The Global Diffusion of Ideas*

Francisco J. Buera
Federal Reserve Bank of Chicago

Ezra Oberfield
Princeton

April 1, 2015

Preliminary and Incomplete

Abstract

We provide a tractable theory of innovation and diffusion of technologies to explore the role of international trade and foreign direct investment (FDI). We model innovation and diffusion as a process involving the combination of new ideas with insights from other industries or countries. We provide conditions under which each country's equilibrium frontier of knowledge converges to a Frechet distribution, and derive a system of differential equations describing the evolution of the scale parameters of these distributions, i.e., countries' stocks of knowledge. In particular, the growth of a country's stock of knowledge depends only on the its trade and FDI shares and the stocks of knowledge of its trading partners. We use this framework to quantify the dynamic gains from trade in the short and long run. We also explore the model's potential to account for the post-war growth of South Korea.

*

Economic miracles are characterized by protracted growth in productivity, per-capita income, and increases in trade and FDI flows. The experiences of Japan and South Korea in the postwar period and the recent performance of China are prominent examples. These experiences suggest an important role played by openness in the process of development.[1] Yet quantitative trade models relying on standard static mechanisms imply relatively small gains from openness, and therefore cannot account for growth miracles.[2] These findings call for alternative channels through which openness can affect development. In this paper we present and analyze a model of an alternative mechanism: the impact of openness on the creation and diffusion of best practices across countries.

We model innovation and diffusion as a process involving the combination of new ideas with insights from other industries and countries. Insights occur randomly due to local interactions among domestic producers. In our theory openness affects the creation and diffusion of ideas by determining the distribution from which domestic producers draw their insights. Our theory is flexible enough to incorporate different channels through which ideas may diffuse across countries. We focus on two main channels: (i) insights are drawn from those that sell goods to a country, (ii) insights are drawn from technologies used domestically, whether foreign or domestically owned. In our model, openness to trade and Foreign Direct Investment (FDI) affects the quality of the insights drawn by domestic producers by selecting different sellers to a country and/or affecting the technologies used to produce domestically.

We use the model to explore several questions. First, we study how barriers to trade and FDI alter the learning process. At the micro level, the insights one draws depend on local interactions. At the aggregate level, the growth of a country's stock of knowledge depends on its trade and FDI shares and the stocks of knowledge of its trading partners, and that of foreign firms operating domestically. Starting from autarky, opening up to trade and FDI results in a higher temporary growth rate, and permanent higher level, of the stock of knowledge, as the producers are exposed to more productive ideas. Nevertheless, the case of costless trade does not necessarily result in the highest short-run growth rate of the stock of knowledge. For example, a country can change the growth rate of its stock of knowledge by altering the composition of its trading partners. If

[1]See Feyrer (2009a,b) for recent estimates of the impact of trade on income, and a review of the empirical literature. See also the discussion in Lucas (2009b).

[2]See Connolly and Yi (2009) for a quantification of the role of trade on Korean's growth miracle. Atkeson and Burstein (2010) also find relatively small effects in a model with innovation.

learning from sellers is important, a country could increase the growth rate of its stock of knowledge by tilting imports toward its higher-wage trade partners; imports from higher-wage countries tend to be produced at higher productivity, as the high wage must be overcome with a low unit labor requirement. In this case, the growth rate of the stock of knowledge is maximized when import costs perfectly offset the wage differences among its trading partners. However, this generically conflicts with maximizing the static gains from trade.

We next use our model to quantify the dynamic gains from openness, studying in particular how opening to trade and FDI shape the diffusion of ideas. In a world that is generally open, if a single closed country opens to trade, it will experience an instantaneous jump in real income, a mechanism that has been well-studied in the trade literature. Following that jump, this country's stock of knowledge will gradually improve as the liberalization leads to an improvement in the composition of insights drawn by its domestic producers. Here, the speed of convergence depends on the nature of learning process. If insights are drawn from goods that are sold to the country, then convergence will be faster, as opening to trade allows producers to draw insight from the relatively productive foreign producers. In contrast, if insights are drawn from technologies that are used locally, the country's stock of knowledge grows more slowly. In that case, a trade liberalization leads to better selection of the domestic producers, but those domestic producers have low productivity relative to foreign firms. Opening to FDI also tends to lead to faster convergence than opening to trade. Opening to FDI provides more immediate access to the ideas of foreign producers. Over time, the exposure to more productive foreign multinationals producing domestically leads to faster growth of the stock of knowledge. In the case of opening to FDI, the stock of knowledge grows faster when learning is from technologies used locally.

We also study whether trade and FDI are substitutes or complements in the diffusion of ideas. For both the static and dynamic gains from opening, this depends crucially on the correlation of multinationals' productivities across potential production locations. When this correlation is high, trade and FDI are substitutes in increasing the speed of learning and raising real incomes.

Finally, we specify a quantitative version of the model that includes non-traded goods and intermediate inputs, and equipped labor with capital and education. Using cross-country data on trade flows and GDP, we study whether the model can account for the post-war growth of South Korea. We first vary a parameter that modulates the strength of diffusion. We find that the role of diffusion is largest for intermediate values of this parameter; When the

Literature Review Our work builds on a large literature modeling innovation and diffusion of technologies as a stochastic process, starting from the earlier work of Jovanovic and Rob (1989), Jovanovic and MacDonald (1994), Kortum (1997), and recent contributions by Alvarez et al. (2008) and Luttmer (2012).[3] We are particularly related to recent applications of these frameworks to study the connection between trade and the diffusion of ideas (Lucas, 2009a; Alvarez et al., 2013; Perla et al., 2013; Sampson, 2014).

Our theory captures the models in Kortum (1997) and Alvarez et al. (2008, 2013) as special cases. When the contribution of insights to the development of new technologies is zero, $\beta = 0$ in our notation, our framework simplifies to a version of Kortum (1997) with exogenous search intensity. As in that paper, ours is a model with semi-endogenous growth. When insights from domestic sellers are the only input to the development of new technologies, $\beta = 1$, our framework simplifies to the model in Alvarez et al. (2008, 2013) with stochastic arrival of ideas. Beyond analyzing the intermediate cases, $\beta \in (0, 1)$, the behavior of the model is qualitatively different from either of the two special cases $\beta = 0$ or $\beta = 1$. With $\beta = 0$, there is no diffusion of ideas and thus no dynamic gains from trade. With $\beta = 1$, changes in trade costs alter a country's growth rate, and the equilibrium frontier of knowledge is closer to a logistic distribution. More importantly, when $\beta = 1$ and trade barriers are finite, changes in trade barriers have no impact on the tail of the distribution of productivity, and therefore, the model has a more limited success in providing a quantitative theory of the level and transitional dynamics of productivity. With $\beta < 1$, the frontier of knowledge converges to a Frechet distribution. This allows us use the machinery of Eaton and Kortum (2002), Bernard et al. (2003), and Alvarez and Lucas (2007) which have been remarkably successful as quantitative trade models. We therefore believe that studying the intermediate case of $\beta \in (0, 1)$ is a step toward a quantitative model of the cross-country diffusion of ideas. Finally, we are able to nest alternative sources of insights, e.g., learning from those who sell goods to a country, learning from those that produce within a country, and study the role of both trade and FDI in determining the distribution of insights.

Eaton and Kortum (1999) also build a model of the diffusion of ideas across countries in which

[3] Lucas and Moll (2014) and Perla and Tonetti (2014) extends these models by studying the case with endogenous search effort, a dimension that we abstract from.

the distribution of productivities in each country is Frechet, and where the evolution of the scale parameter of the Frechet distribution in each country is governed by a system of differential equations. In their work insights are drawn from the distribution of potential producers in each country, according to exogenous diffusion rates which are estimated to be country-pair specific, although countries are assumed to be in autarky otherwise. Therefore, changes in trade and FDI costs do not affect the diffusion of ideas.

The model shares some features with Oberfield (2013) which models the formation of supply chains and the economy's input-output architecture. In that model, entrepreneurs discover methods of producing their goods using other entrepreneurs' goods as inputs.[4]

1 Technology Diffusion with a General Source Distribution

We begin with a description of technology diffusion in a single country given a general source distribution. The source distribution describes the set of insights that domestic producers might access. In the specific examples that we explore later in the paper, the source distribution will depend on the profiles of productivity across all countries in the world, but in this section we take it to be a general function satisfying weak tail properties. Given the assumption on the source distribution, we show that the equilibrium distribution of productivity in a given economy is Frechet, and derive a differential equation describing the evolution of the scale parameter of this distribution.

We consider an economy with a continuum of goods $s \in [0, 1]$. For each good, there are m producers. We will later study an environment in which the producers engage in Bertrand competition, so that (barring ties) at most one of these producers will actively produce. A producer is characterized by her productivity, q. A producer of good s with productivity q has access to a labor-only, linear technology

$$y(s) = ql(s), \tag{1}$$

where $l(s)$ is the labor input and $y(s)$ is output of good s. The state of technology in the economy is described by the function $M_t(q)$, the fraction of producers with knowledge no greater than q. We

[4]Here, the evolution of the distribution of marginal costs depends on a differential equation summarizing the history of insights that were drawn. In Oberfield (2013), the distribution of marginal costs is the solution to a fixed point problem, as each producer's marginal cost depends on her potential suppliers' marginal costs.

call M_t the distribution of knowledge at t.

The economy's productivity depends on the frontier of knowledge. The frontier of knowledge is characterized by the function

$$\tilde{F}_t(q) \equiv M_t(q)^m.$$

$\tilde{F}_t(q)$ is the probability that none of the m producers of a good have productivity better than q.

We now turn to a description of the dynamics of the distribution of knowledge. We model diffusion as a process involving the random interaction among producers of different goods or countries. We assume each producer draws insights from others stochastically at rate α_t. However there is randomness in the adaptation of that insight. More formally, when an insight arrives to a producer with productivity q, the producer learns an idea with random productivity zq'^{β} and adopts the idea if $zq'^{\beta} > q$. The productivity of the idea has two components. There is an insight drawn from another producer, q', which is drawn from the source distribution $\tilde{G}_t(q')$. The second component z is drawn from an exogenous distribution with CDF $H(z)$. We refer to $H(z)$ as the exogenous distribution of ideas.[5]

This process captures the fact that interactions with more productive individuals tend to lead to more useful insights, but it also allows for randomness in the adaptation of others' techniques to alternative uses. The latter is captured by the random variable z. An alternative interpretation of the model is that z represents an innovator's "original" random idea, which is combined with random insights obtained from other technologies.[6]

Given the distribution of knowledge at time t, $M_t(q)$, the source distribution, $\tilde{G}_t(q')$, and the exogenous distribution of ideas, $H(z)$, the distribution of knowledge at time $t + \Delta$ is

$$M_{t+\Delta}(q) = M_t(q) \left[(1 - \alpha_t \Delta) + \alpha_t \Delta \int_0^{\infty} H\left(q/x^{\beta}\right) d\tilde{G}_t(x) \right]$$

The first term on the right hand side is the distribution of knowledge at time t, which gives the fraction of producers with productivity less than q. The second term is the probability that a

[5] From the perspective of this section, both $\tilde{G}_t(q)$ and $H(z)$ are exogenous. The distinction between these distributions will become clear once we consider specific examples of source distributions, in which the source distribution will be an *endogenous* function of the countries' frontiers of knowledge.

[6] If $\beta = 0$ our framework simplifies to a version of the model in Kortum (1997) with exogenous search intensity. The framework also nests the model of diffusion in Alvarez et al. (2008) with stochastic arrival of ideas if $\beta = 1$, H is degenerate, and $\tilde{G}_t = \tilde{F}_t$.

producer did not have an insight between time t and $t + \Delta$ that raised her productivity above q. This can happen if no insight arrived in an interval of time Δ, an event with probability $1 - \alpha_t \Delta$, or if at least one insight arrived but none resulted in a technique with productivity greater than q, an event that occurs with probability $\int_0^\infty H\left(q/x^\beta\right) d\tilde{G}_t(x)$.

Rearranging and taking the limit as $\Delta \to 0$ we obtain

$$\frac{d}{dt} \ln M_t(q) = \lim_{\Delta \to 0} \frac{M_{t+\Delta}(q) - M_t(q)}{\Delta M_t(q)} = -\alpha_t \int_0^\infty \left[1 - H\left(q/x^\beta\right)\right] d\tilde{G}_t(x).$$

With this, we can derive an equation describing the frontier of knowledge. Since $\tilde{F}_t(q) = K_t(q)^m$, the change in the frontier of knowledge evolves as:

$$\frac{d}{dt} \ln \tilde{F}_t(q) = -m\alpha_t \int_0^\infty \left[1 - H\left(q/x^\beta\right)\right] d\tilde{G}_t(x).$$

To gain tractability, we make the following assumption about the exogenous component of ideas.

Assumption 1 *The exogenous distribution has a Pareto right tail with exponent θ, so that*

$$\lim_{z \to \infty} \frac{1 - H(z)}{z^{-\theta}} = 1.$$

We thus assume that the right tail of the exogenous distribution of ideas is regularly varying. The restriction that the limit is equal to 1 rather than some other positive number is without loss of generality; we can always choose units so that the limit is one.

For this section we make one additional assumption: the source distribution $\tilde{G}_t$ has a sufficiently thin tail.[7]

Assumption 2 *At each t, $\lim_{q \to \infty} q^{\beta\theta}[1 - \tilde{G}_t(q)] = 0$.*

We will study economies where the number of producers for each good is large. As such, it will be convenient to study how the frontier of knowledge evolves when normalized by the number of producers for each good. Define $F_t(q) \equiv \tilde{F}_t\left(m^{\frac{1}{(1-\beta)\theta}} q\right)$ and $G_t(q) \equiv \tilde{G}_t\left(m^{\frac{1}{(1-\beta)\theta}} q\right)$

[7] In later sections when we endogenize the source distribution, this assumption will be replaced by an analogous assumption on the right tail of the initial distribution of knowledge, $\lim_{q \to \infty} q^{\beta\theta}[1 - M_0(q)] = 0$ and the restriction that $\beta < 1$.

Proposition 1 *Suppose that Assumption 1 and Assumption 2 hold. Then in the limit as $m \to \infty$, the frontier of knowledge evolves as:*

$$\frac{d \ln F_t(q)}{dt} = -\alpha_t q^{-\theta} \int_0^\infty x^{\beta\theta} dG_t(x)$$

Motivated by the previous proposition, we define $\lambda_t \equiv \int_{-\infty}^t \alpha_\tau \int_0^\infty x^{\beta\theta} dG_\tau(x) d\tau$. With this, one can show that the economy's frontier of knowledge converges asymptotically to a Frechet distribution.

Corollary 2 *Suppose that* $\lim_{t\to\infty} \lambda_t = \infty$. *Then* $\lim_{t\to\infty} F_t(\lambda_t^{1/\theta} q) = e^{-q^{-\theta}}$.

Proof. Solving the differential equation gives $F_t(q) = F_0(q) e^{-(\lambda_t - \lambda_0) q^{-\theta}}$. Evaluating this at $\lambda_t^{1/\theta} q$ gives $F_t(\lambda_t^{1/\theta} q) = F_0(\lambda_t^{1/\theta} q) e^{-(\lambda_t - \lambda_0)\lambda_t^{-1} q^{-\theta}}$. This implies that, asymptotically, $\lim_{t\to\infty} F_t(\lambda_t^{1/\theta} q) = e^{-q^{-\theta}}$ ■

Thus, the distribution of productivities in this economy is asymptotically Frechet and the dynamics of the scale parameter is governed by the differential equation

$$\dot{\lambda}_t = \alpha_t \int_0^\infty x^{\beta\theta} dG_t(x). \tag{2}$$

We call λ_t the stock of knowledge.

In the rest of the paper we analyze alternative models for the source distribution G_t. A simple example that illustrates basic features of more general cases is $G_t(q) = F_t(q)$. This corresponds to the case in which diffusion opportunities are randomly drawn from the set of domestic best practices across all goods. In a closed economy this set equals the set of domestic producers and sellers. In this case equation (2) becomes

$$\dot{\lambda}_t = \alpha_t \Gamma(1-\beta) \lambda_t^\beta$$

where $\Gamma(u) = \int_0^\infty x^{u-1} e^{-x} dx$ is the Gamma function. Growth in the long-run is obtained in this framework by assuming that the arrival rate of insight is growing over time, $\alpha_t = \alpha_0 e^{\gamma t}$. In this case, the scale of the Frechet distribution λ_t grows asymptotically at the rate $\gamma/(1-\beta)$, and per-capita GDP grows at the rate $\gamma/[(1-\beta)\theta]$. In general, the evolution of the de-trended stock of knowledge

$\hat{\lambda}_t = \lambda_t e^{\gamma/(1-\beta)t}$ can be summarized in terms of the de-trended arrival of ideas $\hat{\alpha}_t = \alpha_t e^{\gamma t}$

$$\dot{\hat{\lambda}}_t = \hat{\alpha}_t \Gamma(1-\beta)\hat{\lambda}_t^{\beta} - \frac{\gamma}{1-\beta}\hat{\lambda}_t,$$

and on a balanced growth path on which $\hat{\alpha}$ is constant, the de-trended stock of ideas is

$$\hat{\lambda} = \left[\frac{\hat{\alpha}(1-\beta)}{\gamma}\Gamma(1-\beta)\right]^{\frac{1}{1-\beta}}.$$

In the model that follows, potential producers engage in Bertrand competition. In that environment, an important object is the joint distribution of the productivities of best and second best producers of a good. We denote the CDF of this joint distribution as $\tilde{F}_t^{12}(q_1, q_2)$, which, for $q_1 \geq q_2$, equals[8]

$$\tilde{F}_t^{12}(q_1, q_2) = M(q_2)^m + m\left[M(q_2) - M(q_1)\right] M(q_2)^{m-1}.$$

Since the frontier of knowledge at t satisfies $F_t(q) = M_t(q)^m$, the joint distribution can be written as

$$\tilde{F}_t^{12}(q_1, q_2) = \left[1 + m\left\{\left(\tilde{F}_t(q_1)/F_t(q_2)\right)^{1/m} - 1\right\}\right]\tilde{F}(q_2), \qquad q_1 \geq q_2.$$

Normalizing this joint distribution by the number of producers, $F_t^{12}(q_1, q_2) \equiv \tilde{F}_t^{12}\left(m^{\frac{1}{(1-\beta)\theta}} q_1, m^{\frac{1}{(1-\beta)\theta}} q_2\right)$, we have that for large m,

$$F_t^{12}(q_1, q_2) = \left[1 + \log F_t(q_1) - \log F_t(q_2)\right] F(q_2), \qquad q_1 \geq q_2.$$

2 International Trade

We first consider a world where n economies interact through trade, and ideas diffuse through the contact of domestic producers with those who sell goods to the country as well as with those that produce within the country. Given the results from the previous section, the static trade theory is given by the standard Ricardian model in Eaton and Kortum (2002), Bernard et al. (2003), and Alvarez and Lucas (2007), which we briefly introduce before deriving the equations which

[8]Intuitively, there are two ways the best and second best productivities can be no greater than q_1 and q_2 respectively. Either none of the productivities are greater than q_2, or one of the m draws is between q_1 and q_2 and none of the remaining $m-1$ are greater than q_2.

characterize the evolution of the profile of the distribution of productivities of countries in the world economy.

