{1-\rho}{\theta}} \left(\frac{\sum_l X_{nln}}{X_n}\right)^{-\frac{\rho}{\theta}} \left(\frac{w_n F_n}{X_n}\right)^{\frac{\sigma-1-\theta}{\theta(\sigma-1)}}. \quad (48)$$

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

$$w_l L_l = \frac{1}{\widetilde{\sigma}} Y_l + \frac{\theta - \sigma + 1}{\theta \sigma} X_l$$

and thus we can write real wage as

$$\frac{w_n}{P_n} = \kappa_n \left(T_n^e T_n^p M_n\right)^{1/\theta} \left(\frac{X_{nnn}}{X_n}\right)^{-\frac{1-\rho}{\theta}} \left(\frac{\sum_l X_{nln}}{X_n}\right)^{-\frac{\rho}{\theta}} \left(\frac{1}{\widetilde{\sigma}} \frac{Y_n}{X_n} + \frac{\theta - \sigma + 1}{\theta \sigma}\right)^{\frac{\sigma-1-\theta}{\theta(\sigma-1)}}. \quad (49)$$

To derive X_n / P_n we simply use (49) and the labor market clearing one more time. This last step completes the proof. **QED** ∎

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

$$\frac{w_n}{P_n} = \frac{X_n}{P_n} \frac{1}{L_n} = \kappa_n \left(T_n^e T_n^p M_n\right)^{1/\theta} \left(\frac{X_{nnn}}{X_n}\right)^{-\frac{1-\rho}{\theta}} \left(\frac{\sum_l X_{nln}}{X_n}\right)^{-\frac{\rho}{\theta}}. \quad (50)$$

Proof. This follows immediately from imposing $X_n = w_n L_n$ in (48). **QED** ∎

B.6 Proof of Proposition 3

Since $T_i^e = 1$ then $m_i = M_i / L_i$ for all i. The assumption that $A_i = 1$ for all i implies that $T_{ii} = T_i^p = L_i^{1-\rho}$. Let $m \equiv \sum M_j / \sum L_j$, let W_i be the real wage in country i under frictionless trade and infinite MP costs and let W_i^* be the real wage in country i under frictionless trade and frictionless MP and define $l_i = L_i / \sum_k L_k$. We first need to characterize the expressions for welfare under restricted entry which is done in the following Lemma.

Lemma 4 *Under restricted entry, consider a world where $T_i^e = \left(T_i^p\right)^{1/(1-\rho)} / L_i = 1$, $\forall i$ and assume $\rho \to 1$ and that $m_i \leq l_i^{\frac{(\theta+1)(\theta\sigma-\sigma+1)}{(\theta-\sigma+1)}}$ for all i. The ratio of the real wage under frictionless trade and infinite MP costs to the real wage under free trade and no MP costs, W_i^* / W_i, is given by the expression:*

$$\frac{W_i^*}{W_i} = \frac{m^{\frac{1}{\theta}} \left(\frac{\sigma-1}{1-\eta} \frac{Y_n}{X_n} + 1\right)^{\frac{\sigma-1-\theta}{\theta(\sigma-1)}}}{\left(\frac{\sigma-\eta\sigma}{1-\eta\sigma}\right)^{\frac{\sigma-1-\theta}{\sigma-1}} \left(m_n\right)^{1/(1+\theta)} \left(\sum_k m_k^{1/(1+\theta)} l_k\right)^{\frac{1}{\theta}}} \quad (51)$$

Proof. See online Appendix ■

With the help of this Lemma we can now proceed to prove the two parts of the proposition:

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 where we also need to solve for X_n/Y_n. We have shown that $(1-\eta)\sum_l X_l = \sum_l L_l \equiv L$ so that using (79) given that $\lambda_i^E = M_i/M$ we obtain

$$X_n = \frac{L_n}{1-\eta}\left(1-\eta+\eta\frac{M_n/L_n}{M/L}\right) \Longrightarrow$$
$$X_n/Y_n = 1-\eta+\eta\frac{m_n}{m} \ . \tag{52}$$

Substituting into (51) and rearranging we get

$$\left(W_n^*/W_n\right)^\theta = \left(\frac{m_n}{m}\right)^{-\theta/(1+\theta)}\left[\frac{\sum m_j l_j}{\left(\sum_k m_k^{1/(1+\theta)} l_k\right)^{1+\theta}}\right]^{\frac{1}{1+\theta}}\left(1-\eta+\eta\frac{m_n}{m}\right)^{\upsilon} \tag{53}$$

where $\upsilon \equiv \theta/(\sigma-1)-1$. Rearranging this expression we get

$$\left(W_n^*/W_n\right)^\theta = \frac{\left[(1-\eta)\,m+\eta m_i\right]^{\upsilon} m^{1-\upsilon}}{m_i^{\theta/(1+\theta)}\sum_k m_k^{1/(1+\theta)} l_k} \tag{54}$$

Taking logs of the expression and differentiating w.r.t. to the size of one country m_i and evaluating at symmetry, $m_j = m$ for all j, we get that the sign of this derivative is determined by

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

or simply

$$\upsilon\eta-\frac{\theta}{1+\theta}.$$

The condition $\upsilon\eta > \frac{\theta}{1+\theta}$ is equivalent to $\sigma > \frac{(1+\theta)^2}{1+\theta+\theta^2} \equiv \bar{\theta}$, which proves part i).

Part ii) Now consider real expenditures. With no MP we have

$$x_i = \frac{\omega_i}{1-\eta},$$

whereas with frictionless trade and MP we have

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

Using (52) we then get

$$\frac{x_i^*}{x_i} = \left(1-\eta+\eta\frac{m_i}{m}\right)\frac{\omega_i^*}{\omega_i}$$

and hence

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

This expression is similar to what we had above for real wages, only that instead of υ we now have $\upsilon+\theta$. Thus, the condition for real income to fall is that $(\upsilon+\theta)\,\eta<\frac{\theta}{1+\theta}$. Notice however that this condition is equivalent to $\sigma-1>\theta$, which can never be true since we require $\theta>\sigma-1$ for the various integrals to have a finite mean. Thus, real expenditure must increase with MP.

What happens to real profits? Profits in country i are $\Pi_i=X_i-w_iL_i$, so real profits per person of country i are $\pi_i=x_i-\omega_i$. With frictionless trade and no MP we have $x_i=\omega_i/(1-\eta)$, so $\pi_i=\omega_i\left(\frac{1}{1-\eta}-1\right)=\frac{\eta\omega_i}{1-\eta}$, while with frictionless trade and frictionless MP we have

$$\pi_i^*=\frac{\eta\omega_i^*}{1-\eta}\frac{m_i}{m}.$$

This implies that

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

and hence

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

Taking logs, differentiating, and evaluating at symmetry we get

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

which is always positive. **QED**

B.7 Proof of Proposition 4

To prove proposition 4 we will first compute the real wage for two scenarios: (i) frictionless trade but no MP and (ii) frictionless trade with frictionless MP. Then we will compare the two.

(i) Frictionless trade but no MP. Given that there is no MP, trade is balanced so that $X_i=Y_i$ and $L_i^e=\eta L_i$ for all i. Therefore the current account balance equation (16) together with (21) and $L_i^e=\eta L_i$ implies $w_iL_i=\sum_n\lambda_{in}^EX_n$. But since there is frictionless trade but no MP, then we also have

$$\lambda_{in}^E=\frac{M_iT_i^eT_i^pw_i^{-\theta}}{\sum_k M_kT_k^eT_k^pw_k^{-\theta}}.$$

The current account balance can then be written as

$$w_iL_i=\frac{M_iT_i^eT_i^pw_i^{-\theta}}{\sum_k M_kT_k^eT_k^pw_k^{-\theta}}\sum_n X_n.$$

Choosing country N labor as the numeraire, and using $M_i=r_iL_i/f_i^e$ with $r_i=\eta$, wages are then

$$w_i=\left(\frac{T_i^eT_i^p/f_i^e}{T_N^eT_N^p/f_N^e}\right)^{\frac{1}{1+\theta}}. \tag{55}$$

Also, note that by using (55) and

$$\Psi_{in} \equiv \left[\sum_k \left(T_{ik} (\gamma_{ik} w_k \tau_{kn})^{-\theta} \right)^{\frac{1}{1-\rho}} \right]^{1-\rho} = T_i^e T_i^p w_i^{-\theta}$$

we obtain

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

Finally, using the above relationship, and $X_n = w_n L_n$ inside the price index, given by equation (36), and equation (55) we finally obtain the real wage

$$\frac{w_i}{P_i} = \zeta \eta^{1/\theta} \left[\left(\frac{F_i}{L_i} \right)^{\frac{\theta-\sigma+1}{1-\sigma}} \sum_k L_k \left(\frac{T_k^e T_k^p}{f_k^e} \right)^{\frac{1}{1+\theta}} \right]^{1/\theta} \left(\frac{T_i^e T_i^p}{f_i^e} \right)^{\frac{1}{1+\theta}}$$