In each country, consumers have identical preferences over a continuum of goods. We use $c_i(s)$ to denote the consumption of a representative household in i of good $s \in [0,1]$. Utility is given by $u(C_i)$, where the the consumption aggregate is

$$C_i = \left[\int_0^1 c_i(s)^{\frac{\varepsilon-1}{\varepsilon}} ds \right]^{\varepsilon/(\varepsilon-1)}$$

so goods enter symmetrically and exchangeably. We assume that $\varepsilon - 1 < \theta$, which guarantees the price level is finite. Let $p_i(s)$ be the price of good s in i, so that i's ideal price index is $P_i = \left[\int_0^1 p_i(s)^{1-\varepsilon} ds \right]^{\frac{1}{1-\varepsilon}}$. Letting X_i denote i's total expenditure, i's consumption of good s is $c_i(s) = \frac{p_i(s)^{-\varepsilon}}{P_i^{1-\varepsilon}} X_i$.

In each country, individual goods can be manufactured by many producers, each using a labor-only, linear technology (1). As discussed in the previous section, provided countries share the same exogenous distribution of ideas $H(z)$, the frontier of productivity in each country is described by a Frechet distribution with curvature θ and a country-specific scale λ_i, $F_i(q) = e^{-\lambda_i q^{-\theta}}$. Transportation costs are given by the standard "iceberg" assumption, where κ_{ij} denotes the units that must be shipped from country j to deliver a unit of the good in country i, with $\kappa_{ii} = 1$ and $\kappa_{ij} \geq 1$.

We now briefly present the basic equations that summarize the static trade equilibrium given the vector of scale parameters $\lambda = (\lambda_1, ..., \lambda_n)$. Because the expressions for price indices, trade shares, and profit are identical to Bernard et al. (2003), we relegate the derivation of these expressions to Appendix B.

Given the isoelastic demand, if a producer had no direct competitors, it would set a price with a markup of $\frac{\varepsilon}{\varepsilon-1}$ over marginal cost. Producers engage in Bertrand competition. This means that lowest cost provider of a good to a country will either use this markup or, if necessary, set a limit price to just undercut the next-lowest-cost provider of the good.

Let w_i denote the wage in country i. For a producer with productivity q in country j, the cost of providing one unit of the good in country i is $\frac{w_j \kappa_{ij}}{q}$. The price of good s in country i is determined as follows. Suppose that country j's best and second best producers of good s have productivities

$q_{j1}(s)$ and $q_{j2}(s)$. The country that can provide good s to i at the lowest cost is given by

$$\arg\min_j \frac{w_j \kappa_{ij}}{q_{j1}(s)}$$

If the lowest-cost-provider of good s for i is a producer from country k, the price of good s in i is

$$p_i(s) = \min\left\{\frac{\varepsilon}{\varepsilon-1}\frac{w_k \kappa_{ik}}{q_{k1}(s)}, \frac{w_k \kappa_{ik}}{q_{k2}(s)}, \min_{j\neq k}\frac{w_j \kappa_{ij}}{q_{j1}(s)}\right\}$$

That is, the price is either the monopolist's price or else it equals the cost of the next-lowest-cost provider of the good; the latter is either the second best producer of good s in country k or the best producer in one of the other countries.

In Appendix B, we show that, in equilibrium, i's price index is

$$P_i = B\left\{\sum_j \lambda_j (w_j \kappa_{ij})^{-\theta}\right\}^{-1/\theta}$$

where B is a constant.[9]

Let $S_{ij} \subseteq [0,1]$ be the set of goods for which a producer in j is the lowest-cost-provider for country i. Let π_{ij} denote the share of country i's expenditure that is spent on goods from country j so that $\pi_{ij} = \int_{s\in S_{ij}} (p_i(s)/P_i)^{1-\varepsilon} ds$. In Appendix B, we show that the expenditure share is

$$\pi_{ij} = \frac{\lambda_j \left(w_j \kappa_{ij}\right)^{-\theta}}{\sum_{k=1}^n \lambda_k \left(w_k \kappa_{ik}\right)^{-\theta}}.$$

A static equilibrium is given by a profile of wages $\mathbf{w} = (w_1, ..., w_n)$ such that labor market clears in all countries. The static equilibrium will depend on whether trade is balanced and where profit from producers is spent. For now, we take the each country's expenditure as given and solve for the equilibrium as a function of these expenditures.

Labor in j is used to produce goods for all destinations. To deliver one unit of good $s \in S_{ij}$ to i, the producer in j uses $\kappa_{ij}/q_{j1}(s)$ units of labor. Thus the labor market clearing constraint for

[9] $B^{1-\varepsilon} = \left[\left(1-\frac{\varepsilon-1}{\theta}\right)\left(1-\left(\frac{\varepsilon}{\varepsilon-1}\right)^{-\theta}\right)+\left(\frac{\varepsilon}{\varepsilon-1}\right)^{-\theta}\right]\Gamma\left(1-\frac{\varepsilon-1}{\theta}\right).$

country j is

$$L_j = \sum_i \int_{s \in S_{ij}} \frac{\kappa_{ij}}{q_{j1}(s)} c_i(s) ds.$$

Similarly, the total profit earned by producers in j can be written as

$$\Pi_j = \sum_i \int_{s \in S_{ij}} \left(p_i(s) - \frac{w_j \kappa_{ij}}{q_{j1}(s)} \right) c_i(s) ds.$$

In Appendix B, we show that these can be expressed as

$$w_j L_j = \frac{\theta}{\theta + 1} \sum_i \pi_{ij} X_i$$

and

$$\Pi_j = \frac{1}{\theta + 1} \sum_i \pi_{ij} X_i$$

Under the natural assumption that trade is balanced and that all profit from domestic producers is spent domestically, then $X_i = w_i L_i + \Pi_i$ and the labor market clearing conditions can be expressed as

$$w_j L_j = \sum_i \pi_{ij} w_i L_i$$

As a simple benchmark, it is useful to consider the case with costless trade, $\kappa_{ij} = 1$, all j, and countries of equal size $L_i = L_j$, all $j \neq i$. In this case, relative wages are

$$\frac{w_i^{FT}}{w_{i'}^{FT}} = \left(\frac{\lambda_i}{\lambda_{i'}} \right)^{\frac{1}{1+\theta}}$$

and the relative expenditure shares are

$$\frac{\pi_{ij}^{FT}}{\pi_{ij'}^{FT}} = \left(\frac{\lambda_j}{\lambda_{j'}} \right)^{\frac{1}{1+\theta}}. \tag{3}$$

Given the static equilibria, we next solve for the evolution of the profile of scale parameters $\lambda = (\lambda_1, ..., \lambda_n)$ by specializing (2) for alternative assumptions about source distributions. We consider source distributions that encompass two cases: (i) domestic producers learn from sellers to the country, (ii) domestic producers learn from other producers in the country.

2.1 Learning from Sellers

Following the framework introduced in Section 1, we model the evolution of technologies as the outcome of a process where technology managers combine "own ideas" with random insights from technologies in other sectors or countries. We first consider the case in which insights are drawn from sellers to the country. In particular, we assume that insights are randomly drawn from the distribution of sellers' productivity in proportion to the expenditure on each good.[10] In this case, the source distribution is given by the expenditure weighted distribution of productivity of sellers

$$G_i(q) = G_i^S(q) \equiv \sum_j \int_{s \in S_{ij} | q_j(s) < q} \frac{p_i(s) c_i(s)}{P_i C_i} ds$$

As we show in Appendix B specializing equation (2) to this source distribution, the evolution of the scale of the Frechet distribution, i.e., the stock of ideas, is described by

$$\begin{aligned} \dot{\lambda}_{it} =& \alpha_{it} \int_0^{\infty} x^{\beta\theta} dG_i^S(q) \\ =& \alpha_{it} B^S \sum_j \pi_{ij} \left(\frac{\lambda_j}{\pi_{ij}} \right)^{\beta} \end{aligned} \tag{4}$$

where B^S is a constant that involves β, θ and ε. That is, the evolution of the stock of ideas is a simple weighted sum of the stock of ideas in all countries, where the weights are given by expenditure shares.

For the change in i's stock of knowledge to be finite, it must be that $\beta + \frac{\varepsilon - 1}{\theta} < 1$. When ε is high, consumers substitute toward the lower cost–or higher productivity–goods. If this condition were violated, then consumers would substitute enough toward these goods that the average insight would be unbounded. We therefore assume for the remainder of the paper that this condition holds.

[10] For the case of learning from sellers, the assumption that insights are randomly drawn in proportion to the expenditure on a good is not central. Alternative assumptions, e.g., insights are uniformly drawn from the set of sellers, give the same law of motion for the each country's stock of knowledge up to a constant. See Appendix C.1.1.

Assumption 3 $\beta + \frac{\varepsilon-1}{\theta} < 1$

Equation (4) shows that trade shapes how a country learns in two ways. Trade gives a country access to the ideas of sellers from other countries. In addition, trade leads to tougher competition, so that there is more selection among the producers from which insights are drawn. In fact, the less a country is able to sell, the stronger selection is among its producers. The amount i learns from j is given by $(\lambda_j/\pi_{ij})^{\beta}$, where λ_j/π_{ij} is the average productivity of sellers from j to i. Holding fixed j's stock of knowledge, a smaller π_{ij} reflects more selection into selling, which means that the insights drawn from sellers from j are likely to be higher quality insights.

Nevertheless, the quality of insights is not necessarily maximized in the case of free trade. To optimize the quality of insights a country must bias its trade toward those countries with higher technologies. In particular, in the short run the growth of country i's stock of knowledge is maximized when its expenditure shares are proportional to the stock of ideas of its trading partners.[11]

$$\frac{\pi_{ij}}{\pi_{ij'}} = \frac{\lambda_j}{\lambda_{j'}}. \tag{5}$$

whereas in equilibrium, country i's expenditure shares will satisfy

$$\frac{\pi_{ij}}{\pi_{ij'}} = \frac{\lambda_j (w_j \kappa_{ij})^{-\theta}}{\lambda_{j'} (w_{j'} \kappa_{ij'})^{-\theta}}. \tag{6}$$

Notice that (5) and (6) coincide only if differences in trade costs perfectly offset differences in trading partners' wages. For example, suppose trade costs are symmetric. A country that is the technological leader can implement this by increasing the cost of trading with less technologically developed countries. In turn, countries that are not at the technological frontier need to subsidize trade with technological advanced countries to maximize the growth rate of the stock of ideas.[12]

As discussed before, to obtain growth in the long-run we assume that the arrival rate of insights grow over time, in which case it is convenient to analyze the evolution of the de-trended stock of

[11] This is the solution to $\max_{\{\pi_{ij}\}} \sum_j \pi_{ij}^{1-\beta} \lambda_j^{\beta}$ subject to $\sum_j \pi_{ij} = 1$.

[12] To be clear, iceberg trade costs are not tariffs (which both distort trade costs and provide revenue), so the preceding argument does not show that the distorting trade represents optimal policy. However, if the shadow value of a higher stock of knowledge is positive, a planner that maximizes the present value of a small open economy's real income and can set country-specific tariffs would generically set tariffs that are non-zero and not uniform across countries.

ideas $\hat{\lambda}_{it} = \lambda_{it} e^{-\frac{\gamma}{1-\beta}t}$

$$\dot{\hat{\lambda}}_{it} = \hat{\alpha}_{it} B^S \sum_{j=1}^{n} \pi_{ij}^{1-\beta} \hat{\lambda}_{jt}^{\beta} - \frac{\gamma}{1-\beta} \hat{\lambda}_{it}, \tag{7}$$

On a balanced growth path where the arrival rate of insights grows at rate γ, the de-trended stock of knowledge solves the system of non-linear equations

$$\hat{\lambda}_i = \frac{(1-\beta)\hat{\alpha}_i}{\gamma} B^S \sum_{j=1}^{n} \pi_{ij}^{1-\beta} \hat{\lambda}_j^{\beta}. \tag{8}$$

2.2 Learning from Production

Another natural source of ideas is given by the interaction of technology managers with other domestic producers, or workers employed by these producers. Along these lines, in this section we consider the case in which the insights are drawn from the distribution of productivity among domestic producers, in proportion to the labor used in the production of each good.[13] Under this assumption, the source distribution is given by the labor weighted distribution of productivity of domestic producers.

If a producer in i is the lowest cost supplier of good s to country j ($s \in S_{ji}$), the fraction of i's labor used to produce the good is $\frac{1}{L_i} \frac{\kappa_{ji}}{q_{i1}(s)} c_j(s)$. Summing over all destinations, i's source distribution would then be

$$G_i(q) = G_i^P(q) = \sum_j \int_{s \in S_{ji} | q_{i1}(s) \leq q} \frac{1}{L_i} \frac{\kappa_{ji}}{q_{i1}(s)} c_j(s) ds$$

As we show in Appendix B specializing equation (2) to this source distribution, the evolution of the scale of the Frechet distribution, i.e., the stock of ideas, is described by

$$\begin{aligned} \dot{\lambda}_{it} &= \alpha_{it} \int_0^{\infty} x^{\beta\theta} dG_i^P(q) \\ &= \alpha_{it} B^P \sum_{j=1}^{n} r_{ji} \left(\frac{\lambda_i}{\pi_{ji}} \right)^{\beta}. \end{aligned}$$

[13] For this case, the assumption that insights are randomly drawn in proportion to the labor used or the value added produced, instead of been randomly drawn from the set of producers, is more important. See Appendix C.1.1 for a characterization of the dynamics of the stock of ideas under alternative assumptions.

where $r_{ji} = \frac{\pi_{ji} X_j}{\sum_k \pi_{ki} X_k}$ is the share of i revenue coming from sales to country j and B^P is a constant involving β, θ and ε. Thus, the source distribution of a country i is a function of the fraction of domestic goods purchased by other countries, π_{ji}, the expenditure (and hence the income) of these countries, and the domestic stock of knowledge, λ_i.

How does trade alter a country's stock of knowledge? In autarky, insights are drawn from all domestic producers, including very unproductive ones. As a country opens up to trade the set of domestic producers improves as the unproductive technologies are selected out. This raises the quality of insights drawn and increases the growth rate of the stock of knowledge.[14]

For each of the two specifications of learning, trade induces selection, which alters the composition of insights drawn. However, there are two important differences across the two specifications. First, with learning from sellers, a country gets insights from its trading partners' stocks of knowledge, whereas with learning from producers, insights are drawn from a country's own stock of knowledge. Second, with learning from sellers, the composition of expenditures is important, whereas with learning from producers the composition of sales plays an important role.

As before, the evolution of the de-trended scale $\hat{\lambda}_{it} = \lambda_{it}/e^{\gamma/(1-\beta)t}$ is given by

$$\dot{\hat{\lambda}}_{it} = \hat{\alpha}_{it} B^P \sum_{j=1}^{n} r_{ji} \left(\frac{\hat{\lambda}_{it}}{\pi_{ji}} \right)^{\beta} - \frac{\gamma}{(1-\beta)} \hat{\lambda}_{it}, \tag{9}$$

and on a balanced growth path it solves the following system of non-linear equations

$$\hat{\lambda}_i = \frac{(1-\beta)\hat{\alpha}_i}{\gamma} B^P \sum_{j=1}^{n} r_{ji} \left(\frac{\hat{\lambda}_i}{\pi_{ji}} \right)^{\beta}. \tag{10}$$

3 Gains from Trade

As in other gravity models, a country's real income and welfare can be summarized by its stock of knowledge (or some other measure of aggregate productivity), its expenditure share on domestic goods, and the trade elasticity:

$$y_i \equiv \frac{w_i}{P_i} = B^{-1} \left(\frac{\lambda_i}{\pi_{ii}} \right)^{1/\theta} \tag{11}$$

[14] This mechanism is emphasized by Perla et al. (2013) and Sampson (2014).

In our model gains from trade have a static and dynamic component. The static component, holding each country's stock of knowledge fixed, is the familiar gains from trade in standard Ricardian models, e.g., Eaton and Kortum (2002).[15] The dynamic gains from trade are the ones that operate through the effect of trade on the flow of ideas.

In this section we consider several simple examples that illustrate the determinants of the static and dynamic gains from trade, both in the short and long run. We first consider an example of a world with symmetric countries. We study both the consequences of a simultaneous change in common trade barriers as well as the case of a single deviant country that is more isolated than the rest of the world. We also study how a small open economy responds when its trade barriers change, a case that admits an analytical characterization. The details of each are worked out in Appendix D.

3.1 Gains from Trade in a Symmetric Economy

Consider a world with n symmetric countries in which there is a common iceberg cost κ of shipping a good across any border. Specializing either equation (8) or equation (10), each country's de-trended stock of knowledge on a balanced growth path is

$$\hat{\lambda}(\kappa) = \left[\frac{(1-\beta)\hat{\alpha}}{\gamma} B^X\right]^{\frac{1}{1-\beta}} \left[\frac{1+(n-1)\kappa^{-\theta(1-\beta)}}{\left(1+(n-1)\kappa^{-\theta}\right)^{1-\beta}}\right]^{\frac{1}{1-\beta}}. \tag{12}$$

where B^X equals B^S or B^P depending on the specification.[16] The de-trended real per-capita income is obtained by substituting equation (12) into equation (11)

$$\begin{aligned} \hat{y}(\kappa) &= \frac{1}{B}\hat{\lambda}(\kappa)^{\frac{1}{\theta}}\left(1+(n-1)\kappa^{-\theta}\right)^{\frac{1}{\theta}} \\ &\propto \left[1+(n-1)\kappa^{-\theta(1-\beta)}\right]^{\frac{1}{\theta}\frac{1}{1-\beta}} \end{aligned} \tag{13}$$

Using equation (13), we derive a simple expression for the gains from trade. For the case of a world with symmetric populations, the long-run gains from trade as measured by the per-

[15] See Arkolakis et al. (2012) for other examples.

[16] For the special case of a symmetric world, it turns out that learning from producers or learning from sellers deliver exactly the same law of motion for countries' stocks of knowledge up to a constant. This can be seen by inspecting equation (4) and equation (9) and noticing that $\lambda_i = \lambda_j$ and $r_{ij} = \pi_{ij} = \pi_{ji}$ for all i, j. In general, the two specifications of idea flows are not equivalent.

capita income on a balanced growth path with costless trade relative to one with autarky, can be decomposed as follows

$$\frac{y^{FT}}{y^{AUT}} = n^{\frac{1}{\theta}} \quad \left(\frac{\lambda^{FT}}{\lambda^{AUT}}\right)^{\frac{1}{\theta}}$$
$$= \underbrace{n^{\frac{1}{\theta}}}_{static} \quad \underbrace{n^{\frac{\beta}{(1-\beta)\theta}}}_{dynamic} = n^{\frac{1}{\theta}\frac{1}{1-\beta}}. \tag{14}$$

The gains from trade depend on three parameters n, θ, and β. As in standard trade models, the gains from trade depend on the size of the country $1/n$ and the curvature of the distribution of productivity θ. The smaller each individual economy in this symmetric world, the more they gain by having access to the best producers abroad. In turn, the higher the curvature θ, the thinner the right tails of productivities. That is, there are fewer highly productive producers abroad which individuals can buy from under free-trade. The novel parameter determining the gains from trade is β. The parameter β controls the importance of insights from others in the quality of new ideas, i.e., the extent of technological spillovers associated with trade. With higher β, insights from others are more important, and therefore, more is gained by being exposed to more productive producers in a world with free trade. In the limit as β goes to 1, holding fixed θ, the gains from trade relative to autarky grow arbitrarily large. This limiting case is the one analyzed by Alvarez et al. (2013).[17]

Another way of interpreting equation (14) is that the diffusion of ideas causes the static gains from trade to compound itself. The expression for the static and dynamic gains from trade shares features with an analogous expression in a static world in which production uses intermediate inputs. In that world, a decline in trade costs reduces the costs of production, lowering the cost of intermediate inputs, which lowers the cost of production further, etc. Here, when trade costs decline, producers draw better insights from others, raising stocks of knowledge, and this improves the quality of insights others draw, etc. The parameter β gives the contribution of an insight to a new idea, just as the share of intermediate goods measures the contribution of the cost of intermediate inputs to marginal cost.

More generally, we can analyze the gains from trade for intermediate trade costs. For several

[17] When $\beta = 1$, the steady state gains from moving from autarky to free trade are infinite because integration raises the growth rate of the economy. In contrast, for any $\beta < 1$, integration raises the level of incomes but leaves the growth rate unchanged.

values of β, the left panel of Figure 1 illustrates the common value of each country's stock of knowledge relative to its level under free trade. The right panel shows the corresponding real income per capita. The dashed line in the right panel represents $\beta = 0$, which corresponds to the Eaton and Kortum (2002). As trade costs rise, countries becomes more closed and their stocks of knowledge decline. When β is larger, the dynamic gains from trade are larger.

Note that the dynamic gains from trade are largest when the world is relatively closed, whereas the static gains from trade are largest when the world is relatively open. To understand this, consider a country close to autarky. If trade costs decline, the marginal import tends to be made by foreign producers with high productivity and the marginal export tends to be made by domestic producers with high productivity. While the high trade costs imply that the static gains from trade remain relatively small, the insights drawn from these marginal producers tend to be of high quality. In contrast, for a country close to free trade, the reduction in trade costs leads to large infra-marginal static gains from trade, but the insights drawn from the marginal producers are likely to be lower quality.

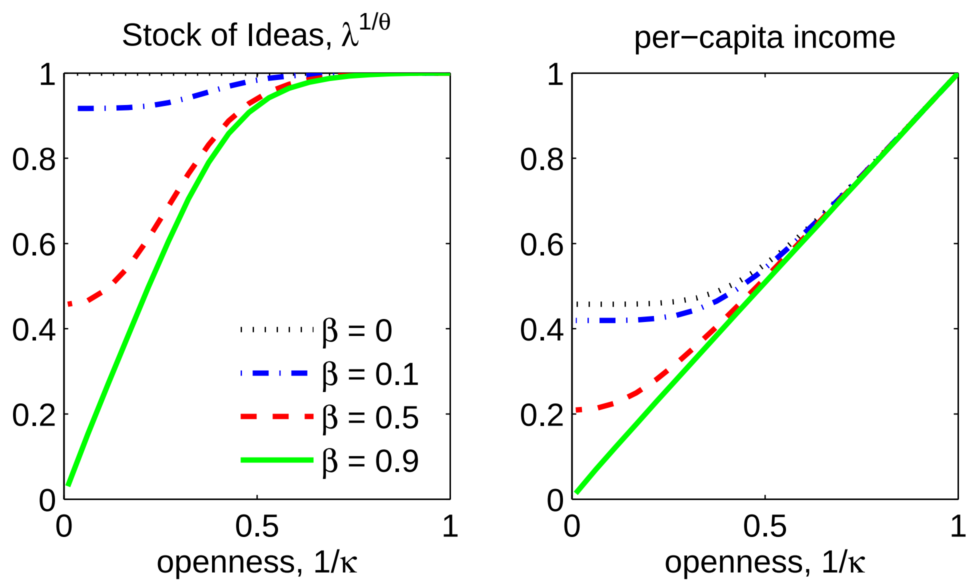

Figure 1: Gain from Reducing common trade barriers

3.2 Asymmetric Economies

What is the fate of a single country that is more open than others? Or one that is closed off from world trade? This section studies gains from trade in an asymmetric world in two simple ways.

We first describe how trade costs affect real income of a small open economy. We then return to example of a symmetric world discussed in the previous section but with a single "deviant" country that is more isolated.

Small Open Economy Consider first a small open economy that is small in the sense its actions do not impact other countries' stocks of knowledge, real wages, or expenditures. Let i be the small open economy, and suppose that all trade costs take the form of $\kappa_{ij} = \kappa\tilde{\kappa}_{ij}$ and $\kappa_{ji} = \kappa\tilde{\kappa}_{ji}$ for $j \neq i$. The claim below summarizes, to a first order, the long-run impact of a change in κ on country i's real income.

Claim 3 *Consider the small open economy described above. The long-run response of steady state income per capita to trade costs in each specification of learning is*

$$\text{Sellers:} \quad \frac{d\log y_i}{d\log \kappa} = -\frac{1+2\theta}{\frac{1-\Omega_{ii}^S\beta}{1-\Omega_{ii}^S}\frac{\pi_{ii}+\theta(1+\pi_{ii})}{(1-\beta)+\beta(1-\pi_{ii})}+1}$$

$$\text{Producers:} \quad \frac{d\log y_i}{d\log \kappa} = -\frac{(1-\beta)(1+2\theta)+\beta(1-\pi_{ii})}{\frac{1-\beta}{1-\Omega_{ii}^P}\left[\pi_{ii}+\theta(1+\pi_{ii})\right]+1+\beta\pi_{ii}}$$

where $\Omega_{ii}^S = \frac{\pi_{ii}^{1-\beta}\lambda_i^\beta}{\sum_j \pi_{ij}^{1-\beta}\lambda_j^\beta}$ *and* $\Omega_{ii}^P = \frac{r_{ii}(\lambda_i/\pi_{ii})^\beta}{\sum_j r_{ji}(\lambda_i/\pi_{ji})^\beta}$.

With learning from sellers, the term Ω_{ii}^S is the share of the growth in i's stock of knowledge that is associated with purchasing goods from i; with learning from producers, Ω_{ii}^P is the share associated with producing goods for i. One implication is that, holding fixed π_{ii}, the response of real income to a decline in trade costs is larger when Ω_{ii}^S is smaller. In words, this means that, among small open economies with the same trade shares, the response of real income to trade will be larger when the country relies more on others for growth in its stock of knowledge. For example, a country with a low stock of knowledge will rely more on others for good quality insights. When such a country reduces trade barriers, the impact on income is larger. This is one form of catch-up growth.

Single Deviant Economy We consider an asymmetric version of the model with $n-1$ open countries, $i = 1, ..., n-1$, and a single deviant economy, $i = n$. The $n-1$ open countries can freely trade among themselves, i.e., $\kappa_{ij} = 1$, $i, j < n$, but trade to and from the deviant economy incurs transportation cost, i.e., $\kappa_{nj} = \kappa_{jn} = \kappa_n \geq 1$, $j < n$.[18]

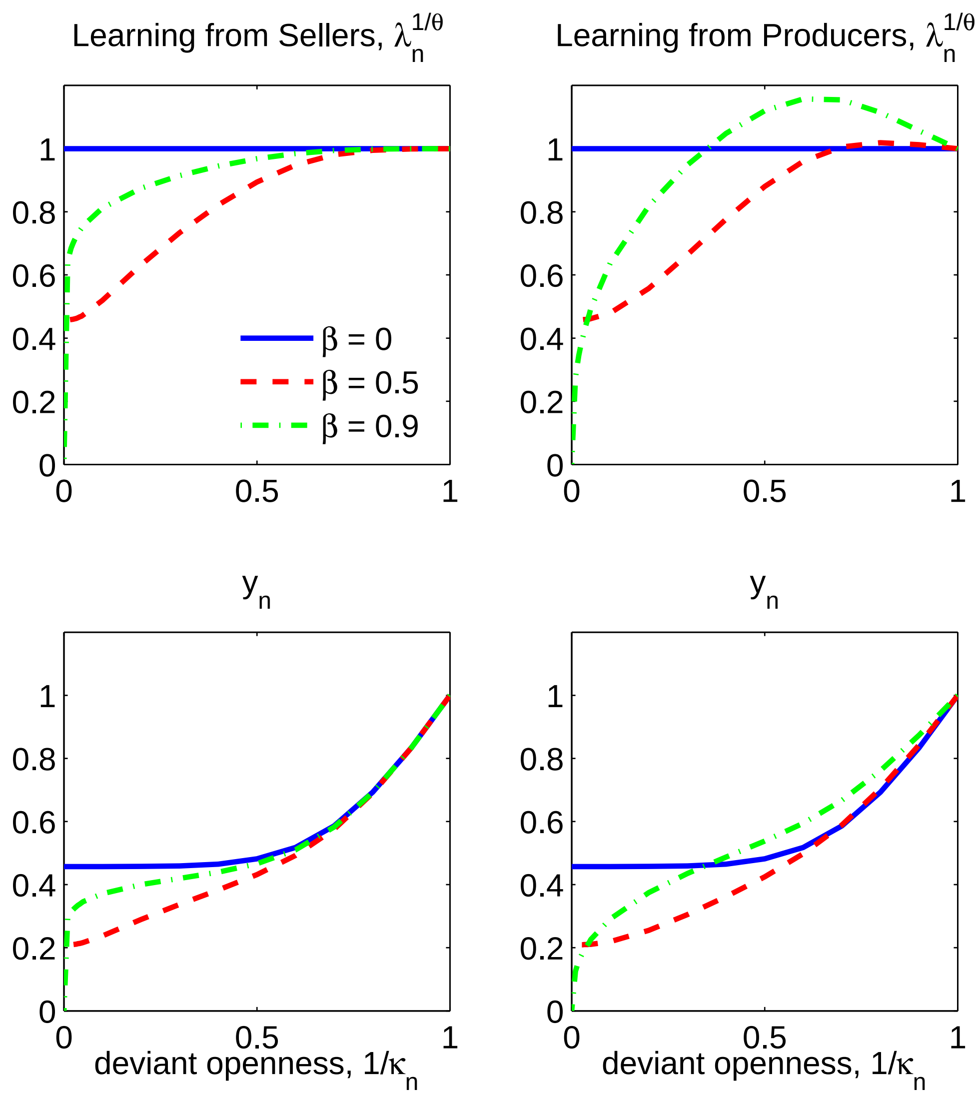

Figure 2: The Stock of Ideas and Per Capita Income of the Deviant Economy.

The top panels of Figure 2 show the deviant country's stock of knowledge changes with the degree of openness. The x-axis measures openness as the inverse of the cost to trade to and from

[18]In the numerical examples that follow, we consider a world with $n = 50$ economies with symmetric populations, so that each country is of the size of Canada or South Korea. We set $\theta = 5$, the curvature of the Frechet distribution, which itself equals the tail of the distribution of exogenous ideas. This value is in the range consistent with estimates of trade elasticities. See Simonovska and Waugh (2014), and the references therein. Given a value of β, the growth rate of the arrival rate of ideas is calibrated so that on the balanced growth path each country's TFP grows at 1%, $\frac{\gamma}{(1-\beta)\theta} = 0.01$. The parameter $\hat{\alpha}$ is normalized so that in the case of costless trade, $\kappa_n = 1$, the de-trended stock of ideas equals 1.

the deviant economy, $1/\kappa_n$. On the y-axis we report the stock of ideas relative to the case with costless trade ($\kappa_n = 1$). The bottom panels graph the corresponding real incomes. The different lines correspond to alternative values of β, which controls the importance of insights from others. The solid line shows the effect of openness in the case with no spillovers, $\beta = 0$, which also equals the effect in the standard static trade theory of Eaton and Kortum (2002). The other two curves correspond to cases with positive technological spillovers. The left panels correspond to learning from sellers while he right panes show learning from domestic producers.

When ideas diffuse through learning from sellers, we see that, across balanced growth paths, as deviant economy becomes more isolated its stock of ideas contracts relative to that of the balanced growth path of n economies engaging in costless trade. As discussed earlier, through their effect on the stock of ideas trade costs have effects on per-capita income beyond the static gains from trade.

Figure 2 has two curious features. The upper right panel shows how a the deviant country's stock of knowledge varies with openness when insights are drawn from domestic producers. Starting from the case of costless trade, the stock of knowledge in the deviant economy initially *improves* as the trade cost to and from this country increases, and it becomes isolated. The dynamic gains from trade are negative in this range. As trading becomes more costly, there is a first order improvement in the quality of insights because of more stringent selection into exporting. For an economy in which most of production is exported, this outweighs the deterioration of the quality from less stringent selection into producing for domestic consumption.[19] Eventually, the negative selection effects of inefficient domestic producers entering dominates, and the stock of ideas in the deviant economy deteriorates. The dynamic gains from trade are positive, and can be very large as the deviant economy approaches autarky. These effects are more pronounced the higher the degree of spillovers as measured by the parameter β.

Second, if the deviant economy is moderately open, the gains from trade are non-monotonic in β. This holds across both specifications of learning. When β is zero, there are no dynamic gains from trade. When β is large, the dynamic gains are very large near autarky but are much smaller once the country is moderately open. This stems from the concavity generated by β in

[19] If insights were to be randomly drawn from the set of active domestic producers and trade barriers satisfy the triangle inequality, the evolution of the stock of knowledge is $\dot{\lambda}_{it} = \Gamma(1-\beta)\alpha_{it}(\lambda_i/\pi_{ii})^{\beta}$. In this case, the growth rate of the stock of ideas would be maximized in the case of costless trade, $\pi_{ii} = 1/n$, provided we abstract from import subsidies. The only effect of higher trade costs in this version of the model is to worsen the selection of domestic producers.

the combination of insights with exogenous components of ideas. When β is large, the difference between a high and low quality insight is magnified. Thus when β is large, a countries growth depends much more heavily on insights from the most productive producers. When a country is only moderately open, most of these most productive producers will have already entered. Indeed, as $\beta \to 1$ then $\lambda_n \to 1$ for any finite κ_n.[20]

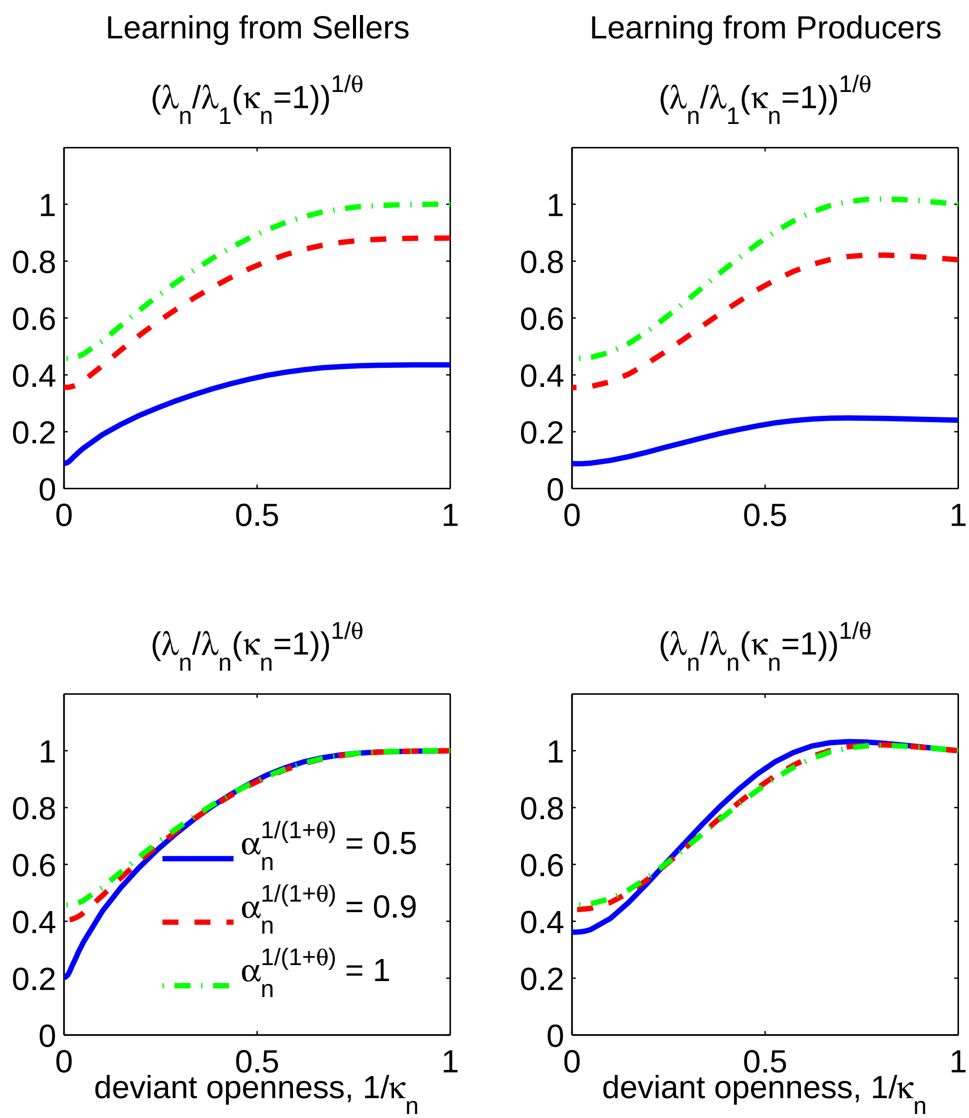

Figure 3: The arrival rate of ideas: The Stock of Ideas and Per Capita Income of the Deviant Economy.