(ii) Frictionless trade and frictionless MP. From (40) and $w_N = 1$ we get

$$w_i = \frac{T_i^e / f_i^e}{T_N^e / f_N^e} \tag{56}$$

The zero profit condition in this case is equation (39) so that

$$\eta \frac{M_i T_i^e}{\sum_k M_k T_k^e} \sum_k w_k L_k = M_i w_i f_i^e,$$

which implies

$$\sum_k M_k T_k^e = \eta \frac{T_i^e}{w_i f_i^e} \sum_k X_k$$

But using $X_n = w_n L_n$ and (56) we then get

$$\sum_k M_k T_k^e = \eta \sum_k \frac{L_k T_k^e}{f_k^e}$$

The above equation, together with equation (37), inside the price index, equation (36), imply that

$$P_n = \left[\zeta^\theta \eta \left(\frac{F_n}{L_n} \right)^{(\theta-\sigma+1)/(1-\sigma)} \left(\frac{T_N^e}{f_N^e} \right)^\theta \left(\sum\nolimits_k \left[T_k^p \left(T_k^e / f_k^e \right)^{-\theta} \right]^{1/(1-\rho)} \right)^{1-\rho} \sum\nolimits_k \frac{L_k T_k^e}{f_k^e} \right]^{-1/\theta}$$

and given equation (56) the real wage in country i is then

$$\frac{w_n}{P_n} = \zeta \eta^{1/\theta} T_n^e / f_n^e \left[\left(\frac{F_n}{L_n} \right)^{(\theta-\sigma+1)/(1-\sigma)} \left(\sum\nolimits_k \left[T_k^p \left(T_k^e / f_k^e \right)^{-\theta} \right]^{1/(1-\rho)} \right)^{1-\rho} \left(\sum\nolimits_k \frac{L_k T_k^e}{f_k^e} \right) \right]^{1/\theta}$$

Comparison of the welfare with without MP and with free MP To prove our result we need to compare the welfare in the two cases, and thus we need to prove

$$\left(\sum_k \frac{L_k T_k^e}{f_k^e} \right) \left(\sum_k \left[T_k^p \left(T_k^e / f_k^e \right)^{-\theta} \right]^{1/(1-\rho)} \right)^{1-\rho} \left(\frac{T_i^e}{f_i^e} \right)^\theta \geq \left[\sum_k L_k \left(\frac{T_k^e T_k^p}{f_k^e} \right)^{\frac{1}{1+\theta}} \right] \left(\frac{T_i^e T_i^p}{f_i^e} \right)^{\frac{\theta}{1+\theta}} \Longrightarrow$$
$$\left(\sum_k \left[T_k^p \left(\frac{T_k^e}{f_k^e} \right)^{-\theta} \right]^{1/(1-\rho)} \right)^{1-\rho} \geq \left[T_i^p \left(\frac{T_i^e}{f_i^e} \right)^{-\theta} \right]^{\frac{\theta}{1+\theta}} \left(\sum_j \frac{L_j \frac{T_j^e}{f_j^e}}{\sum_k \frac{L_k T_k^e}{f_k^e}} \left[T_j^p \left(\frac{T_j^e}{f_j^e} \right)^{-\theta} \right]^{\frac{1}{1+\theta}} \right). \tag{57}$$

Note first that the RHS of this expression is less or equal than $\max_k T_k^p \left(\frac{T_k^e}{f_k^e} \right)^{-\theta}$. Then given that

$$\begin{aligned}
\left(\sum_k \left[T_k^p \left(\frac{T_k^e}{f_k^e}\right)^{-\theta}\right]^{1/(1-\rho)}\right)^{1-\rho} &\geq \max_k T_k^p \left(\frac{T_k^e}{f_k^e}\right)^{-\theta} \implies \\
\sum_k \left[T_k^p \left(\frac{T_k^e}{f_k^e}\right)^{-\theta}\right]^{1/(1-\rho)} &\geq \left[\max_k T_k^p \left(\frac{T_k^e}{f_k^e}\right)^{-\theta}\right]^{1/(1-\rho)}
\end{aligned}$$

which holds always true.

C Additional Appendix (not for the PDF)

C.1 Main Text

AddNoteText1 *** Algebra: Using (12) , (8) we have

$$\begin{aligned}
X_i &= Y_i - \sum_{j,n} M_{jin} w_n F_n + \sum_{j,l} M_{jli} w_i F_i + \sum_{l,n} \Pi_{iln} - \sum_{j,n} \Pi_{jin} \Longrightarrow \\
X_i &= Y_i + \frac{\theta - \sigma + 1}{\sigma\theta} \left[\sum_{j,l} X_{jli} - \sum_{j,n} X_{jin} \right] + \eta \sum_{l,n} X_{iln} - \eta \sum_{j,n} X_{jin} \Longrightarrow \\
X_i &= Y_i + \frac{\theta - \sigma + 1}{\sigma\theta} \left[X_i - Y_i \right] + \eta \sum_{l,n} X_{iln} - \eta Y_i
\end{aligned}$$

AddNoteText1b *** We want to derive $r_i = L_i^e / L_i = M_i f_i^e / L_i$

from the labor market cleraling condition we have

$$\begin{aligned}
\frac{\theta - \sigma + 1}{\sigma\theta} X_i + \frac{1}{\tilde{\tilde{\sigma}}} \sum_n \lambda_{in}^T X_n + w_i M_i f_i^e &= w_i L_{i.} \Longrightarrow \\
\frac{\theta - \sigma + 1}{\sigma\theta} + \frac{1}{\tilde{\tilde{\sigma}}} \frac{Y_i}{X_i} + \frac{M_i f_i^e}{L_{i.}} &= 1 \Longrightarrow \\
r_i &= 1 - \frac{\theta - \sigma + 1}{\sigma\theta} - \frac{1}{\tilde{\tilde{\sigma}}} \frac{Y_i}{X_i} \Longrightarrow \\
r_i - \eta &= \frac{\sigma - 1}{\sigma} - \frac{1}{\tilde{\tilde{\sigma}}} \frac{Y_i}{X_i} \Longrightarrow \\
r_i - \eta &= \frac{1}{\tilde{\tilde{\sigma}}} \left(1 - \frac{Y_i}{X_i} \right) \\
r_i - \eta &= \frac{1}{\tilde{\tilde{\sigma}}} \left(\frac{X_i - Y_i}{X_i} \right)
\end{aligned}$$

AddNoteText1c ***

$$\begin{aligned}
r_i &= M_i = \left[\frac{1}{\tilde{\tilde{\sigma}}} \left(\frac{X_i - Y_i}{X_i} \right) + \eta \right] \frac{L_i}{f_e} \\
\tilde{M}_i &= \eta \frac{L_i}{f_e} \\
\frac{M_i}{\widetilde{M}_i} &= \frac{1}{\eta \tilde{\tilde{\sigma}}} \left(\frac{X_i - Y_i}{X_i} \right) + 1 \\
&= \theta - \theta \frac{Y_i}{X_i} + 1
\end{aligned}$$

C.2 Appendix

AddNoteApp1*** Algebra, to be removed:

$$\begin{aligned}
(1-1/\sigma)\frac{\left(T_i^p w_i^{-\theta}\right)^{1/(1-\rho)}}{\sum_j \left(T_j^p w_j^{-\theta}\right)^{1/(1-\rho)}}W+\eta\frac{T_i^e}{\sum_j M_j T_j^e}WM_i &= \eta\frac{T_i^e L_{i.}/f^e}{\sum_j M_j T_j^e}W\left(1-\frac{\theta-\sigma+1}{\sigma\theta}\right)\\
(1-1/\sigma)\frac{f^e}{\eta}\frac{\left(T_i^p w_i^{-\theta}\right)^{1/(1-\rho)}}{\sum_j \left(T_j^p w_j^{-\theta}\right)^{1/(1-\rho)}}\frac{\sum_j M_j T_j^e}{T_i^e L_i}+\frac{M_i f^e}{L_i} &= \left(1-\frac{\theta-\sigma+1}{\sigma\theta}\right)
\end{aligned}$$

AddNoteApp2*** Algebra, to be removed:

$$\begin{aligned}
r_i &= \eta+1-1/\sigma-(1-1/\sigma)\left(\frac{f^e}{\eta}\frac{\left(T_i^p\left(T_i^e\right)^{-\theta}\right)^{1/(1-\rho)}}{\sum_k\left(T_k^p\left(T_k^e\right)^{-\theta}\right)^{1/(1-\rho)}}\frac{\sum_j M_j T_j^e}{T_i^e L_i}\right)\\
&= \eta+\frac{\sigma-1}{\sigma}\left(1-\frac{f^e}{\eta}\frac{A_i L_i\left(T_i^e\right)^{-\theta/(1-\rho)}}{\sum_k A_k L_k\left(T_k^e\right)^{-\theta/(1-\rho)}}\frac{\sum_j M_j T_j^e}{T_i^e L_i}\right)
\end{aligned}$$