Figure 3 demonstrates how the arrival rate α affects the cost of isolation. Consider a world in

[20] Alvarez et al. (2013) analyzed the limit point $\beta = 1$. In particular, their Proposition 7 and 8 show that the behavior of the tail of the distribution of productivity is independent of trade costs, as long as they are finite.

which $n-1$ of the economies have de-trended arrival rate $\hat{\alpha}=1$ and can trade costlessly among themselves, and a single deviant economy has arrival rate $\hat{\alpha}_n<1$ and faces iceberg costs κ. The top panels show the deviant country's stock of knowledge relative to that of the open countries in a world in which all trade is costless. The bottom panels show the deviant country's stock of knowledge relative to its own stock of knowledge in a world with costless trade. The left panels show learning from sellers while the right panels show learning from producers.

As would be expected, a lower arrival rate reduces the deviant country's stock of knowledge. Inspection of the top panels reveals that when insights come from domestic producers, a reduced arrival rate has a bigger impact on the stock of knowledge, even with costless trade ($\kappa=1$). This happens because the reduced stock of knowledge is compounded; the same producers that have the lower arrival rate are those that provide insights.

Does the arrival rate impact the cost of isolation? The bottom panels reveal that for economies that are open at least a moderate amount, the arrival rate does not impact the dynamic gains from trade. However, for countries close to isolation, the low arrival rate compounds the cost of isolation. The logic is similar to that of the last paragraph: when an economy is close to isolation, learning from sellers resembles learning from producers.

3.3 Trade Liberalization

We now study how a country's stock of knowledge and real income evolve when it opens to trade. Does the country experience a period of protracted growth or does it converge relatively quickly?

We first address this question by studying a small open economy. For this economy, we can log-linearize around its long-run real income to derive an exact expression for the speed of convergence. Even though the world may have arbitrary bilateral trade costs, Claim 4 provides a relatively simple expression for the speeds of convergence in terms of the small open economy's share of expenditures on domestic goods π_{ii} and the share of growth in its stock of knowledge that comes from insights from domestic firms.

Claim 4 *For any variable x, let $\check{x}$ denote the log deviation from its long run value, and let x (no decoration) denote its long run value. When agents learn from sellers, the speed of convergence in* *a small open economy is*

$$\frac{d}{dt}\log\left[\check{w}_i - \check{P}_i\right] \quad = \quad -\gamma\left\{1 - \frac{\Omega_{ii}^S - \pi_{ii}}{1+\theta\left(1+\pi_{ii}\right)} + \frac{\beta}{1-\beta}\left(1-\Omega_{ii}^S\right)\right\}$$

If agents learn from producers, the speed of convergence is

$$\frac{d}{dt}\log\left[\check{w}_i - \check{P}_i\right] \quad = \quad -\gamma\left\{1 - \frac{\Omega_{ii}^P - \pi_{ii}}{1+\theta\left(1+\pi_{ii}\right)} + \frac{\beta}{1-\beta}\frac{\left(1-\Omega_{ii}^P\right)\left(1+\pi_{ii}\right)}{1+\theta\left(1+\pi_{ii}\right)}\right\}$$

The expressions for the speed of convergence gives several immediate implications. First, convergence is faster when diffusion is more important (β is larger), and this effect is magnified when insights from others contributes more to growth (Ω_{ii}^S or Ω_{ii}^P is small). Second, the arrival rate of ideas does not appear anywhere.

In addition, convergence is faster with learning from sellers than with learning from domestic producers under assumptions that make the two specifications comparable.[21] When insights are drawn from sellers, a trade liberalization gives immediate access to insights from goods sold by high productivity foreign producers. In contrast, when insights are drawn from domestic producers, the insights are initially low quality, although they become more selected, and only gradually improve as the country's stock of knowledge increases.

These implications can be illustrated by a world economy that starts with $n-1$ open economies and a single deviant economy that are on a balanced growth path. Initially, the $n-1$ open economies can trade at no cost among themselves, $\kappa_1 = 1$, but trade to and from the deviant economy faces large trade costs, $\kappa_n = 100$. The initial balanced growth path corresponds to a point close to autarky. We then trace the evolution of the stock of ideas and per-capita income as trade costs to and from the (former) deviant economy are eliminated, $\kappa_n = 100 \to \kappa_n' = 1$. The paths of the (de-trended) stock of knowledge solve the differential equations in (7) and (9), depending on whether insights are drawn from sellers or producers.[22]

Figure 4 shows the evolution of the stock of ideas (left panel) and per-capita income (right panel) in the initially deviant economy, following the elimination of trade costs. The solid line

[21] The two specifications are comparable if each is moving toward the same steady state and if $\Omega_{ii}^P = \Omega_{ii}^S$. In that case learning from sellers has a faster transition because $\frac{\left(1+\pi_{ii}^{ss}\right)}{1+\theta\left(1+\pi_{ii}^{ss}\right)} < 1$.

[22] We set $\beta = 0.5$. The rest of the parameters follow the calibration in footnote 18.

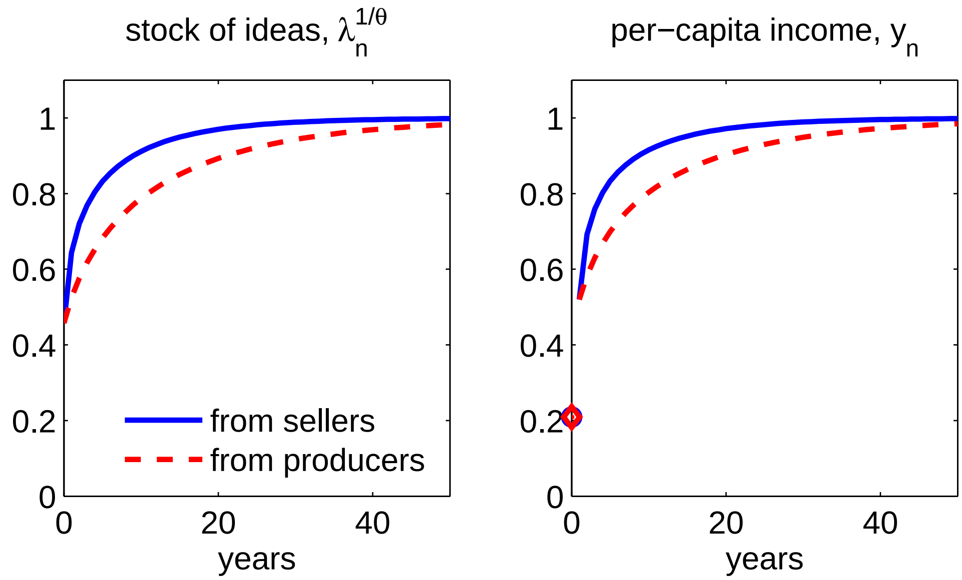

Figure 4: Dynamics Following a Trade Liberalization for Alternative Sources of Insights.

corresponds to the transitional dynamics in the model where learning is from sellers to a country, while the dashed line gives the dynamics for the model in which insights are drawn from domestic producers.

Following the elimination of trade costs, the stock of ideas in the deviant economy slowly converges to that of the $n-1$ open countries. On impact real income jumps as it would in a static model. Furthermore, over time, it continues to increase as the stock of knowledge improves. The speed of this process depends on the sources of insights, as discussed above. Notice that for this example, the long-run effect on per-capita income is substantially larger than the effect on impact, and that the transition is very protracted. These examples suggest that our model provides a promising theory of growth miracles fueled by openness, and underscores the importance of investigating the particular mechanism for the diffusion of ideas.

4 Quantitative Exploration

We now explore the ability of the theory to account for the evolution of the distribution of productivity (TFP) across countries in the post-war period. We extend the simple trade model introduced in Section 2 to incorporate intermediate inputs, non-traded goods, and a broader notion of labor which we refer to as equipped labor.

In particular, we assume that in each country i a producer of good s with productivity q has access to a constant returns to scale technology combining intermediate input aggregate (x) and equipped labor (l)

$$y_i(s) = \frac{1}{\eta^\eta (1-\eta)^{1-\eta}} q x_i(s)^\eta l_i(s)^{1-\eta}$$

All goods use the intermediate good aggregate, or equivalently, the same bundle of intermediate inputs. The intermediate input aggregate is produced using the same Constant Elasticity of Substitution technology as the consumption aggregate, so that the market clearing condition for intermediate inputs for i is

$$\int x_i(s) ds = \left[\int \chi_i(s)^{1-1/\varepsilon} ds \right]^{\varepsilon/(\varepsilon-1)}.$$

where $\chi_i(s)$ denotes the amount of good s used in the production of the intermediate input aggregate. Equipped labor L is produced with an aggregate Cobb-Douglas technology requiring capital and efficiency units of labor

$$L_i = \int l_i(s) ds = K_i^\zeta (h_i \tilde{L}_i)^{1-\zeta}.$$

In our quantitative exercises we take an exogenous path of aggregate physical and human capital, K_i and h_i, from the data, therefore, we abstract from modeling the accumulation of these factors.[23]

In addition to the iceberg transportation costs κ_{ij}, we assume that a fraction μ of the goods are non-tradable, i.e., this subset of the goods face infinite transportation costs. The main effect of introducing non-traded goods is that in the extended model the value of the elasticity of substitution ε affects equilibrium allocations. In the Appendix we derive the expressions for price indices, trade share and profits of this version of the model.

[23] Implicitly, we are assuming that individual technologies are

$$y(s) = \frac{1}{\eta^\eta (1-\eta)^{1-\eta} \zeta^{(1-\eta)\zeta} (1-\zeta)^{(1-\eta)(1-\zeta)}} q x(s)^\eta [k(s)^\zeta (h_i l(s))^{1-\zeta}]^{1-\eta}$$

and that investment can be produced with the same Constant Elasticity of Substitution technology as the consumption and intermediate input aggregates.

4.1 Calibration

We need to calibrate six common parameters, $(\theta, \eta, \zeta, \mu, \gamma, \varepsilon)$, and two set of parameters that are (potentially) country and time specific, the matrix of transportation costs $\mathbf{K}_t = [\kappa_{int}]$ and the vector of arrival rates $\alpha_{\mathbf{t}} = (\alpha_{1t}, ..., \alpha_{nt})$. In addition, we need to assign a value to the diffusion parameter β. Instead of calibrating it, we present results for alternative values of $\beta \in [0, 1)$.

We set $\theta = 5$. This value is in the range consistent with estimates of trade elasticities. See Simonovska and Waugh (2014), and the references therein. We let $\eta = 0.5$ and $\zeta = 1/3$ to match the intermediate share in gross production and the labor share of value added. We consider a share of non-trade goods $\mu = 0.5$. Given values for θ and β, we choose γ to match an average growth rate of TFP in the US of 1.01 percent. We set $\varepsilon = 1$ to guarantee that the constants B^S and B^P are finite, but note that alternative values that satisfy this restriction do not affect the results significantly.

As we show in the Appendix, given values for θ and μ, an assumption that $\kappa_{ij} = \kappa_{ji}$, and data on bilateral trade shares over time, the iceberg cost of shipping a good to country i from country j at time t is

$$\kappa_{ijt} = \left[\frac{1 - \pi_{iit}}{\pi_{ijt}} \frac{1 - \pi_{jjt}}{\pi_{jit}} \left(\frac{Z_{it}}{1 - Z_{it}} \right) \left(\frac{1 - Z_{jt}}{Z_{jt}} \right) \right]^{\frac{1}{2\theta}}$$

where Z_{it} solves

$$\pi_{iit} = \frac{(1 - \mu) + \mu Z_{it}^{1 - \frac{\varepsilon - 1}{\theta}}}{(1 - \mu) + \mu Z_{it}^{-\frac{\varepsilon - 1}{\theta}}}.$$

To operationalize this equation, we use bilateral trade data for 1962-2000 from Feenstra et al. (2005) and GDP data from PWT 8.0.

To assign values to the vector of arrival rates $\hat{\alpha}_{\mathbf{t}} = (\hat{\alpha}_{1t}, ..., \hat{\alpha}_{nt})$ we follow two alternative strategies. First, we consider the case in which countries have common and constant arrival rate of ideas $\hat{\alpha}_{it} = \hat{\alpha}$. Given that we can always normalize one element of the vector $\hat{\alpha}$, we set the common arrival rate of ideas to $\hat{\alpha} = 1$. In this case, the differences in TFP across countries are only generated by their differential degree of openness. The second strategy consists of assuming that countries are on a balanced growth path in 1962, and choose the vector $\hat{\alpha}$ to match the relative TFP across countries in 1962.

4.2 Explaining the Distribution of TFP

We first explore the ability of the theory to account for the initial distribution of TFP. To do so, we assume that countries have common and constant arrival rate of ideas $\hat{\alpha}_{it} = \hat{\alpha}$, and calculate the distribution of TFP in a balance growth path. If we further assume that countries are in a balance growth path, differences in per-capita income can only be explained by the measured differences in transportation costs. We consider the case in which domestic producers learn from sellers to their market.

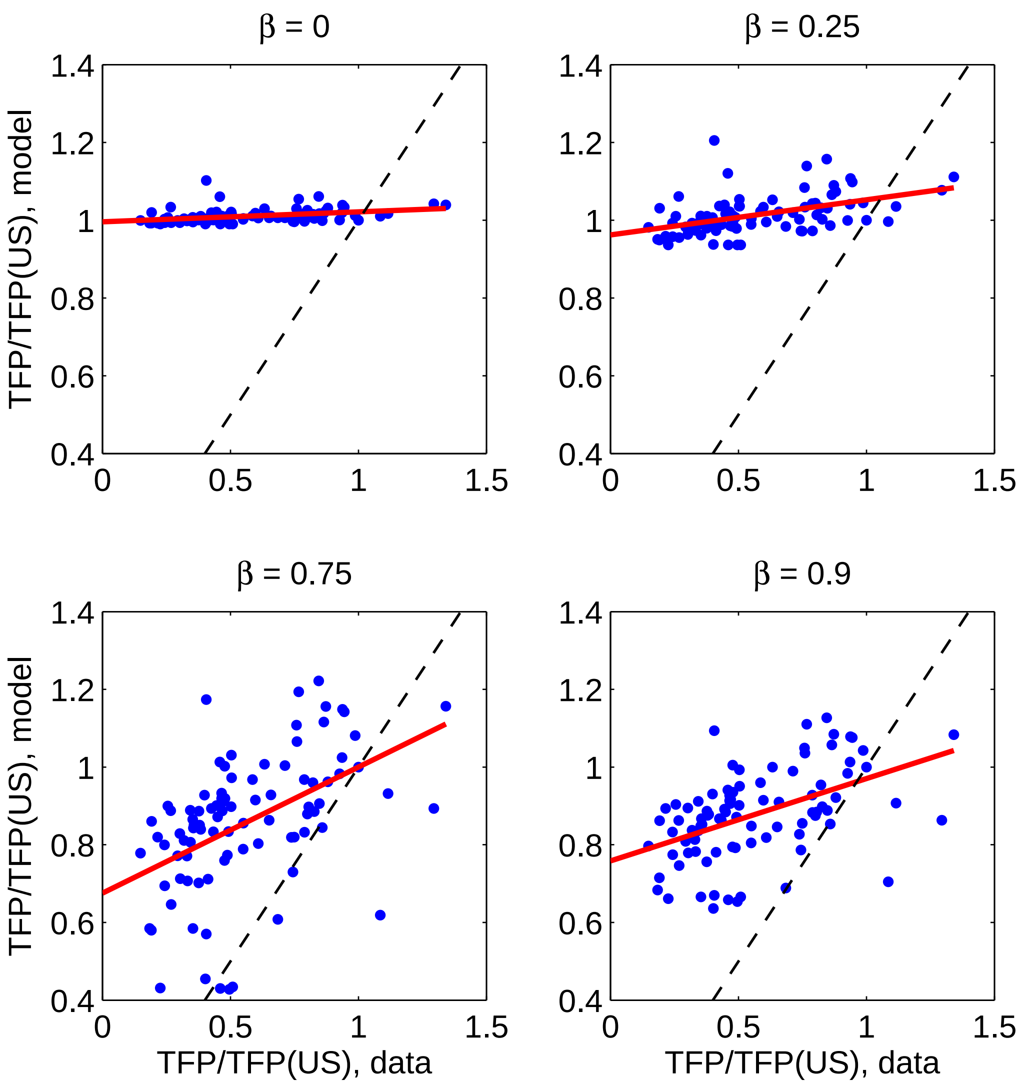

Figure 5: Openness and the Distribution of TFP in 1962. Alternative β, Learning from Sellers.

Common Arrival Rate of Ideas In Figure Figure 5 we compare the implied distribution of TFP in the balance growth path of model with common α's with the the observed distribution of TFP in 1962, the first year in our sample. We consider The

4.3 Explaining the Dynamics of TFP

We illustrate the ability of the theory to account for the evolution of the distribution of productivity (TFP) by comparing the implied evolution of TFP in South Korea and the US. South Korea is a particularly interesting example as it is one of the most successful growth miracles in the post-war period, and a country that became most integrated with the rest of the world, as inferred from the behavior of trade flows. The U.S. economy provides a natural benchmark developed economy.

Common Arrival Rate of Ideas We first explore the implied dynamics of TFP under the assumption of a common arrival rate of ideas, $\alpha_{it} = 1$. Figure 6 shows the evolution of TFP for South Korea (top panels) and the US (bottom panels) for this case. The left and right panels show the implied dynamics of TFP for specifications in which producers learn from sellers and from other domestic producers, respectively. The solid line shows the evolution of TFP in the data, de-trended by the average growth of TFP in the U.S. The other lines correspond to simulations using alternative values of the diffusion parameters β. The case of $\beta = 0$ (dotted line) gives the dynamics of TFP implied by a standard Ricardian trade model, e.g., the dynamics quantified by Connolly and Yi (2009). The other three lines illustrate the dynamic gains from trade implied by the model.

Three clear messages stem from this figure. First, for a wide range of values of the diffusion parameter the dynamic model accounts for a substantial fraction of the TFP dynamics of South Korea. This is particularly true when considering intermediate values of the diffusion parameters β. Recall from Figure 2 that for an economy that is moderately open, dynamic gains from trade are non-monotonic in β. Second, the results are robust to the particular specification of the learning used, i.e., learning from sellers vs. learning from domestic producers. This shows that the differences across the alternative learning models discussed in Section 2 are not quantitatively important when considering realistic examples. Finally, the bottom panels show that changes in the dynamic gains from trade identified by the model are less relevant understanding the growth experience of a

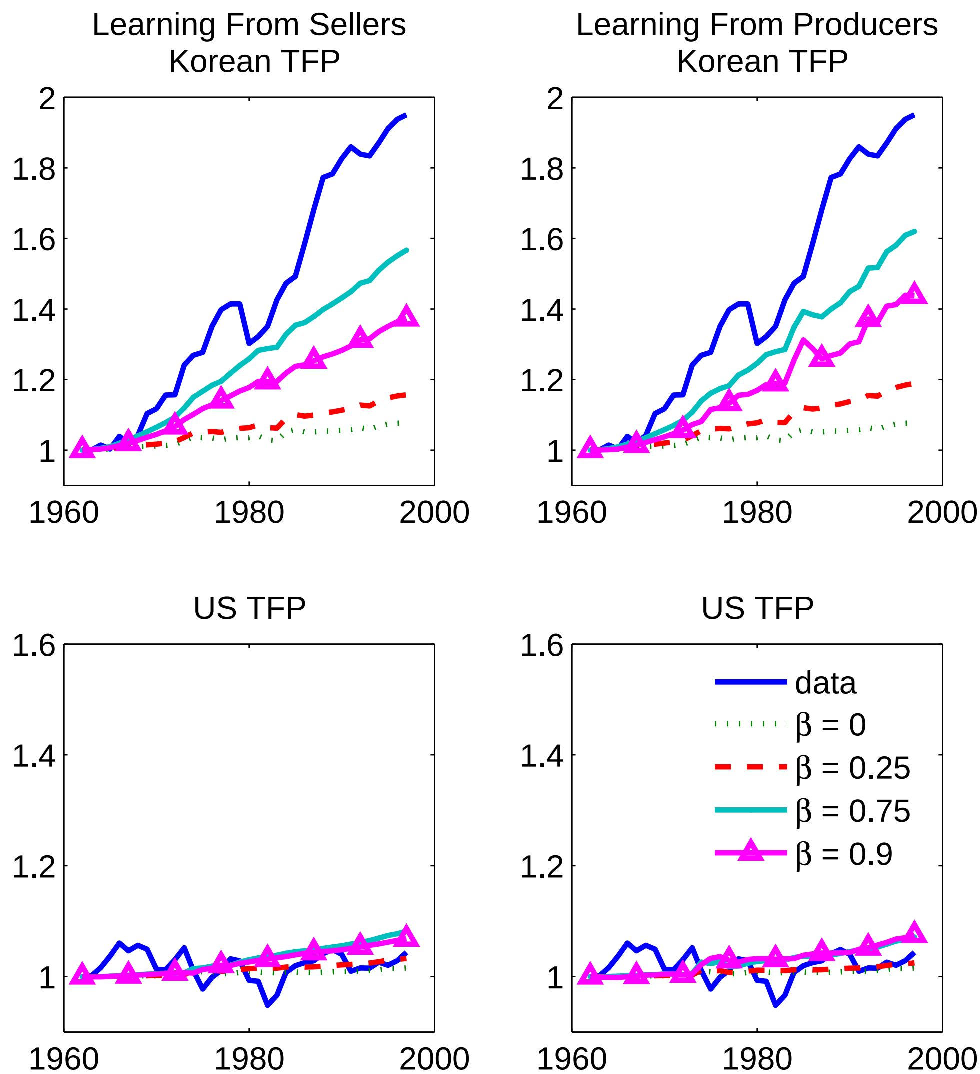

Figure 6: Openness and the Evolution of TFP: Alternative β and learning mechanism.

developed country close to its balanced growth path.

Heterogeneous Arrival Rate of Ideas We next consider the case with country specific arrival rates of ideas $\hat{\alpha}_{i0}$ that are chosen to account for initial difference in TFP across countries, i.e., the difference in TFP that cannot be accounted by the initial differences in openness κ_{ij0}. This exercise underscores an important interaction between the arrival rate of ideas α and trade costs κ that we showed with the simple examples in Figure 3. In addition, we consider cases where we allow for the arrival rate of ideas to change (linearly) over time α_{it} so that our model can account for the entire growth of TFP observed in the sample. For simplicity we focus on the case where producers learn from sellers to their market.

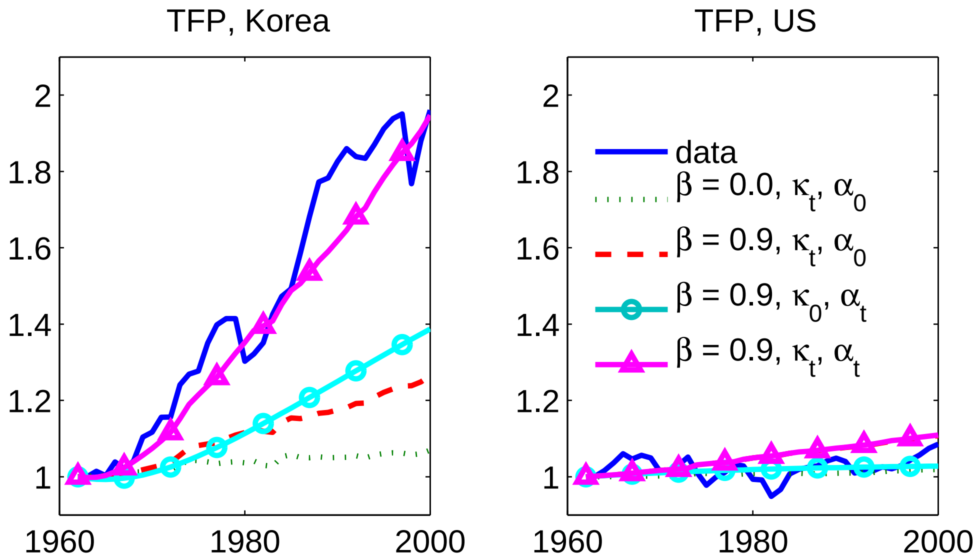

Figure 7: Openness and the Evolution of TFP: The Role of Openness and the Intensity of Ideas Arrival.

Figure 7 shows the evolution of TFP for South Korea (left panel) and the US (right panel). As before, the solid and dotted lines show the dynamics of TFP in the data and the static Ricardian model ($\beta = 0$), respectively. The dashed line shows the impact of openness when the arrival rate of ideas is kept at the initial (low) level. For the case of South Korea, the model requires a very low arrival rate of ideas to account for the low initial TFP, and therefore, the impact of the increase in openness is significantly muted as compared to the one suggested by Figure 6. The line with triangles is the one that we obtain when we fit a linear growth in the arrival rate of ideas so that, given the value of $\beta = 0.9$, the model can account for the entire increase in TFP for South Korea. To measure the contribution of openness the circled line shows the evolution of TFP if we were to keep the initial (large) trade costs, while changing α_{it} as described before. For the case of $\beta = 0.9$ a majority of the TFP is accounted by the endogenous diffusion mechanism in the model.

5 Trade and Multinational Production

The set of producers in a country may be influenced by multinationals' decisions of where to locate production. Some countries have policies to encourage FDI, and this might be related to positive spillovers to local producers. In this section we extend the basic model to allow for multinational production, and insights drawn from foreign technologies used domestically.

We follow Ramondo and Rodriguez-Clare (2013) in modeling a multinational as a producer that can manufacture a single variety with productivity that varies with the location of production. For variety s, a multinational is characterized by a vector of productivities $\mathbf{q}(s) = \{q_j(s)\}_j$, where $q_j(s)$ is the best productivity available to the multinationals when producing variety s in country j.

Producing abroad is further subject to iceberg costs. Suppose a producer of variety s based in i can produce in j with productivity q. That producer's marginal cost will be $\frac{w_j \delta_{ji}}{q_{ji}(s)}$. We assume $\delta_{ji} \geq 1$ and $\delta_{ii} = 1$.

We first study learning from a general source distribution. Among all varieties, let $M_{it}(\{q_1, ..., q_n\})$ be the joint distribution of productivities across locations among multinationals based in i, so that $F_{it}(\{q_1, ..., q_n\})^m$ represents the frontier of knowledge. Suppose an individual production manager draws an insight from a source distribution, $G_{it}(q)$. Upon drawing insight from an idea with productivity $\tilde{q}$, the manager is provided with a vector of ideas with productivities $\left\{z_j \tilde{q}^\beta\right\}_j$ across locations. The idiosyncratic portion of the productivities, $\{z_j\}$, are drawn from a multivariate distribution with joint CDF $H(\{z_1, ..., z_n\})$.

Extending the analysis in Section 1, the distribution of productivity of technologies based on country i

$$\frac{d}{dt} \ln F_{it}(\{q_1, ..., q_n\}) = -\alpha_t m \int_0^\infty \left[1 - H\left(\left\{\frac{q_1}{\tilde{q}^\beta}, ..., \frac{q_n}{\tilde{q}^\beta}\right\}\right)\right] dG_{it}(\tilde{q})$$

We assume that the idiosyncratic components of new ideas are drawn from a joint distribution with right tails that are jointly regularly varying.

Assumption 4

$$\lim_{x\to\infty} \frac{1 - H(x\mathbf{z})}{x^{-\theta}} = \left(\sum_{i=1}^n z_i^{-\frac{\theta}{1-\rho}}\right)^{1-\rho}$$

with $\rho \in [0, 1]$. Each marginal distribution has a Pareto right tail with tail index θ. Using the same logic as the baseline analysis, we can define a $\tilde{F}(\mathbf{q}) \equiv F\left(m^{\frac{1}{\theta(1-\beta)}}\mathbf{q}\right)$ and $\tilde{G}(\mathbf{q}) \equiv G\left(m^{\frac{1}{\theta(1-\beta)}}\mathbf{q}\right)$ to be the frontier distribution and the source distribution scaled by the number of multinationals. The law of motion for the frontier is

$$\frac{d}{dt}\ln \tilde{F}_{it}\left(\{q_1,...,q_n\}\right) = -\alpha_t \left(\sum_j q_j^{-\frac{\theta}{1-\rho}}\right)^{1-\rho} \int_0^\infty \tilde{q}^{\beta\theta} d\tilde{G}_{it}(\tilde{q})$$

Asymptotically, the frontier converges to a Frechet distribution

$$\tilde{F}_{it}\left(\{q_1,...,q_n\}\right) = e^{-\lambda_{it}\left(\sum_j q_j^{-\frac{\theta}{1-\rho}}\right)^{1-\rho}}$$

with

$$\dot{\lambda}_{it} = \alpha_t \int_0^\infty \tilde{q}^{\beta\theta} d\tilde{G}_{it}(\tilde{q})$$

The frontier of knowledge thus takes the form of a multivariate Frechet.[24]

We now describe the equilibrium, leaving the derivation of the equations to Appendix E. Define π_{ijk} to be the share of i's expenditure on goods produced in j by by multinationals based in k. Similarly, let $r_{ijk} \equiv \frac{\pi_{ijk}X_i}{\sum_{\tilde{i}\tilde{k}} \pi_{\tilde{i}j\tilde{k}} X_{\tilde{i}}}$ be the share of j's revenue from sales to i of goods produced by multinationals based in k. Thus $1 = \sum_j \pi_{ij} = \sum_j \sum_k \pi_{ijk}$ and $1 = \sum_i r_{ij} = \sum_i \sum_k r_{ijk}$. Given wages, the trade shares and the price level are

$$\pi_{ijk} = \frac{\lambda_k \left(\sum_l \left[w_l \kappa_{il} \delta_{lk}\right]^{-\frac{\theta}{1-\rho}}\right)^{-\rho} \left[w_j \kappa_{ij} \delta_{jk}\right]^{-\frac{\theta}{1-\rho}}}{\sum_{\tilde{k}} \lambda_{\tilde{k}} \left(\sum_l \left[w_l \kappa_{il} \delta_{l\tilde{k}}\right]^{-\frac{\theta}{1-\rho}}\right)^{1-\rho}}$$

and

$$P_i = B \left(\sum_k \lambda_k \left(\sum_j \left[w_j \kappa_{ij} \delta_{jk}\right]^{-\frac{\theta}{1-\rho}}\right)^{1-\rho}\right)^{-\frac{1}{\theta}}.$$

where B is defined in Section 2. The set of wages that clear each country's labor market solve the set of equations

$$w_j L_j = \frac{\theta}{\theta+1} \sum_i \sum_k \pi_{ijk} X_i = \frac{\theta}{\theta+1} \sum_i \pi_{ij} X_i.$$

If individuals learn from sellers, the evolution of country i's stock of knowledge can be summarized

[24] It is given this name because each marginal distribution is Frechet.

by the following differential equation:

$$\dot{\lambda}_i = B^S \alpha \sum_j \sum_k \pi_{ijk} \left(\frac{\lambda_k}{\pi_{ijk}^{1-\rho} \left[\sum_l \pi_{ilk}\right]^\rho} \right)^\beta$$

whereas if individuals learn from domestic producers, country i's stock of knowledge evolves as

$$\dot{\lambda}_i = B^P \alpha \sum_j \sum_k r_{jik} \left(\frac{\lambda_k}{\pi_{jik}^{1-\rho} \left[\sum_l \pi_{jlk}\right]^\rho} \right)^\beta$$

where B^S and B^P are defined in Section 2.

There are a number of ways that production managers in i can get insights from production managers in k. When learning from sellers is important, a manager in i can draw insight from multinationals based in k from goods produced in any location. When learning from producers is important, a manager in i can learn from multinationals based in k that produce goods in i to export to j.

When a multinational's productivities are uncorrelated across production locations ($\rho = 0$), the logic of selection of ideas is exactly the same as the case without FDI: holding fixed k's stock of knowledge, for each combination of multinational's home and production location, the smaller i's expenditure, the more likely it is that insights are drawn from higher productivity goods. When a multinational's ideas are correlated across locations ($\rho > 0$), the logic is similar.

Whether trade and FDI are complements or substitutes in learning and production depends on the correlation of multinationals' productivities across locations. This can be seen most easily in the case of symmetric economies. Let $y(\kappa, \delta)$ be real income when trade costs between any pair of countries is κ and FDI costs between any pair of countries is δ. In this setting, real income relative to costless trade can be summarized concisely for two polar cases:

$$\lim_{\rho \to 0} \frac{y(\kappa, \delta)}{y(1,1)} = \left[\left(\frac{1 + (n-1)\kappa^{-\theta(1-\beta)}}{n} \right) \left(\frac{1 + (n-1)\delta^{-\theta(1-\beta)}}{n} \right) \right]^{\frac{1}{\theta(1-\beta)}}$$

and

$$\lim_{\rho \to 1} \frac{y(\kappa, \delta)}{y(1,1)} = \max \left\{ \left(\frac{1 + (n-1)\kappa^{-\theta(1-\beta)}}{n} \right), \left(\frac{1 + (n-1)\delta^{-\theta(1-\beta)}}{n} \right) \right\}^{\frac{1}{\theta(1-\beta)}}$$

When multinationals' productivities are uncorrelated across locations ($\rho \to 0$), lower worldwide trade costs and lower worldwide FDI costs are complementary. In contrast, when each multinational has the same productivity in every location, trade and FDI are perfect substitutes. Consider a multinational that has the highest productivity for its variety. If trade costs are lower than FDI costs, it will produce in its home country and export. If FDI costs are lower than trade costs, it will produce in each destination country.

For the case of a symmetric world, the gains from openness (trade and FDI) as measured by the per-capita income on the balanced growth path with costless trade and FDI relative to autarky can be characterized for an arbitrary correlation of multinationals' productivities across locations:

$$\begin{aligned} \frac{y^O}{y^{AUT}} &= n^{\frac{1}{\theta}} \quad \left(\frac{\lambda^{FT}}{\lambda^{AUT}}\right)^{\frac{1}{\theta}} \\ &= \underbrace{n^{\frac{2-\rho}{\theta}}}_{static} \quad \underbrace{n^{\frac{(2-\rho)\beta}{(1-\beta)\theta}}}_{dynamic} \end{aligned}$$

Notice that both, the static and dynamic, gains from openness are larger than those in (14), provided $\rho < 1$.

5.1 Opening to Trade and FDI

Economic miracles are characterized by protracted growth in productivity and per-capita income, and are associated with increases in trade and FDI flows. Our theory features novel mechanisms through which trade and FDI liberalization could result in protracted growth in productivity and per-capita income. In this section we explore quantitatively the dynamic implications of our theory following trade and FDI liberalization. We also assess the relative importance of these flows for the diffusion of technologies.

We consider a world economy that starts with $n-1$ (relatively) open economies and one deviant economy that are on a balanced growth path. We calibrate the trade costs of the $n-1$ open economies so that their trade shares equal 0.50, $\kappa_1 = 2.15$. We set the FDI cost to be 40% higher than the trade costs, $\delta_1 = 3$, to be consistent with the estimates in Ramondo and Rodriguez-Clare (2013). Trade to and from, and operation in, the deviant economy face large trade and FDI costs, $\kappa_n = \delta_n = 100$. We trace the evolution of the stock of ideas and per-capita income as trade and

FDI costs are eliminated, $\kappa_n = \delta_n = 100 \to \kappa'_n = \delta'_n = 1$. We also consider the cases in which only one of the costs are eliminated. In all these examples, we set $\beta = 0.5$, $\rho = 0.5$, and $\psi = 0.1$.[25]

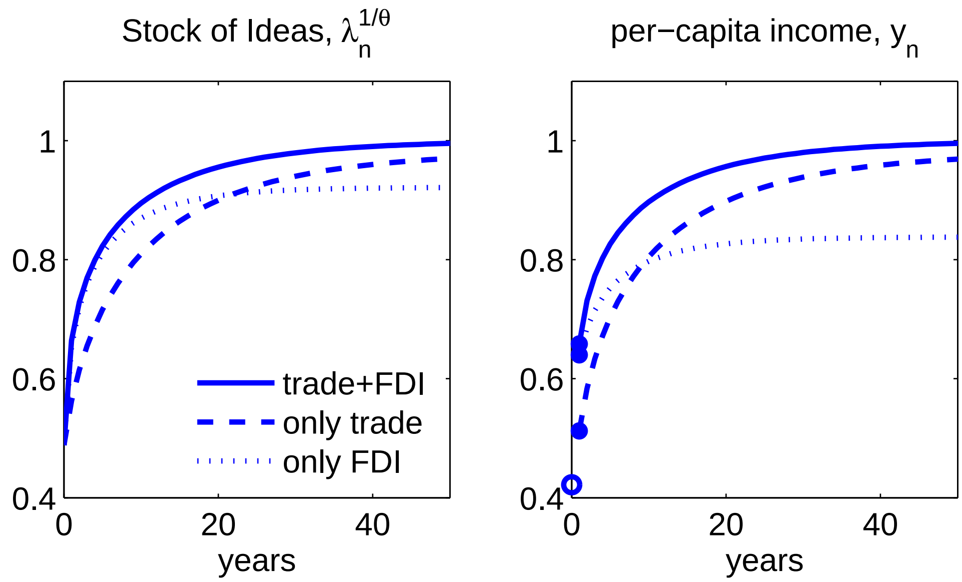

Figure 8: Dynamics after Opening to Trade and/or FDI.