AddNoteApp3*** Algebra, to be removed:

$$\begin{aligned}
\frac{f^e M_i}{L_{i.}} &= \frac{\sigma-1}{\sigma}\left(1-\frac{f_e}{\eta}\frac{A_i/\left(T_i^e\right)^{\theta/(1-\rho)}}{\sum_j A_j L_j/\left(T_j^e\right)^{\theta/(1-\rho)}}\frac{\sum_j M_j T_j^e}{T_i^e}\right)+\eta\\
M_i T_i^e &= \frac{\sigma-1}{\sigma}\left(\frac{L_i T_i^e}{f^e}-\frac{1}{\eta}\frac{A_i L_i/\left(T_i^e\right)^{\theta/(1-\rho)}}{\sum_j A_j L_j/\left(T_j^e\right)^{\theta/(1-\rho)}}\sum_j M_j T_j^e\right)+\frac{\eta L_i T_i^e}{f^e}\\
\sum_i M_i T_i^e &= \sum_i\frac{\sigma-1}{\sigma}\left(\frac{L_i T_i^e}{f^e}-\frac{1}{\eta}\frac{L_i A_i/\left(T_l^e\right)^{\theta/(1-\rho)}}{\sum_j L_j A_j/\left(T_j^e\right)^{\theta/(1-\rho)}}\sum_j M_j T_j^e\right)+\eta\sum_i\frac{L_i T_i^e}{f^e}\\
\sum_i M_i T_i^e &= \left(\frac{\sigma-1}{\sigma}+\eta\right)\left(\sum_i\frac{L_i T_i^e}{f^e}\right)-\frac{\sigma-1}{\sigma}\frac{1}{\eta}\sum_j M_j T_j^e\\
\sum_i M_i T_i^e &= \frac{\frac{\sigma-1}{\sigma}+\eta}{1+\frac{\sigma-1}{\sigma}\frac{1}{\eta}}\left(\sum_i\frac{L_i T_i^e}{f^e}\right)=\eta\sum_i\frac{L_i T_i^e}{f^e}
\end{aligned}$$

AddNoteApp4*** Algebra, to be removed: To derive

$$r_i=\frac{\sigma-1}{\sigma}\left(1-\frac{A_i/\left(T_l^e\right)^{\theta/(1-\rho)+1}}{\sum_j\delta_j A_j/\left(T_j^e\right)^{\theta/(1-\rho)+1}}\right)+\eta.$$

Notice that the equilibrium is true only if it satisfies the condition that $0\le L_i^e\le L_i$ for all i.

$$\begin{aligned}
r_i &= \frac{\sigma-1}{\sigma}\left(1-\frac{f_e}{\eta}\frac{A_i/\left(T_i^e\right)^{\theta/(1-\rho)}}{\sum_j A_j L_j/\left(T_j^e\right)^{\theta/(1-\rho)}}\frac{\eta\frac{\sum_i L_i T_i^e}{f^e}}{T_i^e}\right)+\eta\\
&= \frac{\sigma-1}{\sigma}\left(1-\frac{A_i/\left(T_i^e\right)^{\theta/(1-\rho)+1}}{\sum_j A_j L_j T_j^e/\left(T_j^e\right)^{\theta/(1-\rho)+1}}\sum_i L_i T_i^e\right)+\eta
\end{aligned}$$

AddNoteApp5*** Algebra, to be removed:The following expression

$$1-\frac{(1-\eta)\sigma}{\sigma-1}<\frac{A_i/\left(T_l^e\right)^{\theta/(1-\rho)+1}}{\sum_j\delta_j A_j/\left(T_j^e\right)^{\theta/(1-\rho)+1}}<1+\eta\frac{\sigma}{\sigma-1}$$

comes from

$$
\begin{aligned}
0 &< \frac{\sigma-1}{\sigma}\left(1-\frac{A_i/\left(T_l^e\right)^{\theta/(1-\rho)+1}}{\sum_j \delta_j A_j/\left(T_j^e\right)^{\theta/(1-\rho)+1}}\right)+\eta<1 \\
-\eta\frac{\sigma}{\sigma-1} &< 1-\frac{A_i/\left(T_l^e\right)^{\theta/(1-\rho)+1}}{\sum_j \delta_j A_j/\left(T_j^e\right)^{\theta/(1-\rho)+1}}<(1-\eta)\frac{\sigma}{\sigma-1} \\
1+\eta\frac{\sigma}{\sigma-1} &> \frac{A_i/\left(T_l^e\right)^{\theta/(1-\rho)+1}}{\sum_j \delta_j A_j/\left(T_j^e\right)^{\theta/(1-\rho)+1}}>1-(1-\eta)\frac{\sigma}{\sigma-1}
\end{aligned}
$$

C.3 Proof of Proposition 2

AddNoteApp6*** Algebra, to be removed:

$$
\begin{aligned}
M_1 f^e &= \eta\lambda_1^E(L_1+L_2/\bar{w}) \\
M_2 f^e/\bar{w} &= \eta\lambda_2^E(L_1+L_2/\bar{w}) \\
\frac{M_1}{M_2}\bar{w} &= \frac{\lambda_{11}^E}{\lambda_{22}^E}=\frac{M_1\Psi_1}{M_2\Psi_2}
\end{aligned}
$$

AddNoteApp7*** Algebra, to be removed: using $\Psi_{iv}=\left[\sum_v\left(T_{iv}\xi_{ivn}^{-\theta}\right)^{\frac{1}{1-\rho}}\right]^{1-\rho}$ and the assumptions in the proposition we get

$$
\frac{\Psi_1}{\Psi_2}=\frac{\left[\sum_v\left(T^eT_v^p\left(\gamma_{1v}w_v\right)^{-\theta}\right)^{\frac{1}{1-\rho}}\right]^{1-\rho}}{\left[\sum_v\left(T^eT_v^p\left(\gamma_{2v}w_v\right)^{-\theta}\right)^{\frac{1}{1-\rho}}\right]^{1-\rho}}
$$

AddNoteApp8*** Algebra, to be removed: We have

$$
\bar{w}=\left(\frac{\left(T_1^p\right)^{1/(1-\rho)}+\left(T_2^p\right)^{1/(1-\rho)}\phi\bar{w}^{\theta/(1-\rho)}}{\left(T_1^p\right)^{1/(1-\rho)}\phi+\left(T_2^p\right)^{1/(1-\rho)}\bar{w}^{\theta/(1-\rho)}}\right)^{1-\rho}.
$$

$A_1=A_2$ implies that

$$
\begin{aligned}
\bar{w} &= \left(\frac{L_1+L_2\phi\bar{w}^{\theta/(1-\rho)}}{L_1\phi+L_2\bar{w}^{\theta/(1-\rho)}}\right)^{1-\rho} \\
\bar{w}^{1/(1-\rho)}\left(L_1\phi+L_2\bar{w}^{\theta/(1-\rho)}\right) &= L_1+L_2\phi\bar{w}^{\theta/(1-\rho)} \\
L_2\bar{w}^{\theta/(1-\rho)}\left(\bar{w}^{1/(1-\rho)}-\phi\right) &= L_1\left(1-\phi\bar{w}^{1/(1-\rho)}\right) \\
\frac{\bar{w}^{\theta/(1-\rho)}\left(\bar{w}^{1/(1-\rho)}-\phi\right)}{1-\phi\bar{w}^{1/(1-\rho)}} &= \frac{L_1}{L_2}
\end{aligned}
$$

AddNoteApp9*** Algebra: We have $\lambda_{ln}^T=\sum_i\psi_{iln}\lambda_{in}^E$ with

$$
\psi_{iln}=\psi_{il}\equiv\left(\frac{T_{il}\xi_{iln}^{-\theta}}{\Psi_{in}}\right)^{\frac{1}{1-\rho}}=\frac{\left(T^eT_l^p\left(\gamma_{il}w_l\right)^{-\theta}\right)^{1/(1-\rho)}}{\sum_v\left(T^eT_v^p\left(\gamma_{iv}w_v\right)^{-\theta}\right)^{1/(1-\rho)}}=\frac{T_e^{1/(1-\rho)}L_l\left(\gamma_{il}w_l\right)^{-\theta/(1-\rho)}}{\left(\Psi_i\right)^{1/(1-\rho)}}
$$

Since

$$\begin{aligned}\lambda_1^T &= \psi_{11}\lambda_1^E + \psi_{21}\lambda_2^E \\ \lambda_2^T &= \psi_{12}\lambda_1^E + \psi_{22}\lambda_2^E\end{aligned}$$

then (45) implies

$$\begin{aligned}\frac{\psi_{11}\lambda_1^E + \psi_{21}\lambda_2^E}{w_1 L_1} &> \frac{\psi_{12}\lambda_1^E + \psi_{22}\lambda_2^E}{w_2 L_2} \\ L_1 w_1^{-\theta/(1-\rho)} \frac{\lambda_1^E \Psi_1^{-1/(1-\rho)} + \phi\lambda_2^E \Psi_2^{-1/(1-\rho)}}{w_1 L_1} &> L_2 w_2^{-\theta/(1-\rho)} \frac{\phi\lambda_1^E \Psi_1^{-1/(1-\rho)} + \lambda_2^E \Psi_2^{-1/(1-\rho)}}{w_2 L_2} \\ w_1^{-\theta/(1-\rho)-1}\left(\lambda_1^E \Psi_1^{-1/(1-\rho)} + \phi\lambda_2^E \Psi_2^{-1/(1-\rho)}\right) &> w_2^{-\theta/(1-\rho)-1}\left(\phi\lambda_1^E \Psi_1^{-1/(1-\rho)} + \lambda_2^E \Psi_2^{-1/(1-\rho)}\right) \\ w_1^{-\theta/(1-\rho)-1}\left(M_1 \Psi_1^{-\rho/(1-\rho)} + \phi M_2 \Psi_2^{-\rho/(1-\rho)}\right) &> w_2^{-\theta/(1-\rho)-1}\left(\phi M_1 \Psi_1^{-\rho/(1-\rho)} + M_2 \Psi_2^{-\rho/(1-\rho)}\right)\end{aligned}$$

AddNoteApp9b*** Note that when $L_1 > L_2$ we have $\bar{w} > 1$, so that the term in parentheses on left-hand-side of this inequality is negative. If $\bar{w}^{\theta/(1-\rho)+1}\phi \geq 1$, then the inequality must be violated and the desired contradiction is shown. Alternatively, if $\bar{w}^{\theta/(1-\rho)+1}\phi < 1$ then the above inequality implies that

$$\frac{r_1}{r_2} > \frac{L_2 \bar{w}^{\frac{\rho}{1-\rho}}\left(\bar{w}^{\theta/(1-\rho)+1} - \phi\right)}{L_1\left(1 - \phi\bar{w}^{\theta/(1-\rho)+1}\right)}. \tag{58}$$

But using (44) then we have that

$$\frac{L_2 \bar{w}^{\frac{\rho}{1-\rho}}\left(\bar{w}^{\theta/(1-\rho)+1} - \phi\right)}{L_1\left(1 - \phi\bar{w}^{\theta/(1-\rho)+1}\right)} = \frac{\bar{w}^{\frac{1}{1-\rho}} - \phi\bar{w}^{-\frac{\theta-\rho}{1-\rho}}}{\bar{w}^{1/(1-\rho)} - \phi} \frac{1 - \phi\bar{w}^{1/(1-\rho)}}{1 - \phi\bar{w}^{\theta/(1-\rho)+1}}$$

But given assumption $\theta > 1$, then $\bar{w} > 1$ implies that $\bar{w}^{-\frac{\theta-\rho}{1-\rho}} < 1$ and $\bar{w}^{\theta/(1-\rho)+1} < \bar{w}^{1/(1-\rho)}$, so both fractions on the RHS are higher than one, which implies

$$\frac{L_2 \bar{w}^{\frac{\rho}{1-\rho}}\left(\bar{w}^{\theta/(1-\rho)+1} - \phi\right)}{L_1\left(1 - \phi\bar{w}^{\theta/(1-\rho)+1}\right)} > 1.$$

AddNoteApp10*** Algebra: *** Algebra:

$$\begin{aligned}\frac{L_1^p}{L_1} &= \frac{\theta-\sigma+1}{\theta\sigma} + \frac{\sigma-1}{\sigma} A\left[\frac{L_1}{AL_1 + AL_2 t} + \frac{tL_2}{AL_1 t + AL_2}\right] \\ &= \frac{\theta-\sigma+1}{\theta\sigma} + \frac{\sigma-1}{\sigma}\left[\frac{l_1}{l_1 + (1-l_1)t} + \frac{t(1-l_1)}{l_1 t + (1-l_1)}\right]\end{aligned}$$

Repeating this algebra for country 2 yields

$$\frac{L_2^p}{L_2} = \frac{\theta-\sigma+1}{\theta\sigma} + \frac{\sigma-1}{\sigma}\left[\frac{tl_1}{l_1 + (1-l_1)t} + \frac{(1-l_1)}{l_1 t + (1-l_1)}\right]$$

C.4 Real Wages

AddNoteApp10b*** The labor market clearing condition under restricted entry gives

$$
\begin{aligned}
w_l L_l &= \left(\frac{\sigma-1}{\sigma}\right) Y_l + \frac{\theta-\sigma+1}{\theta\sigma} X_l \implies \\
\frac{w_l L_l}{X_l} &= \left[\left(\frac{\sigma-1}{\sigma}\right)\frac{Y_l}{X_l} + \frac{\theta-\sigma+1}{\theta\sigma}\right] \implies
\end{aligned}
$$

so that

$$
\begin{aligned}
\frac{w_n}{P_n} &= \zeta \left(T_n^e T_n^p M_n\right)^{1/\theta} \left(\frac{X_{nnn}}{X_n}\right)^{-\frac{1-\rho}{\theta}} \left(\frac{\sum_l X_{nln}}{X_n}\right)^{-\frac{\rho}{\theta}} \left(\left[\left(\frac{\sigma-1}{\sigma}\right)\frac{Y_n}{X_n} + \frac{\theta-\sigma+1}{\theta\sigma}\right]\frac{F_n}{L_n}\right)^{\frac{\sigma-1-\theta}{\theta(\sigma-1)}} \\
&= \left(\frac{F_n}{L_n}\right)^{\frac{\sigma-1-\theta}{\theta(\sigma-1)}} \zeta \left(T_n^e T_n^p M_n\right)^{1/\theta} \left(\frac{X_{nnn}}{X_n}\right)^{-\frac{1-\rho}{\theta}} \left(\frac{\sum_l X_{nln}}{X_n}\right)^{-\frac{\rho}{\theta}} \left[\left(\frac{\sigma-1}{\sigma}\right)\frac{Y_n}{X_n} + \frac{\theta-\sigma+1}{\theta\sigma}\right]^{\frac{\sigma-1-\theta}{\theta(\sigma-1)}}
\end{aligned}
$$

and also

$$
\begin{aligned}
\frac{X_n}{w_n L_n}\frac{w_n}{P_n} &= \frac{X_n}{w_n L_n} \zeta \left(T_n^e T_n^p M_n\right)^{1/\theta} \left(\frac{X_{nnn}}{X_n}\right)^{-\frac{1-\rho}{\theta}} \left(\frac{\sum_l X_{nln}}{X_n}\right)^{-\frac{\rho}{\theta}} \left(\frac{w_n F_n}{X_n}\right)^{\frac{\sigma-1-\theta}{\theta(\sigma-1)}} \\
&= \left(\left(\frac{\sigma-1}{\sigma}\right)\frac{Y_n}{X_n} + \frac{\theta-\sigma+1}{\theta\sigma}\right)^{-1} \zeta \left(T_n^e T_n^p M_n\right)^{1/\theta} \left(\frac{X_{nnn}}{X_n}\right)^{-\frac{1-\rho}{\theta}} \left(\frac{\sum_l X_{nln}}{X_n}\right)^{-\frac{\rho}{\theta}} \left(\left[\left(\frac{\sigma-1}{\sigma}\right)\frac{Y_n}{X_n} + \right.\right. \\
&= \left(\frac{F_n}{L_n}\right)^{\frac{\sigma-1-\theta}{\theta(\sigma-1)}} \zeta \left(T_n^e T_n^p M_n\right)^{1/\theta} \left(\frac{X_{nnn}}{X_n}\right)^{-\frac{1-\rho}{\theta}} \left(\frac{\sum_l X_{nln}}{X_n}\right)^{-\frac{\rho}{\theta}} \left[\left(\frac{\sigma-1}{\sigma}\right)\frac{Y_n}{X_n} + \frac{\theta-\sigma+1}{\theta\sigma}\right]^{\frac{\sigma-1-\theta}{\theta(\sigma-1)}-1}
\end{aligned}
$$

AddNoteApp10c***

$$
\begin{aligned}
X_i\left(1-\frac{\theta-\sigma+1}{\sigma\theta}\right) &= \left(\frac{\sigma-1}{\sigma}\right) Y_i + \eta \sum_n \lambda_{in}^E X_n \implies \\
\sigma - 1 + \eta\sigma - \sigma\eta\lambda_i^E \frac{\sum_n X_n}{X_i} &= (\sigma-1)\frac{Y_i}{X_i} \implies \\
\frac{\sigma - 1 + \eta\sigma - \sigma\eta\lambda_i^E \frac{\sum_n X_n}{X_i}}{1-\eta\sigma} &= \frac{\sigma-1}{1-\eta\sigma}\frac{Y_i}{X_i} \\
\frac{\frac{\sigma-\sigma\eta\lambda_i^E\frac{\sum_n X_n}{X_i}}{1-\eta\sigma}}{\frac{\sigma-\eta\sigma}{1-\eta\sigma}} &= \frac{\frac{\sigma-1}{\sigma}\frac{Y_n}{X_n} + \frac{\theta-\sigma+1}{\theta\sigma}}{1-\eta} \\
\frac{1-\eta\lambda_i^E\frac{\sum_n X_n}{X_i}}{1-\eta} &= \frac{\frac{\sigma-1}{\sigma}\frac{Y_n}{X_n} + \frac{\theta-\sigma+1}{\theta\sigma}}{1-\eta}
\end{aligned}
$$

AddNoteApp10d*** Thus, we finally have

$$
\begin{aligned}
W_n^*/W_n &= \frac{m^{\frac{1}{\theta}}\left(\frac{\sigma-1}{1-\eta\sigma}\frac{Y_n}{X_n}+1\right)^{\frac{\sigma-1-\theta}{\theta(\sigma-1)}}}{\left(\frac{\sigma-\eta\sigma}{1-\eta\sigma}\right)^{\frac{\sigma-1-\theta}{\sigma-1}}(m_n)^{1/(1+\theta)}\left(\sum_j m_j^{1/(1+\theta)}l_j\right)^{\frac{1}{\theta}}} \\
&= \frac{m^{\frac{1}{\theta}}\left(\frac{m}{(1-\eta)m+\eta m_i}\right)^{\frac{\sigma-1-\theta}{\theta(\sigma-1)}}}{(m_n)^{1/(1+\theta)}\left(\sum_j m_j^{1/(1+\theta)}l_j\right)^{\frac{1}{\theta}}}
\end{aligned}
$$

$$
(W_n^*/W_n)^\theta = \frac{m\left(\frac{m}{(1-\eta)m+\eta m_i}\right)^{\frac{\sigma-1-\theta}{(\sigma-1)}}}{(m_n)^{\theta/(1+\theta)}\left(\sum_j m_j^{1/(1+\theta)}l_j\right)}
$$