Figure 8 shows the dynamics of the stock of ideas (left panel) and per-capita income (right panel) following the reduction of trade and/or FDI cost with the deviant economy. The y-axis measures the variables relative to their values on a balanced growth path where all countries have the same, low trade and FDI costs, $\kappa_1 = \kappa_n = 2.15$ and $\delta_1 = \delta_n = 3$. The solid line shows the dynamics after trade and FDI costs with the deviant country are reduced, while the dashed and dotted lines correspond to the cases where only trade or FDI costs are reduced, respectively.

Following the opening to trade and FDI, the stock of ideas in the (formerly) deviant country undergoes a sustained period of growth, converging to that of the $n-1$ open economies. This process is substantially slower when only trade costs are reduced, since in this example we consider the case where most of the insights are drawn from domestic producers, $\psi = 0.1$. If only trade costs are reduced, most of the improvements in the quality of insights come from the selection of better domestic producers, which are themselves relatively unproductive to start with. On the contrary, when FDI costs are reduced, the initial growth of the stock of knowledge is much faster, as more productive foreign multinational start producing in the (formerly) deviant country, resulting in a better distribution of insights. In the long run, since trade costs are reduced by more than FDI

[25]The rest of the parameters are calibrated as discussed in footnote 18.

costs, the gains associated with a reduction in the former is larger, albeit these gains materialize at a slower pace.

In the right panel we present the dynamics of per-capita income following a the reduction in trade and/or FDI costs. On impact the gains are larger when FDI cost are reduced, as in this case the more advanced foreign technologies get to be used. This case attains most of the static gains from openness. Over time, the growth of per-capita income reflects the improvement in the technology shown in the left panel.

6 Conclusions

To be written.

A Technology Diffusion

Lemma 5 *Under Assumption 1, there is a* $K < \infty$ *such that for all* z, $\frac{1-H(z)}{z^{-\theta}} \leq K$

Proof. Choose $\delta > 0$ arbitrarily. Since $\lim_{z\to\infty} \frac{1-H(z)}{z^{-\theta}} = 1$, there is a z^* such that $z > z^*$ implies $\frac{1-H(z)}{z^{-\theta}} < 1 + \delta$. For $z < z^*$, we have that $z^\theta \left[1 - H(z)\right] \leq (z^*)^\theta \left[1 - H(z)\right] \leq (z^*)^\theta$. Thus for any z, $\frac{1-H(z)}{z^{-\theta}} \leq K \equiv \max\left\{1 + \delta, (z^*)^\theta\right\}$ ■

Claim 6 *Suppose that Assumption 1 and Assumption 2 hold. Then in the limit as* $m \to \infty$*, the frontier of knowledge evolves as:*

$$\frac{d \ln F_t(q)}{dt} = -\alpha_t q^{-\theta} \int_0^\infty x^{\beta\theta} dG_t(x)$$

Proof. Evaluating the law of motion at $m^{\frac{1}{(1-\beta)\theta}} q$ and using the change of variables $w = m^{-\frac{1}{(1-\beta)\theta}} x$ we get

$$\begin{aligned}
\frac{\partial}{\partial t} \ln \tilde{F}_t\left(m^{\frac{1}{(1-\beta)\theta}} q\right) &= -m\alpha_t \int_0^\infty \left[1 - H\left(\frac{m^{\frac{1}{(1-\beta)\theta}} q}{x^\beta}\right)\right] d\tilde{G}_t(x) \\
&= -m\alpha_t \int_0^\infty \left[1 - H\left(\frac{m^{\frac{1}{(1-\beta)\theta}} q}{\left(m^{\frac{1}{(1-\beta)\theta}} w\right)^\beta}\right)\right] d\tilde{G}_t\left(m^{\frac{1}{(1-\beta)\theta}} w\right)
\end{aligned}$$

From above, we have that $F_t(q) \equiv \tilde{F}_t\left(m^{\frac{1}{(1-\beta)\theta}} q\right)$, and $G_t \equiv \tilde{G}_t\left(m^{\frac{1}{(1-\beta)\theta}} q\right)$ which have the following derivatives

$$\begin{aligned}
G_t'(q) &= m^{\frac{1}{(1-\beta)\theta}} \tilde{G}_t'\left(m^{\frac{1}{(1-\beta)\theta}} q\right) \\
F_t'(q) &= m^{\frac{1}{(1-\beta)\theta}} \tilde{F}_t'\left(m^{\frac{1}{(1-\beta)\theta}} q\right) \\
\frac{\partial F_t(q)}{\partial t} &= \frac{\partial \tilde{F}_t\left(m^{\frac{1}{(1-\beta)\theta}} q\right)}{\partial t}
\end{aligned}$$

The equation becomes

$$\frac{\partial \ln F_t(q)}{\partial t} = -m\alpha_t \int_0^\infty \left[1 - H\left(\frac{m^{\frac{1}{(1-\beta)\theta}} q}{\left(m^{\frac{1}{(1-\beta)\theta}} w\right)^\beta}\right)\right] dG_t(w)$$

This can be rearranged as

$$\frac{\partial \ln F_t(q)}{\partial t} = -\alpha_t q^{-\theta} \int_0^\infty \left[\frac{1 - H\left(m^{1/\theta} q w^{-\beta}\right)}{\left[m^{1/\theta} q w^{-\beta}\right]^{-\theta}} \right] w^{\beta\theta} dG_t(w)$$

We want to take a limit as $m \to \infty$. To do this, we must show that we can take the limit inside the integral. By Lemma 5, there is a $K < \infty$ such that for any z, $\frac{1-H(z)}{z^{-\theta}} \le K$. Further, given the assumptions on the tail of G_t, the integral $\int_0^\infty K w^{\beta\theta} d\tilde{G}_t(w)$ is finite. Thus we can take the limit inside the integral using the dominated convergence theorem to get

$$\frac{\partial \ln F_t(q)}{\partial t} = -\alpha_t q^{-\theta} \int_0^\infty w^{\beta\theta} dG_t(w)$$

■

B Trade

B.1 Equilibrium

This section derives expressions for price indices, trade shares, and market clearing conditions that determine equilibrium wages. The total expenditure in i is X_i. Throughout this section, we maintain that $F_i^{12}(q_1, q_2) = \left[1 + \lambda_i q_1^{-\theta} - \lambda_i q_2^{-\theta}\right] e^{-\lambda_i q^{-\theta}}$.

For a variety $s \in S_{ij}$ (produced in j and exported to i) that is produced with productivity q, the equilibrium price in i is $p_i(s) = \frac{w_j \kappa_{ij}}{q}$, the expenditure on consumption in i is $\left(\frac{p_i(s)}{P_i}\right)^{1-\varepsilon} X_i$, consumption is $\frac{1}{p_i(s)} \left(\frac{p_i(s)}{P_i}\right)^{1-\varepsilon} X_i$, and the labor used in j to produce variety s for i is $\frac{\kappa_{ij}/q_{j1}(s)}{p_i(s)} \left(\frac{p_i(s)}{P_i}\right)^{1-\varepsilon} X_i$

Define $\pi_{ij} \equiv \frac{\lambda_j (w_j \kappa_{ij})^{-\theta}}{\sum_k \lambda_k (w_k \kappa_{ik})^{-\theta}}$. We will eventually show this is the share of i's total expenditure that is spent on goods from j.

We begin with a lemma which will be useful in deriving a number of results.

Lemma 7 *Suppose τ_1 and τ_2 satisfy $\tau_1 < 1$ and $\tau_1 + \tau_2 < 1$. Then*

$$\int_{s \in S_{ij}} q_{j1}(s)^{\tau_1 \theta} p_i(s)^{-\tau_2 \theta} ds = \tilde{B}(\tau_1, \tau_2) \left[\sum_k \lambda_k \left(w_k \kappa_{ik}\right)^{-\theta}\right]^{\tau_2} \pi_{ij} \left(\frac{\lambda_j}{\pi_{ij}}\right)^{\tau_1}$$

where $\tilde{B}(\tau_1,\tau_2) \equiv \left\{1 - \frac{\tau_2}{1-\tau_1} + \frac{\tau_2}{1-\tau_1}\left(\frac{\varepsilon}{\varepsilon-1}\right)^{-\theta(1-\tau_1)}\right\}\Gamma\left(1-\tau_1-\tau_2\right)$

Proof. We begin by defining the measure $\mathcal{F}_{ij}$ to satisfy

$$\mathcal{F}_{ij}(q_1,q_2) = \int_0^{q_2} \prod_{k\neq j} F_k^{12}\left(\frac{w_k\kappa_{ik}x}{w_j\kappa_{ij}}, \frac{w_k\kappa_{ik}x}{w_j\kappa_{ij}}\right) F_j^{12}(dx,x) + \int_{q_2}^{q_1} \prod_{k\neq j} F_k^{12}\left(\frac{w_k\kappa_{ik}q_2}{w_j\kappa_{ij}}, \frac{w_k\kappa_{ik}q_2}{w_j\kappa_{ij}}\right) F_j^{12}(dx,q_2) \tag{15}$$

$\mathcal{F}_{ij}(q_1,q_2)$ is the fraction of varieties that i purchases from j with productivity no greater than q_1 and second best provider of the good to i has marginal cost no smaller than $\frac{w_j\kappa_{ij}}{q_2}$. There are two terms in the sum. The first term integrates over goods where j's lowest-cost producer has productivity no greater than q_2, and the second over goods where j's lowest cost producer has productivity between q_1 and q_2. The corresponding density $\frac{\partial^2}{\partial q_1 \partial q_2}\mathcal{F}_{ij}(q_1,q_2)$ will be useful because it is the measure of firms in j with productivity q that are the lowest cost providers to i and for which the next-lowest-cost provider has marginal cost $w_j\kappa_{ij}/q_2$.

We first show that

$$\mathcal{F}_{ij}(q_1,q_2) = \left[\pi_{ij} + \lambda_j\left(q_2^{-\theta} - q_1^{-\theta}\right)\right] e^{-\frac{1}{\pi_{ij}}\lambda_j q_2^{-\theta}}$$

The first term of equation (15) can be written as

$$\begin{aligned}
\int_0^{q_2} \prod_{k\neq j} F_k^{12}\left(\frac{w_k\kappa_{ik}x}{w_j\kappa_{ij}}, \frac{w_k\kappa_{ik}x}{w_j\kappa_{ij}}\right) F_j^{12}(dx,x) &= \int_0^{q_2} e^{-\sum_{k\neq j}\lambda_k\left(\frac{w_k\kappa_{ik}}{w_j\kappa_{ij}}\right)^{-\theta} x^{-\theta}} \theta\lambda_j x^{-\theta-1} e^{-\lambda_j x^{-\theta}} dx \\
&= \frac{\lambda_j\left(w_j\kappa_{ij}\right)^{-\theta}}{\sum_k \lambda_k\left(w_k\kappa_{ik}\right)^{-\theta}} e^{-\sum_k \lambda_k\left(\frac{w_k\kappa_{ik}}{w_j\kappa_{ij}}\right)^{-\theta} q_2^{-\theta}} \\
&= \pi_{ij} e^{-\frac{\lambda_j}{\pi_{ij}} q_2^{-\theta}}
\end{aligned}$$

The second term is

$$\begin{aligned}
\int_{q_2}^{q_1} \prod_{k\neq i} F_k^{12}\left(\frac{w_k\kappa_{ik}q_2}{w_j\kappa_{ij}}, \frac{w_k\kappa_{ik}q_2}{w_j\kappa_{ij}}\right) F_j^{12}\left(dx,q_2\right) &= e^{-\sum_{k\neq j}\lambda_k\left(\frac{w_k\kappa_{ik}}{w_j\kappa_{ij}}\right)^{-\theta} q_2^{-\theta}} \int_{q_2}^{q_1} \theta\lambda_j x^{-\theta-1} e^{-\lambda_j q_2^{-\theta}} dx \\
&= e^{-\sum_k \lambda_k\left(\frac{w_k\kappa_{ik}}{w_j\kappa_{ij}}\right)^{-\theta} q_2^{-\theta}} \lambda_j\left[q_2^{-\theta} - q_1^{-\theta}\right] \\
&= e^{-\frac{\lambda_j}{\pi_{ij}} q_2^{-\theta}} \lambda_j\left[q_2^{-\theta} - q_1^{-\theta}\right]
\end{aligned}$$

Together, these give the expression for $\mathcal{F}_{ij}$, so the joint density is

$$\frac{\partial^2}{\partial q_1 \partial q_2}\mathcal{F}_{ij}(q_1, q_2) = \frac{1}{\pi_{ij}}\left(\theta\lambda_j q_1^{-\theta-1}\right)\left(\theta\lambda_j q_2^{-\theta-1}\right)e^{-\frac{1}{\pi_{ij}}\lambda_j q_2^{-\theta}}$$

We next turn to the integral $\int_{s\in S_{ij}} q_{j1}(s)^{\theta\tau_1} p_i(s)^{-\theta\tau_2} ds$. Since the price of good s is set at either a markup of $\frac{\varepsilon}{\varepsilon-1}$ over marginal cost or at the cost of the next lowest cost provider, this integral equals

$$\begin{aligned}
& \int_0^\infty \int_{q_2}^\infty q_1^{\theta\tau_1} \min\left\{\frac{w_j\kappa_{ij}}{q_2}, \frac{\varepsilon}{\varepsilon-1}\frac{w_j\kappa_{ij}}{q_1}\right\}^{-\theta\tau_2} \frac{\partial^2\mathcal{F}_{ij}\left(q_1, q_2\right)}{\partial q_1 \partial q_2} dq_1 dq_2 \\
= & \int_0^\infty \int_{q_2}^\infty q_1^{\theta\tau_1} \min\left\{\frac{w_j\kappa_{ij}}{q_2}, \frac{\varepsilon}{\varepsilon-1}\frac{w_j\kappa_{ij}}{q_1}\right\}^{-\theta\tau_2} \frac{1}{\pi_{ij}}\left(\theta\lambda_j q_1^{-\theta-1}\right)\left(\theta\lambda_j q_2^{-\theta-1}\right)e^{-\frac{1}{\pi_{ij}}\lambda_j q_2^{-\theta}} dq_1 dq_2
\end{aligned}$$

Using the change of variables $x_1 = \frac{\lambda_j}{\pi_{ij}} q_1^{-\theta}$ and $x_2 = \frac{\lambda_j}{\pi_{ij}} q_2^{-\theta}$, this becomes

$$\left(w_j\kappa_{ij}\right)^{-\theta\tau_2}\pi_{ij}\left(\frac{\lambda_j}{\pi_{ij}}\right)^{\tau_1+\tau_2}\int_0^\infty\int_0^{x_2} x_1^{-\tau_1}\min\left\{x_2, \left(\frac{\varepsilon}{\varepsilon-1}\right)^\theta x_1\right\}^{-\tau_2} e^{-x_2} dx_1 dx_2$$

Define $\tilde{B}(\tau_1, \tau_2) \equiv \int_0^\infty \int_0^{x_2} x_1^{-\tau_1} \min\left\{x_2, \left(\frac{\varepsilon}{\varepsilon-1}\right)^\theta x_1\right\}^{-\tau_2} e^{-x_2} dx_1 dx_2$, so that the integral is

$$\int_{s\in S_{ij}} q_{j1}(s)^{\theta\tau_1} p_i(s)^{-\theta\tau_2} ds = B(\tau_1, \tau_2)\left(w_j\kappa_{ij}\right)^{-\theta\tau_2}\pi_{ij}\left(\frac{\lambda_j}{\pi_{ij}}\right)^{\tau_1+\tau_2}$$

Using $\pi_{ij} = \frac{\lambda_j(w_j\kappa_{ij})^{-\theta}}{\sum_k \lambda_k(w_k\kappa_{ik})^{-\theta}}$, we have $\left(w_j\kappa_{ij}\right)^{-\theta\tau_2}\left(\frac{\lambda_j}{\pi_{ij}}\right)^{\tau_2} = \left[\sum_k \lambda_k\left(w_k\kappa_{ik}\right)^{-\theta}\right]^{\tau_2}$. Finally we complete the proof by providing an expression for $\tilde{B}(\tau_1,\tau_2)$:

$$
\begin{aligned}
\tilde{B}(\tau_1,\tau_2) &= \int_0^\infty \int_0^{x_2} x_1^{-\tau_1} \min\left\{x_2, \left(\frac{\varepsilon}{\varepsilon-1}\right)^\theta x_1\right\}^{-\tau_2} e^{-x_2} dx_1 dx_2 \\
&= \int_0^\infty \int_{\left(\frac{\varepsilon}{\varepsilon-1}\right)^{-\theta} x_2}^{x_2} x_1^{-\tau_1} x_2^{-\tau_2} e^{-x_2} dx_1 dx_2 + \int_0^\infty \int_0^{\left(\frac{\varepsilon}{\varepsilon-1}\right)^{-\theta} x_2} x_1^{-\tau_1} \left\{\left(\frac{\varepsilon}{\varepsilon-1}\right)^\theta x_1\right\}^{-\tau_2} e^{-x_2} dx_1 dx_2 \\
&= \int_0^\infty \frac{x_2^{1-\tau_1} - \left(\left(\frac{\varepsilon}{\varepsilon-1}\right)^{-\theta} x_2\right)^{1-\tau_1}}{1-\tau_1} x_2^{-\tau_2} e^{-x_2} dx_2 + \left(\frac{\varepsilon}{\varepsilon-1}\right)^{-\theta\tau_2} \int_0^\infty \frac{\left(\left(\frac{\varepsilon}{\varepsilon-1}\right)^{-\theta} x_2\right)^{1-\tau_1-\tau_2}}{1-\tau_1-\tau_2} e^{-x_2} dx_2 \\
&= \frac{1-\left(\frac{\varepsilon}{\varepsilon-1}\right)^{-\theta(1-\tau_1)}}{1-\tau_1} \int_0^\infty x_2^{1-\tau_1-\tau_2} e^{-x_2} dx_2 + \frac{\left(\frac{\varepsilon}{\varepsilon-1}\right)^{-\theta(1-\tau_1)}}{1-\tau_1-\tau_2} \int_0^\infty x_2^{1-\tau_1-\tau_2} e^{-x_2} dx_2 \\
&= \left\{\frac{1-\left(\frac{\varepsilon}{\varepsilon-1}\right)^{-\theta(1-\tau_1)}}{1-\tau_1} + \frac{\left(\frac{\varepsilon}{\varepsilon-1}\right)^{-\theta(1-\tau_1)}}{1-\tau_1-\tau_2}\right\} \Gamma(2-\tau_1-\tau_2) \\
&= \left\{1 - \frac{\tau_2}{1-\tau_1} + \frac{\tau_2}{1-\tau_1}\left(\frac{\varepsilon}{\varepsilon-1}\right)^{-\theta(1-\tau_1)}\right\} \Gamma(1-\tau_1-\tau_2)
\end{aligned}
$$

where the final equality uses the fact that for any x, $\Gamma(x+1) = x\Gamma(x)$. ■

We first use this lemma to provide expressions for the price index in i and the share of i's expenditure on goods from j.