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

$$
(W_n^*/W_n)^\theta = \frac{(m)^{1-\upsilon}\left((1-\eta)m+\eta m_n\right)^{\upsilon}}{(m_n)^{\theta/(1+\theta)}\left(\sum_j m_j^{1/(1+\theta)}l_j\right)}
$$

AddNoteApp11*** Algebra, to be removed:

$$
\begin{aligned}
P_n &= \zeta^{-1}\left[\left(\frac{w_nF_n}{X_n}\right)^{(\theta-\sigma+1)/(1-\sigma)}\sum_i M_i\Psi_{in}\right]^{-1/\theta} \\
&= \zeta^{-1}\left(\left(\frac{w_nF_n}{X_n}\right)^{(\theta-\sigma+1)/(1-\sigma)}\frac{M_n\Psi_{nn}}{\lambda_{nn}^E}\right)^{-1/\theta}
\end{aligned}
$$

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

$$
\frac{w_n}{P_n} = \zeta\left(T_n^pM_n\right)^{1/\theta}\left(\lambda_{nn}^T\right)^{\frac{\rho-1}{\theta}}\left(\lambda_{nn}^E\right)^{-\frac{1}{\theta}}\left[\sum_i\left(T_i^e\gamma_{in}^{-\theta}\frac{\Psi_{nn}}{\Psi_{in}}\right)^{\frac{1}{1-\rho}}\lambda_{in}^E\right]^{\frac{1-\rho}{\theta}}\left(\frac{w_nF_n}{X_n}\right)^{\frac{\sigma-1-\theta}{\theta(\sigma-1)}}. \tag{59}
$$

Notice that with Free entry we have

$$
\frac{X_n}{L_n}\frac{1}{P_n} = \left(\frac{F_n}{L_n}\right)^{\frac{\sigma-1-\theta}{\theta(\sigma-1)}}\zeta\left(T_n^pM_n\right)^{1/\theta}\left(\lambda_{nn}^T\right)^{\frac{\rho-1}{\theta}}\left(\lambda_{nn}^E\right)^{-\frac{1}{\theta}}\left[\sum_i\left(T_i^e\gamma_{in}^{-\theta}\frac{\Psi_{nn}}{\Psi_{in}}\right)^{\frac{1}{1-\rho}}\lambda_{in}^E\right]^{\frac{1-\rho}{\theta}}
$$

AddNoteApp12*** Algebra, to be removed:

$$
\begin{aligned}
\frac{w_n}{P_n} &= \zeta\left[\left(\frac{w_nF_n}{X_n}\right)^{1-\theta/(\sigma-1)}\frac{M_n\Psi_{nn}}{\lambda_{nn}^E}\right]^{1/\theta}\left[\frac{\lambda_{nn}^T}{\sum_i\left(T_{in}\gamma_{in}^{-\theta}/\Psi_{in}\right)^{1/(1-\rho)}\lambda_{in}^E}\right]^{-(1-\rho)/\theta} \\
&= \zeta\left(\lambda_{nn}^T\right)^{-(1-\rho)/\theta}\left(\lambda_{nn}^E\right)^{-1/\theta}\left[M_n\Psi_{nn}\right]^{1/\theta}\left[\sum_i\left(T_i^eT_n^p\gamma_{in}^{-\theta}/\Psi_{in}\right)^{1/(1-\rho)}\lambda_{in}^E\right]^{(1-\rho)/\theta}\left(\frac{w_nF_n}{X_n}\right)^{\frac{\sigma-1-\theta}{\theta(\sigma-1)}} \\
&= \zeta\left(\lambda_{nn}^T\right)^{-(1-\rho)/\theta}\left(\lambda_{nn}^E\right)^{-1/\theta}\left[T_n^pM_n\right]^{1/\theta}\left[(\Psi_{nn})^{1/(1-\rho)}\sum_i\left(T_i^e\gamma_{in}^{-\theta}/\Psi_{in}\right)^{1/(1-\rho)}\lambda_{in}^E\right]^{(1-\rho)/\theta}\left(\frac{w_nF_n}{X_n}\right)^{\frac{\sigma-1-\theta}{\theta(\sigma-1)}}
\end{aligned}
$$

Using $X_{iln} = \psi_{iln}\lambda^E_{in}X_n$ and $\psi_{iln} \equiv \left(T_{il}\xi^{-\theta}_{iln}/\Psi_{in}\right)^{\frac{1}{1-\rho}}$ and simplifying we get

$$\sum_i \left(T^e_i \gamma^{-\theta}_{in} \frac{\Psi_{nn}}{\Psi_{in}}\right)^{\frac{1}{1-\rho}} \lambda^E_{in} = \frac{\left(T^e_n\right)^{\frac{1}{1-\rho}} \lambda^E_{nn}}{X_{nnn}/\sum_i X_{inn}}.$$

AddNoteApp13*** Algebra, to be removed:

$$\begin{aligned}
\frac{X_{nnn}}{\sum_i X_{inn}} &= \frac{\psi_{nnn}\lambda^E_{nn}X_n}{\sum_i \psi_{inn}\lambda^E_{in}X_n} \\
&= \frac{\left(T^e_n T^p_n \xi^{-\theta}_{nnn}/\Psi_{nn}\right)^{\frac{1}{1-\rho}} \lambda^E_{nn}}{\sum_i \left(T^e_i T^p_n \xi^{-\theta}_{inn}/\Psi_{in}\right)^{\frac{1}{1-\rho}} \lambda^E_{in}} = \frac{\left(T^e_n/\Psi_{nn}\right)^{\frac{1}{1-\rho}} \lambda^E_{nn}}{\sum_i \left(T^e_i \gamma^{-\theta}_{in}/\Psi_{in}\right)^{\frac{1}{1-\rho}} \lambda^E_{in}},
\end{aligned}$$

Plugging this into (47), and using the definitions of λ^T_{nn}, λ^E_{nn}, we get that the real wage is given by

$$\frac{w_n}{P_n} = \zeta \left(T^e_n T^p_n M_n\right)^{1/\theta} \left(\frac{X_{nnn}}{X_n}\right)^{-\frac{1-\rho}{\theta}} \left(\frac{\sum_l X_{nln}}{X_n}\right)^{-\frac{\rho}{\theta}} \left(\frac{w_n F_n}{X_n}\right)^{\frac{\sigma-1-\theta}{\theta(\sigma-1)}}. \tag{60}$$

AddNoteApp14*** Algebra, to be removed:

$$\begin{aligned}
\frac{w_n}{P_n} &= \zeta \left(T^p_n M_n\right)^{1/\theta} \left(\lambda^T_{nn}\right)^{\frac{\rho-1}{\theta}} \left(\lambda^E_{nn}\right)^{-\frac{1}{\theta}} \left[\frac{\left(T^e_n\right)^{\frac{1}{1-\rho}} \lambda^E_{nn}}{X_{nnn}/\sum_i X_{inn}}\right]^{\frac{1-\rho}{\theta}} \left(\frac{w_n F_n}{X_n}\right)^{\frac{\sigma-1-\theta}{\theta(\sigma-1)}} \\
&= \zeta \left(T^e_n T^p_n M_n\right)^{1/\theta} \left(\lambda^T_{nn}\right)^{\frac{\rho-1}{\theta}} \left(\lambda^E_{nn}\right)^{-\frac{\rho}{\theta}} \left[\frac{\sum_i X_{inn}}{X_{nnn}}\right]^{\frac{1-\rho}{\theta}} \left(\frac{w_n F_n}{X_n}\right)^{\frac{\sigma-1-\theta}{\theta(\sigma-1)}} \\
&= \zeta \left(T^e_n T^p_n M_n\right)^{1/\theta} \left(\frac{\sum_i X_{inn}}{X_n}\right)^{-\frac{1-\rho}{\theta}} \left(\frac{\sum_l X_{nln}}{X_n}\right)^{-\frac{\rho}{\theta}} \left(\frac{X_{nnn}}{\sum_i X_{inn}}\right)^{-\frac{1-\rho}{\theta}} \left(\frac{w_n F_n}{X_n}\right)^{\frac{\sigma-1-\theta}{\theta(\sigma-1)}}
\end{aligned}$$

AddNoteApp15*** Algebra, to be removed: We start with

$$\frac{w_n}{P_n} = \zeta \left(\frac{F_n}{L_n}\right)^{\frac{\sigma-1-\theta}{\theta(\sigma-1)}} \left(T^e_n T^p_n M_n\right)^{1/\theta} \left(\frac{X_{nnn}}{X_n}\right)^{-\frac{1-\rho}{\theta}} \left(\frac{\sum_l X_{nln}}{X_n}\right)^{-\frac{\rho}{\theta}} \left(\frac{\frac{\sigma-1}{\sigma}Y_n + \frac{\theta-\sigma+1}{\theta\sigma}X_n}{X_n}\right)^{\frac{\sigma-1-\theta}{\theta(\sigma-1)}}$$

and note that

$$\frac{\frac{\sigma-1}{\sigma}Y_n + \frac{\theta-\sigma+1}{\theta\sigma}X_n}{X_n} = \frac{\theta-\sigma+1}{\theta\sigma}\left(\frac{\sigma-1}{\sigma}\frac{\theta\sigma}{\theta-\sigma+1}\frac{Y_n}{X_n} + 1\right)$$

But $1 - \frac{\theta-\sigma+1}{\sigma\theta} = \eta + 1 - 1/\sigma$, so $\frac{\theta-\sigma+1}{\sigma\theta} = 1/\sigma - \eta$, and hence

$$\frac{w_n}{P_n} = \kappa^R_n \left(T^e_n T^p_n M_n\right)^{1/\theta} \left(\frac{X_{nnn}}{X_n}\right)^{-\frac{1-\rho}{\theta}} \left(\frac{\sum_l X_{nln}}{X_n}\right)^{-\frac{\rho}{\theta}} \left(\frac{\sigma-1}{1-\eta\sigma}\frac{Y_n}{X_n} + 1\right)^{\frac{\sigma-1-\theta}{\theta(\sigma-1)}}$$

To derive X_n/P_n we simply use (49) and the labor market clearing one more time.

AddNoteApp16******* Algebra:

$$\begin{aligned}
\frac{X_n}{P_n}\frac{1}{L_n} &= \frac{X_n}{w_n L_n}\frac{w_n}{P_n} = \kappa_n \left(T^e_n T^p_n M_n\right)^{1/\theta} \left(\frac{X_{nnn}}{X_n}\right)^{-\frac{1-\rho}{\theta}} \left(\frac{\sum_l X_{nln}}{X_n}\right)^{-\frac{\rho}{\theta}} \left(\frac{\sigma-1}{1-\eta\sigma}\frac{Y_n}{X_n} + 1\right)^{\frac{\sigma-1-\theta}{\theta(\sigma-1)}} \frac{X_n}{\left(\frac{\sigma-1}{\sigma}\right) Y_n + \frac{\theta-\sigma}{\theta\sigma}} \\
&= \kappa_n \left(T^e_n T^p_n M_n\right)^{1/\theta} \left(\frac{X_{nnn}}{X_n}\right)^{-\frac{1-\rho}{\theta}} \left(\frac{\sum_l X_{nln}}{X_n}\right)^{-\frac{\rho}{\theta}} \left(\frac{\sigma-1}{1-\eta\sigma}\frac{Y_n}{X_n} + 1\right)^{\frac{\sigma-1-\theta}{\theta(\sigma-1)}} \frac{1}{\frac{\theta-\sigma+1}{\theta\sigma}\left(\frac{\sigma-1}{1-\eta\sigma}\frac{Y_n}{X_n} + 1\right)} \\
&= \frac{\kappa_n}{\frac{\theta-\sigma+1}{\theta\sigma}} \left(T^e_n T^p_n M_n\right)^{1/\theta} \left(\frac{X_{nnn}}{X_n}\right)^{-\frac{1-\rho}{\theta}} \left(\frac{\sum_l X_{nln}}{X_n}\right)^{-\frac{\rho}{\theta}} \left(\frac{\sigma-1}{1-\eta\sigma}\frac{Y_n}{X_n} + 1\right)^{\frac{\sigma-1-\theta}{\theta(\sigma-1)}-1}
\end{aligned}$$

**** Costas: this is not the same as what we have above in the lemma... did I make a mistake* *or do we have a typo? ****

Finally, we obtain

$$\frac{X_n}{P_n}\frac{1}{L_n} = \kappa_n \left(T_{nn}M_n\right)^{1/\theta}\left(\frac{X_{nnn}}{X_n}\right)^{-\frac{1-\rho}{\theta}}\left(\frac{\sum_l X_{nln}}{X_n}\right)^{-\frac{\rho}{\theta}}\left[\boldsymbol{\chi}\left(\frac{\sigma-1}{\sigma}\frac{Y_n}{X_n}+\frac{\theta-\sigma+1}{\theta\sigma}\right)^{\frac{\sigma-1-\theta}{\theta(\sigma-1)}-1}+1-\boldsymbol{\chi}\right]. \tag{61}$$

where $\boldsymbol{\chi}$ is an indicator

AddNoteApp18**********The only questionable point of the proof is whether we can take (or maybe under which conditions) the limit $\rho \to 1$. Below in between the *** there is a serious effort to show that but not yet completed, the rest of the proof works as below***

AddNoteApp19********** Let's construct an equilibrium for the limit $\rho \to 1$. Notice that

$$\begin{aligned} \frac{\lambda_l^T}{\lambda_i^T} &= \frac{\left(T_l^p\right)^{1/(1-\rho)} w_l^{-\theta/(1-\rho)}}{\left(T_i^p\right)^{1/(1-\rho)} w_i^{-\theta/(1-\rho)}} \to \\ \left(\frac{\lambda_l^T}{\lambda_i^T}\right)^{1-\rho} &= \frac{T_l^p w_l^{-\theta}}{T_i^p w_i^{-\theta}} \end{aligned}$$

AddNoteApp20****for $\rho \to 1$ we have that the expression (75) becomes

$$W_n^* = \kappa_n \left(\sum M_i\right)^{\frac{1}{\theta}}\left(\frac{\sigma-1}{1-\eta\sigma}\frac{Y_n}{X_n}+1\right)^{\frac{\sigma-1-\theta}{\theta(\sigma-1)}}$$

Using $M \equiv \sum M_i$ we finally get

$$W_n^* = \kappa_n M^{\frac{1}{\theta}}\left(\frac{\sigma-1}{1-\eta\sigma}\frac{Y_n}{X_n}+1\right)^{\frac{\sigma-1-\theta}{\theta(\sigma-1)}} \tag{62}$$

C.5 Proof of Proposition 4

AddNoteApp21b*******we derive real wage as

$$P_n = \left(\zeta^\theta \eta \left(\frac{F_n}{L_n}\right)^{\frac{\theta-\sigma+1}{1-\sigma}}\left(T_N^e T_N^p/f_N^e\right)^{\frac{\theta}{1+\theta}}\sum_i L_i\left(\frac{T_i^e T_i^p}{f_i^e}\right)^{\frac{1}{1+\theta}}\right)^{-1/\theta}$$

and using 55)

$$\frac{w_i}{P_i} = \zeta\eta^{1/\theta}\left(\left(\frac{F_i}{L_i}\right)^{\frac{\theta-\sigma+1}{1-\sigma}}\sum_j L_j\left(\frac{T_j^e T_j^p}{f_j^e}\right)^{\frac{1}{1+\theta}}\right)^{1/\theta}\left(\frac{T_i^e T_i^p}{f_i^e}\right)^{\frac{1}{1+\theta}}$$

AddNoteApp22******* Algebra:

$$\begin{aligned} w_i f_i^e &= \eta\frac{T_i^e}{\sum_v M_v T_v^e}\sum_n w_n L_n \\ \sum_v M_v T_v^e &= \eta\frac{T_i^e}{\frac{T_i^e/f_i^e}{T_N^e/f_N^e}f_i^e}\sum_n\frac{T_n^e/f_n^e}{T_N^e/f_N^e}L_n = \eta\sum_n\frac{L_n T_n^e}{f_n^e} \end{aligned}$$

AddNoteApp23*** Algebra:

$$
\begin{aligned}
P_n &= \left[\zeta^{\theta}\left(\frac{w_n F_n}{X_n}\right)^{(\theta-\sigma+1)/(1-\sigma)} \sum_i M_i \Psi_{in}\right]^{-1/\theta} \\
&= \left[\zeta^{\theta}\left(\frac{F_n}{L_n}\right)^{(\theta-\sigma+1)/(1-\sigma)} \left(\sum_v \left(T_v^p w_v^{-\theta}\right)^{1/(1-\rho)}\right)^{1-\rho} \sum_i M_i T_i^e\right]^{-1/\theta} \\
&= \left[\zeta^{\theta}\left(\frac{F_n}{L_n}\right)^{(\theta-\sigma+1)/(1-\sigma)} \left(\sum_v \left[T_v^p \left(\frac{T_v^e/f_v^e}{T_N^e/f_N^e}\right)^{-\theta}\right]^{1/(1-\rho)}\right)^{1-\rho} \eta \sum_j \frac{L_j T_j^e}{f_j^e}\right]^{-1/\theta} \\
&= \left[\zeta^{\theta}\eta\left(\frac{F_n}{L_n}\right)^{(\theta-\sigma+1)/(1-\sigma)} \left(\frac{T_N^e}{f_N^e}\right)^{\theta} \left(\sum_v \left[T_v^p \left(T_v^e/f_v^e\right)^{-\theta}\right]^{1/(1-\rho)}\right)^{1-\rho} \sum_j \frac{L_j T_j^e}{f_j^e}\right]^{-1/\theta}
\end{aligned}
$$

D Online Appendix

Lemma 4 Under restricted entry, consider a world where $T_i^e = (T_i^p)^{1/(1-\rho)}/L_i = 1$, $\forall i$ and where $m_i \leq l_i \frac{(\theta+1)(\theta\sigma-\sigma+1)}{(\theta-\sigma+1)}$ $\forall i$ and assume $\rho \to 1$ for all i. The ratio of the real wage under frictionless trade and infinite MP costs to the real wage under free trade and no MP costs, W_i^*/W_i, is given by the expression:

$$\frac{W_i^*}{W_i} = \frac{m^{\frac{1}{\theta}}\left(\frac{m}{(1-\eta)m+\eta m_i}\right)^{\frac{\sigma-1-\theta}{\theta(\sigma-1)}}}{(m_i)^{1/(1+\theta)}\left(\sum_j m_j^{1/(1+\theta)} l_j\right)^{\frac{1}{\theta}}} \tag{63}$$