Claim 8 *The price index for i satisfies*

$$P_i = \tilde{B}\left(0, \frac{\varepsilon-1}{\theta}\right)^{\frac{1}{1-\varepsilon-1}} \left[\sum_k \lambda_k (w_k \kappa_{ik})^{-\theta}\right]^{-\frac{1}{\theta}}$$

where $B \equiv \left\{\left(1-\frac{\varepsilon-1}{\theta}\right)\left(1-\left(\frac{\varepsilon}{\varepsilon-1}\right)^{-\theta}\right) + \left(\frac{\varepsilon}{\varepsilon-1}\right)^{-\theta}\right\}\Gamma\left(1-\frac{\varepsilon-1}{\theta}\right)$. $\pi_{ij} = \frac{\lambda_j (w_j\kappa_{ij})^{-\theta}}{\sum_k \lambda_k (w_k \kappa_{ik})^{-\theta}}$ *is the share of i's expenditure on goods from j.*

Proof. The price aggregate of goods provided to i by j is $\int_{s\in S_{ij}} p_i(s)^{1-\varepsilon} ds$. Using Lemma 7, this equals

$$\int_{s\in S_{ij}} p_i(s)^{1-\varepsilon} ds = \tilde{B}\left(0, \frac{\varepsilon-1}{\theta}\right) \left[\sum_k \lambda_k (w_k \kappa_{ik})^{-\theta}\right]^{\frac{\varepsilon-1}{\theta}} \pi_{ij}$$

The price index for i therefore satisfies

$$P_i^{1-\varepsilon} = \sum_j \int_{s\in S_{ij}} p_i(s)^{1-\varepsilon} ds = \tilde{B}\left(0, \frac{\varepsilon-1}{\theta}\right)\left[\sum_k \lambda_k \left(w_k \kappa_{ik}\right)^{-\theta}\right]^{\frac{\varepsilon-1}{\theta}}$$

and i's expenditure share on goods from j is

$$\frac{\int_{s\in S_{ij}} p_i\left(s\right)^{1-\varepsilon} ds}{P_i^{1-\varepsilon}} = \pi_{ij}$$

■

We next turn to the market clearing conditions.

Claim 9 *Country j's expenditure on labor is* $\frac{\theta}{\theta+1}\sum_i \pi_{ij} X_i$.

Proof. i's consumption of good s is $p_i(s)^{-\varepsilon}\frac{X_i}{P_i^{1-\varepsilon}}$. If j is the lowest-cost provider to i, then j's expenditure on labor per unit delivered is $w_j\frac{\kappa_{ij}}{q_{j1}(s)}$. The total expenditure on labor in j to produce goods for i is then $\int_{s\in S_{ij}}\frac{w_j\kappa_{ij}}{q_{j1}(s)}p_i(s)^{-\varepsilon}\frac{X_i}{P_i^{1-\varepsilon}}ds$. Using Lemma 7, the total expenditure on labor in j is thus

$$\begin{aligned}\sum_i \int_{s\in S_{ij}} \frac{w_j\kappa_{ij}}{q_{j1}(s)} p_i(s)^{-\varepsilon}\frac{X_i}{P_i^{1-\varepsilon}}ds &= \sum_i w_j\kappa_{ij}\frac{X_i}{P_i^{1-\varepsilon}}\int_{s\in S_{ij}} q_{j1}\left(s\right)^{-1} p_i\left(s\right)^{-\varepsilon} ds \\ &= \tilde{B}\left(-\frac{1}{\theta},\frac{\varepsilon}{\theta}\right)\sum_i w_j\kappa_{ij}\frac{X_i}{P_i^{1-\varepsilon}}\left[\sum_k \lambda_k\left(w_k\kappa_{ik}\right)^{-\theta}\right]^{\frac{\varepsilon}{\theta}}\pi_{ij}\left(\frac{\lambda_j}{\pi_{ij}}\right)^{-\frac{1}{\theta}}\end{aligned}$$

The result follows from $\tilde{B}\left(-\frac{1}{\theta},\frac{\varepsilon}{\theta}\right) = \frac{\theta}{\theta+1}\tilde{B}\left(0,\frac{\varepsilon-1}{\theta}\right)$ and $\frac{w_j\kappa_{ij}}{P_i^{1-\varepsilon}}\left[\sum_k \lambda_k\left(w_k\kappa_{ik}\right)^{-\theta}\right]^{\frac{\varepsilon}{\theta}}\left(\frac{\lambda_j}{\pi_{ij}}\right)^{-\frac{1}{\theta}} = \tilde{B}\left(0,\frac{\varepsilon-1}{\theta}\right)^{-1}$.

■

C Source Distributions

This appendix derives expressions for the source distributions under various specifications. We begin by describing learning from sellers.

C.1 Learning from Sellers

Here we characterize the learning process when insights are drawn from sellers in proportion to the expenditure on each seller's good. Consider a variety that can be produced in j at productivity q. Since the share of i's expenditure on good s is $(p_i(s)/P_i)^{1-\varepsilon}$, the source distribution is

$$G_i(q) = \sum_j \int_{\{s\in S_{ij}|q_{j1}(s)\leq q\}} (p_i(s)/P_i)^{1-\varepsilon}\, ds$$

The change in i's stock of knowledge depends on

$$\int_0^\infty q^{\beta\theta} dG_i(q) = \sum_j \int_{s\in S_{ij}} q_{j1}(s)^{\beta\theta} (p_i(s)/P_i)^{1-\varepsilon}\, ds$$

Using Lemma 7, this is

$$\begin{aligned}\int_0^\infty q^{\beta\theta} dG_i(q) &= \sum_j \frac{1}{P_i^{1-\varepsilon}} \tilde{B}\left(\beta, \frac{\varepsilon-1}{\theta}\right) \left[\sum_k \lambda_k (w_k\kappa_{ik})^{-\theta}\right]^{\frac{\varepsilon-1}{\theta}} \pi_{ij} \left(\frac{\lambda_j}{\pi_{ij}}\right)^\beta \\ &= \frac{\tilde{B}\left(\beta, \frac{\varepsilon-1}{\theta}\right)}{\tilde{B}\left(0, \frac{\varepsilon-1}{\theta}\right)} \sum_j \pi_{ij} \left(\frac{\lambda_j}{\pi_{ij}}\right)^\beta \qquad (16)\end{aligned}$$

C.1.1 Alternative Weights of Sellers

Here we explore two alternative processes by which individuals can learn from sellers. In the first case, individuals are equally likely to learn from all active sellers, independently of how much of the seller's variety they consume. In the second case, insights are drawn from sellers in proportion to consumption of each sellers' goods. In each case, the speed of learning is the same as our baseline (equation (16)) up to a constant.

Learning from All Active Sellers Equally

If producers are equally likely to learn from all active sellers, the source distribution is

$$G_i(q) = \frac{\sum_j \int_{\{s\in S_{ij}|q_{j1}(s)\leq q\}} ds}{\sum_j \int_{s\in S_{ij}} ds}$$

The change in i's stock of knowledge depends on $\int_0^\infty q^{\beta\theta} dG_i(q) = \frac{\sum_j \int_{s\in S_{ij}} q_{j1}(s)^{\beta\theta} ds}{\sum_j \int_{s\in S_{ij}} ds}$. Using Lemma 7, this is

$$\int_0^\infty q^{\beta\theta} dG_i(q) \quad = \quad \frac{\tilde{B}(\beta,0)}{\tilde{B}(0,0)} \sum_j \pi_{ij} \left(\frac{\lambda_j}{\pi_{ij}}\right)^\beta = \Gamma(1-\beta) \sum_j \pi_{ij} \left(\frac{\lambda_j}{\pi_{ij}}\right)^\beta$$

Learning from Sellers in Proportion to Consumption

i's consumption of goods s is $c_i(s) = p_i(s)^{-\varepsilon} \frac{X_i}{P_i^{1-\varepsilon}}$. If producers learn in proportion to consumption, then the source distribution is

$$G_i(q) = \frac{\sum_j \int_{\{s \in S_{ij} | q_{j1}(s) \leq q\}} p_i(s)^{-\varepsilon} \frac{X_i}{P_i^{1-\varepsilon}} ds}{\sum_j \int_{\{s \in S_{ij}\}} p_i(s)^{-\varepsilon} \frac{X_i}{P_i^{1-\varepsilon}} ds}$$

The change in i's stock of knowledge depends on

$$\int_0^\infty q^{\beta\theta} dG_i(q) = \frac{\sum_j \int_{s \in S_{ij}} q_{j1}(s)^{\beta\theta} p_i(s)^{-\varepsilon} ds}{\sum_j \int_{s \in S_{ij}} p_i(s)^{-\varepsilon} ds}$$

Using Lemma 7, this is

$$\begin{aligned} \int_0^\infty q^{\beta\theta} dG_i(q) \quad &= \quad \frac{\sum_j \tilde{B}\left(\beta, \frac{\varepsilon}{\theta}\right) \left[\sum_k \lambda_k \left(w_k \kappa_{ik}\right)^{-\theta}\right]^{\frac{\varepsilon}{\theta}} \pi_{ij} \left(\frac{\lambda_j}{\pi_{ij}}\right)^\beta}{\sum_j \tilde{B}\left(0, \frac{\varepsilon}{\theta}\right) \left[\sum_k \lambda_k \left(w_k \kappa_{ik}\right)^{-\theta}\right]^{\frac{\varepsilon}{\theta}} \pi_{ij}} \\ &= \quad \frac{\tilde{B}\left(\beta, \frac{\varepsilon}{\theta}\right)}{\tilde{B}\left(0, \frac{\varepsilon}{\theta}\right)} \sum_j \pi_{ij} \left(\frac{\lambda_j}{\pi_{ij}}\right)^\beta \end{aligned}$$

C.2 Learning from Producers

Here we characterize the learning process when insights are drawn from domestic producers in proportion to labor used in production. For each $s \in S_{ij}$, the fraction of j's labor used to produce the good is $\frac{1}{L_j} \frac{\kappa_{ij}}{q_{j1}(s)} c_i(s)$ with $c_i(s) = p_i(s)^{-\varepsilon} \frac{X_i}{P_i^{1-\varepsilon}}$. Summing over all destinations, the source distribution would then be

$$G_j(q) = \sum_i \int_{s \in S_{ij} | q_{j1}(s) \leq q} \frac{1}{L_j} \frac{\kappa_{ij}}{q_{j1}(s)} p_i(s)^{-\varepsilon} \frac{X_i}{P_i^{1-\varepsilon}} ds$$

The change in j's stock of knowledge depends on

$$\int_0^\infty q^{\beta\theta} dG_j(q) = \sum_i \int_{s \in S_{ij}} q^{\beta\theta} \frac{1}{L_j} \frac{\kappa_{ij}}{q_{j1}(s)} p_i(s)^{-\varepsilon} \frac{X_i}{P_i^{1-\varepsilon}} ds$$

Using Lemma 7, this is

$$\int_0^\infty q^{\beta\theta} dG_j(q) \quad = \quad \sum_j \frac{\kappa_{ij}}{L_j} \frac{X_i}{P_i^{1-\varepsilon}} \tilde{B}\left(\beta - \frac{1}{\theta}, \frac{\varepsilon}{\theta}\right) \left[\sum_k \lambda_k \left(w_k \kappa_{ik}\right)^{-\theta}\right]^{\frac{\varepsilon}{\theta}} \pi_{ij} \left(\frac{\lambda_j}{\pi_{ij}}\right)^{\beta - \frac{1}{\theta}}$$

Using the expressions for P_i and π_{ij} from above, this becomes

$$\int_0^\infty q^{\beta\theta} dG_j(q) = \frac{\tilde{B}\left(\beta - \frac{1}{\theta}, \frac{\varepsilon}{\theta}\right)}{\tilde{B}\left(0, \frac{\varepsilon-1}{\theta}\right)} \frac{1}{w_j L_j} \sum_j \pi_{ij} X_i \left(\frac{\lambda_j}{\pi_{ij}}\right)^\beta$$

C.2.1 Alternative Weights of Producers

Here we briefly describe the alternative learning process in which insights are equally likely to be dawn from all active domestic producers. We consider only the case in which trade costs satisfy the triangle inequality $\kappa_{jk} < \kappa_{ji}\kappa_{ik}, \forall i, j, k$ such that $i \neq j \neq k \neq i$. We will show that, in this case, all producers that export also sell domestically. This greatly simplifies characterizing the learning process.

Towards a contradiction, suppose there is a variety s such that i exports to j and k exports to i. This means that $\frac{w_i \kappa_{ji}}{q_i(s)} \leq \frac{w_k \kappa_{jk}}{q_k(s)}$ and $\frac{w_k \kappa_{ik}}{q_k(s)} \leq \frac{w_i \kappa_{ii}}{q_i(s)}$. Since $\kappa_{ii} = 1$, these imply that $\kappa_{ji}\kappa_{ik} \leq \kappa_{jk}$, a violation of the triangle inequality and thus a contradiction.

In this case, the source distribution is $G_i(q) = \frac{\int_{s \in S_{ii} | q_{i1} \leq q} ds}{\int_{s \in S_{ii}} ds}$. The change in i's stock of knowledge depends on

$$\int_0^\infty q^{\beta\theta} dG_i(q) = \frac{\int_{s \in S_{ii}} q^{\beta\theta} ds}{\int_{s \in S_{ii}} ds}$$

Using Lemma 7, this is

$$\int_0^\infty q^{\beta\theta} dG_i(q) = \frac{\tilde{B}(\beta, 0) \pi_{ij} \left(\frac{\lambda_j}{\pi_{ij}}\right)^\beta}{\tilde{B}(0,0) \pi_{ij}} = \Gamma(1 - \beta) \pi_{ij} \left(\frac{\lambda_j}{\pi_{ij}}\right)^\beta$$

D Simple Examples

D.1 Symmetric Countries

If countries are symmetric, there are two possible values of π_{ij}:

$$\begin{aligned} \pi_{ii} &= \frac{1}{1+(n-1)\kappa^{-\theta}} \\ \pi_{ij} &= \frac{\kappa^{-\theta}}{1+(n-1)\kappa^{-\theta}}, \qquad i \neq j \end{aligned}$$

Normalizing the wage to unity, the price level is

$$P = B\lambda^{-\frac{1}{\theta}} \left(1+(n-1)\kappa^{-\theta}\right)^{-\frac{1}{\theta}}$$

The de-trended scale parameter on a balance growth path is

$$\hat{\lambda}(\kappa) = \left[(1-\beta)\frac{\alpha}{\gamma}\frac{\Gamma\left(1-\beta-\frac{\varepsilon-1}{\theta}\right)}{\Gamma\left(1-\frac{\varepsilon-1}{\theta}\right)}\frac{1+(n-1)\kappa^{-\theta(1-\beta)}}{\left(1+(n-1)\kappa^{-\theta}\right)^{1-\beta}}\right]^{\frac{1}{1-\beta}}$$

Using this expression, per-capita income, $y_i = w_i/P_i$, is

$$\begin{aligned} y(\kappa) &= \Gamma\left(1-\frac{\varepsilon-1}{\theta}\right)^{-\frac{1}{1-\varepsilon}} \lambda^{\frac{1}{\theta}} \left(1+(n-1)\kappa^{-\theta}\right)^{\frac{1}{\theta}} \\ &= \Gamma\left(1-\frac{\varepsilon-1}{\theta}\right)^{-\frac{1}{1-\varepsilon}} \left[(1-\beta)\frac{\alpha}{\gamma}\frac{\Gamma\left(1-\beta-\frac{\varepsilon-1}{\theta}\right)}{\Gamma\left(1-\frac{\varepsilon-1}{\theta}\right)}\right]^{\frac{1}{\theta(1-\beta)}} \left(1+(n-1)\kappa^{-\theta(1-\beta)}\right)^{\frac{1}{\theta(1-\beta)}} \end{aligned}$$

The de-trended stock of knowledge and per-capita income relative to costless trade are

$$\begin{aligned} \frac{\hat{\lambda}(\kappa)}{\hat{\lambda}(1)} &= \left[\frac{1+(n-1)\kappa^{-\theta(1-\beta)}}{\left(1+(n-1)\kappa^{-\theta}\right)^{1-\beta}}\right]^{\frac{1}{1-\beta}} n^{-\frac{\beta}{1-\beta}} \\ \frac{y(\kappa)}{y(1)} &= \left(\frac{1+(n-1)\kappa^{-\theta}}{n}\right)^{\frac{1}{\theta}} \left(\frac{\lambda(\kappa)}{\lambda(1)}\right)^{\frac{1}{\theta}} \end{aligned}$$

In particular, per-capita income in autarky relative to the case with costless trade

$$\frac{y(\infty)}{y(1)} = \underbrace{n^{-\frac{1}{\theta}}}_{static} \underbrace{n^{-\frac{\beta}{\theta(1-\beta)}}}_{dynamic}$$

D.2 A Small Open Economy

Consider a small open economy. The economy is small in the sense that actions in the economy have no impact on other countries' expenditures, price levels, wages, or stocks of knowledge.