Proof. Since we focus on frictionless trade, imposing $T_i^e = 1$ we have

$$\Psi_{in} = \left[\sum_k \left(T_k^p(\gamma_{ik}w_k)^{-\theta}\right)^{\frac{1}{1-\rho}}\right]^{1-\rho} \equiv \Psi_i, \tag{64}$$

$$\psi_{iln} = \left[T_l^p\left(\gamma_{il}w_l\right)^{-\theta}/\Psi_i\right]^{\frac{1}{1-\rho}} \equiv \psi_{il}, \tag{65}$$

and

$$\lambda_{in}^E = \frac{M_i\Psi_i}{\sum M_j\Psi_j} \equiv \lambda_i^E.$$

■

Using the definition of λ_{ln}^T and imposing free trade we have

$$w_n = \left[\frac{\lambda_{nn}^T}{\sum_i\left(T_n^p\gamma_{in}^{-\theta}/\Psi_{in}\right)^{1/(1-\rho)}\lambda_{in}^E}\right]^{-(1-\rho)/\theta}.$$

and given (64) and $\lambda_{in}^E = \lambda_i^E$, we can write this expression as

$$w_n = \left[\frac{\lambda_n^T}{(T_n^p)^{1/(1-\rho)}\sum_i\left(\gamma_{in}^{-\theta}/\Psi_i\right)^{1/(1-\rho)}\lambda_i^E}\right]^{-(1-\rho)/\theta} \tag{66}$$

we will use this relationship below and consider separately the two cases: infinite MP costs and zero MP costs.

Infinite MP costs: In this case, we additionally have $\lambda_n^T = \lambda_n^E$, and $\Psi_n = T_n^p w_n^{-\theta}$ so that expression (66) yields

$$w_n = \left[\frac{\lambda_n^T}{(T_n^p/\Psi_n)^{1/(1-\rho)}\lambda_n^E}\right]^{-(1-\rho)/\theta},$$

whereas equation (23) with $\tau_{in} = 1$ for all i, n and infinite MP costs becomes

$$\lambda_l^E = \lambda_l^T = \frac{M_l T_l^p w_l^{-\theta}}{\sum_j M_j T_j^p w_j^{-\theta}}, \tag{67}$$

which gives us

$$\left(\frac{w_i}{w_l}\right)^{-\theta} = \frac{\lambda_i^T}{M_i T_i^p} / \frac{\lambda_l^T}{M_l T_l^p}. \tag{68}$$

Exogenous entry implies $L_l^p = L_l$, and no MP implies $X_l = Y_l$, hence (13) implies

$$w_l L_l = (1-\eta)\,\lambda_l^T \sum_n X_n,$$

and hence

$$\frac{w_l}{w_i} = \frac{\lambda_l^T}{L_l} / \frac{\lambda_i^T}{L_i}.$$

Combined with (68) this equation yields

$$\frac{w_l}{w_i} = \left(\frac{L_i}{L_l}\frac{M_l T_l^p}{M_i T_i^p}\right)^{1/(1+\theta)} \tag{69}$$

Using $T_i^p = L_i^{1-\rho}$ and setting $w_n = 1$ by choice of numeraire, we then have

$$w_i = \left(M_n / L_n^\rho\right)^{-1/(1+\theta)} \left(M_i / L_i^\rho\right)^{1/(1+\theta)}. \tag{70}$$

Also, notice that expression (49) gives the real wage under exogenous entry. With no MP we have $X_{nln} = 0$ except for $l = n$, $\lambda_{nn}^T = X_{nnn}/X_n$, and $X_n = Y_n$, so that the expression for welfare in this case is

$$W_n = \kappa_n \left(1-\eta\right)^{\frac{\sigma-1-\theta}{\theta(\sigma-1)}} \left(T_n^p M_n\right)^{\frac{1}{\theta}} \left(\lambda_{nn}^T\right)^{-\frac{1}{\theta}}$$

But using expression (67), welfare in the case of no MP can be written as

$$W_n = \kappa_n \left(1-\eta\right)^{\frac{\sigma-1-\theta}{\theta(\sigma-1)}} \left(T_n^p M_n\right)^{\frac{1}{\theta}} \left(\frac{M_n T_n^p w_n^{-\theta}}{\sum_k M_k T_k^p w_k^{-\theta}}\right)^{-\frac{1}{\theta}}$$

Now substituting for expression (70) and $T_i^p = L_i^{1-\rho}$, performing some simplifications and considering the limit $\rho \to 1$ we obtain

$$W_n = \kappa_n \left(1-\eta\right)^{\frac{\sigma-1-\theta}{\theta(\sigma-1)}} \left(M_n / L_n\right)^{1/(1+\theta)} \left(\sum\nolimits_k M_k^{1/(1+\theta)} L_k^{\theta/(1+\theta)}\right)^{\frac{1}{\theta}} \tag{71}$$

Zero MP costs: Using (9) and (10) together with the definition of ψ_{iln} and imposing zero MP costs we get

$$\lambda_l^T = \sum_k \left[\frac{T_l^p w_l^{-\theta}}{\Psi_k}\right]^{\frac{1}{1-\rho}} \frac{M_k \Psi_k}{\sum_j M_j \Psi_j}.$$

But now we have $\Psi_i = \Psi \equiv \left[\sum_k \left(T_k^p w_k^{-\theta}\right)^{\frac{1}{1-\rho}}\right]^{1-\rho}$, hence

$$\lambda_i^E = \frac{M_i}{\sum_k M_k} \tag{72}$$

and

$$\lambda_l^T = \frac{\left(T_l^p w_l^{-\theta}\right)^{\frac{1}{1-\rho}}}{\sum_k \left(T_k^p w_k^{-\theta}\right)^{\frac{1}{1-\rho}}}. \tag{73}$$

Therefore, relative trade shares can be given by

$$\frac{\lambda_l^T}{\lambda_i^T} = \frac{\left(T_l^p\right)^{1/(1-\rho)} w_l^{-\theta/(1-\rho)}}{\left(T_i^p\right)^{1/(1-\rho)} w_i^{-\theta/(1-\rho)}}. \tag{74}$$

Using (8) and noting that $\frac{\sum_l X_{nln}}{X_n} = \lambda_{nn}^E$ (from the definition of λ_{in}^E) and recalling that $\lambda_{nn}^E = \lambda_n^E$, then equation (49) can be rewritten as

$$W_n^* = \kappa_n \left(T_n^p M_n\right)^{\frac{1}{\theta}} \left(\lambda_n^E\right)^{-\frac{1}{\theta}} \left(\psi_{nnn}\right)^{-\frac{1-\rho}{\theta}} \left(\frac{1}{\widetilde{\sigma}}\frac{Y_n}{X_n} + \frac{\theta-\sigma+1}{\theta\sigma}\right)^{\frac{\sigma-1-\theta}{\theta(\sigma-1)}},$$

But $\psi_{iln} = \frac{\left(T_l^p (w_l)^{-\theta}\right)^{\frac{1}{1-\rho}}}{\sum_k \left(T_k^p w_k^{-\theta}\right)^{\frac{1}{1-\rho}}} \equiv \psi_n$, $\lambda_n^E = M_n / \sum M_i$ and $T_l^p = L_l^{1-\rho}$ hence this implies

$$W_n^* = \kappa_n \left(L_n^{1-\rho}\right)^{1/\theta} \left(\sum M_i\right)^{\frac{1}{\theta}} \left(\psi_n\right)^{-\frac{1-\rho}{\theta}} \left(\frac{1}{\widetilde{\sigma}}\frac{Y_n}{X_n} + \frac{\theta-\sigma+1}{\theta\sigma}\right)^{\frac{\sigma-1-\theta}{\theta(\sigma-1)}} \tag{75}$$

We want to find the expression for W_n^* when $\rho \to 1$. We first conjecture that under this limit wages equalize and we *a)* derive an expression for the last parenthetical term of the welfare expression *b)* show that ψ_n is finite and bounded away from zero *c)* show that the wage equalization conjecture is true. Combining these three results the limit of the expression (75) as $\rho \to 1$ is

$$W_n^* = \kappa_n M^{\frac{1}{\theta}} \left(1-\eta\right)^{\frac{\sigma-1-\theta}{\theta(\sigma-1)}} \left[\frac{m}{\left(1-\eta\right)m + \eta m_i}\right]^{\frac{\sigma-1-\theta}{\theta(\sigma-1)}} \tag{76}$$

a) We first postulate that wages equalize as $\rho \to 1$ and show that

$$\frac{1}{\widetilde{\sigma}}\frac{Y_n}{X_n} + \frac{\theta-\sigma+1}{\theta\sigma} = \left(1-\eta\right)\frac{m}{\left(1-\eta\right)m + \eta m_i}.$$

From the current account balance condition (16) combined with the labor market clearing condition (19) we obtain

$$X_i\left(1 - \frac{\theta-\sigma+1}{\sigma\theta}\right) = \frac{1}{\widetilde{\sigma}}Y_i + \eta\sum_n \lambda_{in}^E X_n$$

and given that $\lambda_{in}^E = \lambda_i^E$,

$$\frac{X_i}{\sum_k X_k}\left(1 - \frac{\theta-\sigma+1}{\sigma\theta}\right) = \frac{1}{\widetilde{\sigma}}\frac{Y_i}{\sum_k X_k} + \eta\lambda_i^E\ . \tag{77}$$

But labor market clearing (19) implies (using $\sum_l \lambda_{ln}^T X_n = Y_l$)

$$-\frac{1}{\widetilde{\sigma}}\frac{Y_l}{\sum_k X_k} = \frac{\theta-\sigma+1}{\theta\sigma}\frac{X_l}{\sum_k X_k} - \frac{w_l L_l}{\sum_k X_k}\ . \tag{78}$$

Note that $\sum_k Y_k = \sum_k X_k$ combined with the labor market clearing condition implies $\sum_k w_k L_k = (1-\eta)\sum_k Y_k$. Combining (77) and (78) and using this result together with $\lambda_i^E = M_i / \sum_k M_k$ yields

$$\frac{X_i}{\sum_k X_k}\left(1 - \frac{\theta-\sigma+1}{\sigma\theta}\right) = \frac{1}{\widetilde{\sigma}}\frac{\frac{\theta-\sigma+1}{\theta\sigma}\frac{X_i}{\sum_k X_k} - \frac{w_i L_i}{\sum_k X_k}}{-\frac{1}{\widetilde{\sigma}}} + \eta\lambda_i^E \implies$$

$$\frac{X_i}{\sum_k X_k} = \frac{w_i L_i}{\sum_k X_k} + \eta\lambda_i^E \implies \frac{\sum_k X_k}{X_i} = \frac{1}{\left(1-\eta\right)\frac{w_i L_i}{\sum_k w_k L_k} + \eta\frac{M_i}{\sum_k M_k}} \tag{79}$$

But (77) with $\lambda_i^E = M_i / \sum_j M$ implies

$$1 - \eta\frac{M_i}{\sum_k M_k}\frac{\sum_n X_n}{X_i} = \frac{1}{\widetilde{\sigma}}\frac{Y_i}{X_i} + \frac{\theta-\sigma+1}{\sigma\theta}\ .$$

Using (79) we then get

$$\begin{aligned}
\frac{1}{\widetilde{\sigma}}\frac{Y_i}{X_i}+\frac{\theta-\sigma+1}{\sigma\theta} &= 1-\eta\frac{M_i}{\sum_j M}\frac{1}{(1-\eta)\frac{w_iL_i}{\sum_k w_kL_k}+\eta\frac{M_i}{\sum_k M_k}} \\
&= 1-\frac{\eta}{(1-\eta)\frac{w_iL_i}{\sum_k w_kL_k}\frac{\sum_k M_k}{M_i}+\eta}
\end{aligned}$$

If wages are equalized, then the RHS becomes

$$\begin{aligned}
1-\frac{\eta}{(1-\eta)\frac{w_iL_i}{\sum_k w_kL_k}\frac{\sum_k M_k}{M_i}+\eta} &= 1-\frac{\eta}{(1-\eta)\,m/m_i+\eta} \\
&= 1-\frac{\eta m_i}{(1-\eta)\,m+\eta m_i} \\
&= (1-\eta)\frac{m}{(1-\eta)\,m+\eta m_i}
\end{aligned}$$

b) Now we want to show that under the condition in the proposition in this limit equilibrium all countries have , $\psi_n > 0$. To show that we want to find the limit

$$\begin{aligned}
\lim\psi_n &= \lim_{\rho\to1}\frac{\left(T_l^p\,(w_l)^{-\theta}\right)^{\frac{1}{1-\rho}}}{\sum_k\left(T_k^p w_k^{-\theta}\right)^{\frac{1}{1-\rho}}} \\
&= \lim_{\rho\to1}\frac{1}{\sum_k\left(\frac{T_k^p w_k^{-\theta}}{T_l^p(w_l)^{-\theta}}\right)^{\frac{1}{1-\rho}}} \\
&= \lim\lambda_l^T
\end{aligned}$$

Thus, we simply need to construct the trade shares in the case of wage equalization with $\rho\to1$. The equilibrium conditions in a frictionless equilibrium are current account balance,

$$X_i = w_iL_i+\eta\frac{M_i}{M}\sum_n X_n,$$

and labor market clearing,

$$\frac{\theta-\sigma+1}{\sigma\theta}X_i+(1-1/\sigma)\,\lambda_i^T\sum_n X_n = w_iL_i$$

with

$$\lambda_l^T = \frac{\left(T_l^p w_l^{-\theta}\right)^{\frac{1}{1-\rho}}}{\sum_v\left(T_v^p w_v^{-\theta}\right)^{\frac{1}{1-\rho}}} = \frac{L_iw_i^{-\theta/(1-\rho)}}{\sum_v L_vw_v^{-\theta/(1-\rho)}}$$

Adding up across the current account balance conditions implies

$$\sum\nolimits_k X_k = \frac{1}{1-\eta}\sum\nolimits_k w_kL_k$$

Combining current account balance with labor market clearing and using this last result together witht he expression for λ_i^T implies

$$\frac{\theta-\sigma+1}{\sigma\theta}\left(w_iL_i+\frac{M_i}{M}\frac{\eta}{1-\eta}\sum\nolimits_k w_kL_k\right)+(1-1/\sigma)\frac{L_iw_i^{-\theta/(1-\rho)}}{\sum_k L_kw_k^{-\theta/(1-\rho)}}\frac{1}{1-\eta}\sum\nolimits_k w_kL_k = w_iL_i$$

We can set the numeraire by imposing $\sum w_j L_j = L$ so then we have

$$\frac{\eta}{1-\eta}\frac{\theta-\sigma+1}{\sigma\theta}\frac{M_i}{M}+\frac{1}{\widetilde{\sigma}}\frac{1}{1-\eta}\frac{l_i w_i^{-\theta/(1-\rho)}}{\sum_k l_k w_k^{-\theta/(1-\rho)}}=\left(1-\frac{\theta-\sigma+1}{\sigma\theta}\right)w_i l_i \tag{80}$$

Together with $\sum w_j L_j = L$ this is a system of N non-linear equations in N unknowns. If wages are equalized in the limit as $\rho \to 1$, then we must have

$$\lambda_i^T = \lim_{\rho\to 1}\frac{l_i w_i^{-\theta/(1-\rho)}}{\sum_k l_k w_k^{-\theta/(1-\rho)}} = \widetilde{\sigma}\left(1-\eta\right)\left(1-\frac{\theta-\sigma+1}{\sigma\theta}\right)l_i-\widetilde{\sigma}\frac{\theta-\sigma+1}{\sigma\theta}\eta\frac{m_i}{m}$$

so we simply need to assume that

$$0<\widetilde{\sigma}\left(1-\eta\right)\left(1-\frac{\theta-\sigma+1}{\sigma\theta}\right)l_i-\widetilde{\sigma}\frac{\theta-\sigma+1}{\sigma\theta}\eta\frac{m_i}{m}<1 \tag{81}$$

which is the condition of the proposition

c) In the last step we want to show that in the limit equilibrium wages are equalized. The system can be rewritten as

$$a_i/l_i+\left(\frac{w_i}{\left(\sum_k l_k w_k^{-\theta/(1-\rho)}\right)^{-(1-\rho)/\theta}}\right)^{-\theta/(1-\rho)}=bw_i$$

where

$$a_i \equiv \widetilde{\sigma}\frac{\theta-\sigma+1}{\sigma\theta}\eta\frac{m_i/l_i}{m} \text{ and } b\equiv\widetilde{\sigma}\left(1-\eta\right)\left(1-\frac{\theta-\sigma+1}{\sigma\theta}\right)$$

Assumption (81) is then

$$0\le b-a_i/l_i\le 1$$

Since $\max w_v \ge 1$ then, letting $j=\arg\max_v w_v$, we then have $b\max w_v - a_j/l_j \ge 0$, then (80) implies

$$\left(\frac{\max w_v}{\left(\sum_k l_k w_k^{-\theta/(1-\rho)}\right)^{-(1-\rho)/\theta}}\right)=\left(b\max w_v-a_j/l_j\right)^{-(1-\rho)/\theta} \tag{82}$$

Note that

$$\lim_{\rho\to 1}\left(\sum\nolimits_k l_k w_k^{-\theta/(1-\rho)}\right)^{-(1-\rho)/\theta}=\min w_k$$

This implies that

$$\lim_{\rho\to 1}\left(\frac{\max w_v}{\left(\sum_k l_k w_k^{-\theta/(1-\rho)}\right)^{-(1-\rho)/\theta}}\right)=\lim_{\rho\to 1}\frac{\max w_v}{\min w_v}$$

But

$$\lim_{\rho\to 1}\left(b\max w_v-a_j/l_j\right)^{-(1-\rho)/\theta}=1$$

hence, taking limits of (82) we have

$$\lim_{\rho\to 1}\frac{\max w_v}{\min w_v}=1$$

We have completed the derivations of the two formulas for no MP and frictionless MP. From the two formulas (71) and (76) we obtain

$$W_n^*/W_n = \frac{(m)^{1-\upsilon}\left(\left(1-\eta\right)m+\eta m_n\right)^{\upsilon}}{\left(m_n\right)^{\theta/(1+\theta)}\left(\sum_j m_j^{1/(1+\theta)} l_j\right)} \tag{83}$$

with $\upsilon \equiv \theta/\left(\sigma-1\right)-1$. Notice that under symmetry $W_n^*/W_n = 1$. Also notice that the above expression is the same as (54). This last derivation completes the Lemma.