D.2.1 Steady State Gains from Trade

D.2.2 Speed of Convergence

We use the notation $\check{x}$ to denote log-deviation from of x from its steady state (or BGP) value. To derive the speed of convergence, we want expressions for how the trade shares and wages change over time. The trade shares and market clearing condition for i are

$$\begin{aligned}
\pi_{ij} &= \frac{\lambda_j \left(w_j \kappa_{ij}\right)^{-\theta}}{\sum_{k=1}^{n} \lambda_k \left(w_k \kappa_{ik}\right)^{-\theta}} \\
r_{ji} &= \frac{X_j \pi_{ji}}{w_j L_j} \\
w_i L_i &= \sum_j X_j \pi_{ji}
\end{aligned}$$

For a small open economy, we have

$$\begin{aligned}
& \check{\pi}_{ii} = \left(1 - \pi_{ii}\right) \left[\check{\lambda}_i - \theta \check{w}_i\right] \\
j \neq i \quad : \quad & \check{\pi}_{ij} = -\pi_{ii} \left[\check{\lambda}_i - \theta \check{w}_i\right] \\
i \neq j \quad : \quad & \check{\pi}_{ji} = \check{\lambda}_i - \theta \check{w}_i \\
j \neq i \quad : \quad & \check{r}_{ji} = \check{\pi}_{ji} - \check{w}_i = \left[\check{\lambda}_i - \theta \check{w}_i\right] - \check{w}_i
\end{aligned}$$

The change in the wage can be found from linearizing the labor market clearing condition:

$$\begin{aligned}
\check{w}_i &= r_{ii} \left[\check{w}_i + \left(1 - \pi_{ii}\right) \left(\check{\lambda}_i - \theta \check{w}_i\right)\right] + \sum_{j \neq i} r_{ji} \left[\check{\lambda}_i - \theta \check{w}_i\right] \\
\check{w}_i &= \pi_{ii} \left[\check{w}_i - \pi_{ii} \left(\check{\lambda}_i - \theta \check{w}_i\right)\right] + \left[\check{\lambda}_i - \theta \check{w}_i\right] \\
\left(1 - \pi_{ii}\right) \check{w}_i &= \left(1 - \pi_{ii}\right)^2 \left(\check{\lambda}_i - \theta \check{w}_i\right) \\
\check{w}_i &= \left(1 + \pi_{ii}\right) \left(\check{\lambda}_i - \theta \check{w}_i\right)
\end{aligned}$$

This last equation can be expressed in two ways:

$$\check{\lambda}_i - \theta\check{w}_i = \frac{\check{\lambda}_i}{1+\theta(1+\pi_{ii})}$$

$$\check{w}_i = \frac{(1+\pi_{ii})}{1+\theta(1+\pi_{ii})}\check{\lambda}_i$$

Plugging these back into the shares, we have

$$\begin{aligned}
& \check{\pi}_{ii} = \frac{(1-\pi_{ii})}{1+\theta\,(1+\pi_{ii})}\check{\lambda}_i \\
j \neq i \quad : \quad & \check{\pi}_{ij} = -\frac{\pi_{ii}}{1+\theta(1+\pi_{ii})}\check{\lambda}_i \\
i \neq j \quad : \quad & \check{\pi}_{ji} = \frac{1}{1+\theta\,(1+\pi_{ii})}\check{\lambda}_i \\
j \neq i \quad : \quad & \check{r}_{ji} = \check{\pi}_{ji} - \check{w}_i = \frac{1}{1+\theta\,(1+\pi_{ii})} - \frac{(1+\pi_{ii})}{1+\theta\,(1+\pi_{ii})}\check{\lambda}_i = \frac{-\pi_{ii}}{1+\theta\,(1+\pi_{ii})}\check{\lambda}_i
\end{aligned}$$

We now proceed to characterizing transition dynamics for the stock of knowledge, $\check{\lambda}_i$.

Learning from Sellers Let $\Omega^S_{ij} \equiv \frac{\pi_{ij}^{1-\beta}\hat{\lambda}_j^{\beta}}{\sum_k \pi_{ik}^{1-\beta}\hat{\lambda}_k^{\beta}}$. The change in the the deviation of i's stock of knowledge from the BGP is

$$\frac{\partial\check{\lambda}_i}{\partial t} = \frac{1}{\hat{\lambda}_i}\frac{\partial\hat{\lambda}_i}{\partial t} = \frac{B^S\hat{\alpha}_i}{\hat{\lambda}_i}\sum_j \pi_{ij}^{1-\beta}\hat{\lambda}_j^{\beta} - \frac{\gamma}{1-\beta}$$

Log-linearizing around the steady state gives

$$\begin{aligned}
\frac{\partial\check{\lambda}_i}{\partial t} &\approx \frac{B^S\hat{\alpha}_i}{\hat{\lambda}_i}\sum_j \pi_{ij}^{1-\beta}\hat{\lambda}_j^{\beta}\left[(1-\beta)\check{\pi}_{ij} + \beta\check{\lambda}_j - \check{\lambda}_i\right] \\
&= \frac{\gamma}{1-\beta}\frac{\sum_j \pi_{ij}^{1-\beta}\hat{\lambda}_j^{\beta}\left[(1-\beta)\check{\pi}_{ij} + \beta\check{\lambda}_j - \check{\lambda}_i\right]}{\sum_j \pi_{ij}^{1-\beta}\hat{\lambda}_j^{\beta}} \\
&= \frac{\gamma}{1-\beta}\sum_j \Omega^S_{ij}\left[(1-\beta)\,\check{\pi}_{ij} + \beta\check{\lambda}_j - \check{\lambda}_i\right] \\
\frac{1-\beta}{\gamma}\frac{\partial\check{\lambda}_i}{\partial t} &= \sum_j \Omega^S_{ij}\left[(1-\beta)\,\check{\pi}_{ij} + \beta\check{\lambda}_j\right] - \check{\lambda}_i
\end{aligned}$$

For a small open economy, we have $\check{\lambda}_j = 0$, $\check{\pi}_{ii} = \frac{(1-\pi_{ii})\check{\lambda}_i}{1+\theta(1+\pi_{ii})}$, and $\check{\pi}_{ij} = \frac{-\pi_{ii}\check{\lambda}_i}{1+\theta(1+\pi_{ii})}$ for $j \neq i$. The law of motion can be written as

$$\begin{aligned}
\frac{1-\beta}{\gamma}\frac{\partial\check{\lambda}_i}{\partial t} &= \Omega_{ii}^S\left[(1-\beta)\check{\pi}_{ii} + \beta\check{\lambda}_i\right] + (1-\beta)\sum_{j\neq i}\Omega_{ij}^S\check{\pi}_{ij} - \check{\lambda}_i \\
&= \Omega_{ii}^S\left[(1-\beta)\frac{(1-\pi_{ii})\check{\lambda}_i}{1+\theta(1+\pi_{ii})} + \beta\check{\lambda}_i\right] + (1-\beta)\sum_{j\neq i}\Omega_{ij}^S\frac{-\pi_{ii}\check{\lambda}_i}{1+\theta(1+\pi_{ii})} - \check{\lambda}_i \\
&= \left\{\Omega_{ii}^S\left[(1-\beta)\frac{(1-\pi_{ii})}{1+\theta(1+\pi_{ii})} + \beta\right] + (1-\beta)(1-\Omega_{ii}^S)\frac{-\pi_{ii}}{1+\theta(1+\pi_{ii})} - 1\right\}\check{\lambda}_i \\
&= -\left\{1-(1-\beta)\frac{\Omega_{ii}^S - \pi_{ii}}{1+\theta(1+\pi_{ii})} - \beta\Omega_{ii}^S\right\}\check{\lambda}_i \\
&= -\left\{(1-\beta)-(1-\beta)\frac{\Omega_{ii}^S - \pi_{ii}}{1+\theta(1+\pi_{ii})} + \beta\left(1-\Omega_{ii}^S\right)\right\}\check{\lambda}_i \\
\frac{\partial\check{\lambda}_i}{\partial t} &= -\gamma\left\{1-\frac{\Omega_{ii}^S - \pi_{ii}}{1+\theta(1+\pi_{ii})} + \frac{\beta}{1-\beta}\left(1-\Omega_{ii}^S\right)\right\}\check{\lambda}_i
\end{aligned}$$

Finally, we can use this to get at the speed of convergence for real income:

$$\begin{aligned}
\check{w}_i - \check{P}_i &= \frac{1}{\theta}\left[\check{\lambda}_i - \check{\pi}_{ii}\right] = \frac{1}{\theta}\left[1-\frac{(1-\pi_{ii})}{1+\theta(1+\pi_{ii})}\right]\check{\lambda}_i = A\check{\lambda}_i \\
\frac{d}{dt}\left[\check{w}_i - \check{P}_i\right] &= A\frac{d\check{\lambda}_i}{dt} = -\gamma\left\{1-\frac{\Omega_{ii}^S - \pi_{ii}}{1+\theta(1+\pi_{ii})} + \frac{\beta}{1-\beta}\left(1-\Omega_{ii}^S\right)\right\}A\check{\lambda}_i \\
&= -\gamma\left\{1-\frac{\Omega_{ii}^S - \pi_{ii}}{1+\theta(1+\pi_{ii})} + \frac{\beta}{1-\beta}\left(1-\Omega_{ii}^S\right)\right\}\left[\check{w}_i - \check{P}_i\right]
\end{aligned}$$

Learning from Producers Let $\Omega_{ij}^P \equiv \frac{r_{ji}\left(\hat{\lambda}_i/\pi_{ji}\right)^\beta}{\sum_k r_{ki}\left(\hat{\lambda}_i/\pi_{ki}\right)^\beta}$. The change in the the deviation of i's stock of knowledge from the BGP is

$$\frac{\partial\check{\lambda}_i}{\partial t} = \frac{1}{\hat{\lambda}_i}\frac{\partial\hat{\lambda}_i}{\partial t} = \frac{B^P\hat{\alpha}_i}{\hat{\lambda}_i}\sum_j r_{ji}\left(\hat{\lambda}_i/\pi_{ji}\right)^\beta - \frac{\gamma}{1-\beta}$$

Log-linearizing around the steady state gives

$$\begin{aligned}
\frac{\partial \check{\lambda}_i}{\partial t} &\approx \frac{B^P \hat{\alpha}_i}{\hat{\lambda}_i} \sum_j r_{ji} \left(\hat{\lambda}_i / \pi_{ji}\right)^{\beta} \left[\check{r}_{ji} - \beta \check{\pi}_{ij} - (1-\beta)\check{\lambda}_i\right] \\
&= \frac{\gamma}{1-\beta} \sum_j \Omega_{ij}^P \left[\check{r}_{ji} - \beta \check{\pi}_{ij} - (1-\beta)\check{\lambda}_i\right] \\
\frac{1-\beta}{\gamma} \frac{\partial \check{\lambda}_i}{\partial t} &= \sum_j \Omega_{ij}^P \left[\check{r}_{ji} - \beta \check{\pi}_{ij}\right] - (1-\beta)\check{\lambda}_i
\end{aligned}$$

Using the expressions for $\check{\pi}_{ii}$, $\check{\pi}_{ji}$ and $\check{r}_{ji}$, along with $\pi_{ii} = r_{ii}$, the law of motion can be written as

$$\begin{aligned}
\frac{1-\beta}{\gamma} \frac{\partial \check{\lambda}_i}{\partial t} &= \Omega_{ii}^P \left[\check{r}_{ii} - \beta \check{\pi}_{ii}\right] + \sum_{j \neq i} \Omega_{ij}^P \left[\check{r}_{ji} - \beta \check{\pi}_{ij}\right] - (1-\beta)\check{\lambda}_i \\
&= \Omega_{ii}^P (1-\beta) \frac{(1-\pi_{ii})}{1+\theta(1+\pi_{ii})} \check{\lambda}_i + \sum_{j \neq i} \Omega_{ij}^P \left[\frac{-\pi_{ii}}{1+\theta(1+\pi_{ii})} \check{\lambda}_i - \beta \frac{1}{1+\theta(1+\pi_{ii})} \check{\lambda}_i\right] - (1-\beta)\check{\lambda}_i \\
&= \left\{\frac{\Omega_{ii}^P (1-\beta)(1-\pi_{ii}) + \left(1-\Omega_{ii}^P\right)(-\pi_{ii} - \beta)}{1+\theta(1+\pi_{ii})} - (1-\beta)\right\} \check{\lambda}_i \\
&= \left\{\frac{\Omega_{ii}^P (1-\beta)(1-\pi_{ii}) - \pi_{ii}\left(1-\Omega_{ii}^P\right)(1-\beta) - \beta\left(1-\Omega_{ii}^P\right)(1+\pi_{ii})}{1+\theta(1+\pi_{ii})} - (1-\beta)\right\} \check{\lambda}_i \\
\frac{\partial \check{\lambda}_i}{\partial t} &= -\gamma \left\{1 - \frac{\Omega_{ii}^P - \pi_{ii}}{1+\theta(1+\pi_{ii})} + \frac{\beta}{1-\beta} \frac{\left(1-\Omega_{ii}^P\right)(1+\pi_{ii})}{1+\theta(1+\pi_{ii})}\right\} \check{\lambda}_i
\end{aligned}$$

E Multinationals

This section derives expressions for price indices, trade shares, and market clearing conditions. Across multinationals, let $v_{i1}(s)$ and $v_2(s)$ be the lowest and second lowest marginal costs of supplying good s to i. Then the price of good s in i is

$$p_i(s) = \min\left\{\frac{\varepsilon}{\varepsilon - 1} v_{i1}(s), v_{i2}(s)\right\}$$

Define

$$\begin{aligned}
\varphi_i &= \sum_{\tilde{k}} \lambda_{\tilde{k}} \left(\sum_{l=1}^{n} \left(w_l \kappa_{il} \delta_{l\tilde{k}}\right)^{-\theta/[1-\rho]}\right)^{1-\rho} \\
\pi_{ijk} &= \frac{1}{\varphi_i} \lambda_k \left(\sum_{l=1}^{n} \left(w_l \kappa_{il} \delta_{lk}\right)^{-\theta/[1-\rho]}\right)^{1-\rho} \frac{\left(w_j \kappa_{ij} \delta_{jk}\right)^{-\theta/[1-\rho]}}{\sum_{l=1}^{n} \left(w_l \kappa_{il} \delta_{lk}\right)^{-\theta/[1-\rho]}}
\end{aligned}$$

As with, Lemma 7, we begin with a lemma which will be useful intermediate step.

Lemma 10 *Suppose τ_1 and τ_2 satisfy $\tau_1 < 1$ and $\tau_1 + \tau_2 < 1$. Then*

$$\int_{s\in S_{ijk}} q_{jk1}(s)^{\tau_1\theta} p_i(s)^{-\tau_2\theta} ds = \tilde{B}(\tau_1,\tau_2)\pi_{ijk}\varphi_i^{\tau_1+\tau_2}\left(w_j\kappa_{ij}\delta_{jk}\right)^{\tau_1\theta}$$

where $\tilde{B}(\tau_1,\tau_2)$ is defined as in Lemma 7.

Proof. For $v_1 \le v_2$, define $\tilde{\mathcal{V}}_{ijk}\left(v_1,v_2\right)$ to be the fraction of goods for which the lowest cost source for i is a multinational based in k producing in j, and for which that multinational is unable to supply the good at cost lower than v_1, and for which no other multinational can supply the good at cost lower than v_2. We can then express $\tilde{\mathcal{V}}_{ijk}\left(v_1,v_2\right)$ using $\tilde{F}_{kj}$ and M_{kj} to represent the derivatives of $\tilde{F}_k$ and M_k with respect the jth argument:

$$\begin{aligned}\tilde{\mathcal{V}}_{ijk}(v_1,v_2) &= m\int_{v_2}^{\infty}\frac{w_l\kappa_{il}\delta_{lk}}{x^2}M_{kj}\left(\left\{\frac{w_l\kappa_{il}\delta_{lk}}{x}\right\}_l\right)M_k\left(\left\{\frac{w_l\kappa_{il}\delta_{lk}}{x}\right\}_l\right)^{m-1}\prod_{\tilde{k}\neq k}\tilde{F}_{\tilde{k}}\left(\left\{\frac{w_l\kappa_{il}\delta_{l\tilde{k}}}{x}\right\}_l\right)dx \\ &\quad +m\int_{v_1}^{v_2}\frac{w_l\kappa_{il}\delta_{lk}}{x^2}M_{kj}\left(\left\{\frac{w_l\kappa_{il}\delta_{lk}}{x}\right\}_l\right)M_k\left(\left\{\frac{w_l\kappa_{il}\delta_{lk}}{v_2}\right\}_l\right)^{m-1}\prod_{\tilde{k}\neq k}\tilde{F}_{\tilde{k}}\left(\left\{\frac{w_l\kappa_{il}\delta_{l\tilde{k}}}{v_2}\right\}_l\right)dx\end{aligned}$$

The two terms divide $\tilde{\mathcal{V}}_{ijk}\left(v_1,v_2\right)$ into instances where the lowest cost provider has cost less than v_2 and between v_1 and v_2 respectively. To interpret the first term, note that for each of the m multinationals based in k, the term $\frac{w_l\kappa_{il}\delta_{lk}}{x^2}M_{kj}\left(\left\{\frac{w_l\kappa_{il}\delta_{lk}}{x}\right\}_l\right)$ is the probability that that particular multinational can provide the good to i by producing in j at cost $x(> v_2)$ and cannot provide the good at a lower cost by producing elsewhere, while $M_k\left(\left\{\frac{w_l\kappa_{il}\delta_{lk}}{x}\right\}_l\right)^{m-1}\prod_{\tilde{k}\neq k}\tilde{F}_{\tilde{k}}\left(\left\{\frac{w_l\kappa_{il}\delta_{l\tilde{k}}}{x}\right\}_l\right)$ is the probability that none of the other multinationals (the other $m-1$ based in k or any based in another country) can provide the good to i at cost lower than x. The second term is similar, except since $x \in [v_1, v_2]$, it uses the probability that none of the other multinationals can provide the good to i at cost lower than v_2. This can be written more concisely as

$$\tilde{\mathcal{V}}_{ijk}(v_1,v_2) = m\int_{v_1}^{\infty}\frac{w_l\kappa_{il}\delta_{lk}}{x^2}M_{kj}\left(\left\{\frac{w_l\kappa_{il}\delta_{lk}}{x}\right\}_l\right)M_k\left(\left\{\frac{w_l\kappa_{il}\delta_{lk}}{\max\{x,v_2\}}\right\}_l\right)^{m-1}\prod_{\tilde{k}\neq k}\tilde{F}_{\tilde{k}}\left(\left\{\frac{w_l\kappa_{il}\delta_{l\tilde{k}}}{\max\{x,v_2\}}\right\}_l\right)dx$$

Note that since $\tilde{F}_k(\mathbf{q})^{\frac{1}{m}} = M_k(\mathbf{q})$, we can differentiate each with respect to the jth argument to get $M_{kj}(\mathbf{q}) = \frac{1}{m}\frac{\tilde{F}_{kj}(\mathbf{q})}{\tilde{F}_k(\mathbf{q})^{\frac{m-1}{m}}}$. We can therefore express $\tilde{\mathcal{V}}_{ijk}$ as

$$\tilde{\mathcal{V}}_{ijk}(v_1, v_2) = \int_{v_1}^{\infty} \frac{\frac{w_l\kappa_{il}\delta_{lk}}{x^2}\tilde{F}_{kj}\left(\left\{\frac{w_l\kappa_{il}\delta_{lk}}{x}\right\}_l\right)}{\tilde{F}_k\left(\left\{\frac{w_l\kappa_{il}\delta_{lk}}{x}\right\}_l\right)^{1-1/m}} \frac{\prod_{\tilde{k}}\tilde{F}_{\tilde{k}}\left(\left\{\frac{w_l\kappa_{il}\delta_{l\tilde{k}}}{\max\{x,v_2\}}\right\}_l\right)}{\tilde{F}_k\left(\left\{\frac{w_l\kappa_{il}\delta_{lk}}{\max\{x,v_2\}}\right\}_l\right)^{1/m}} dx$$

Define $\mathcal{V}_{ijk}(v_1, v_2) \equiv \tilde{\mathcal{V}}_{ijk}\left(m^{-\frac{1}{1-\beta}\theta}v_1, m^{-\frac{1}{1-\beta}\theta}v_2\right)$. As m grows large, this becomes

$$\mathcal{V}_{ijk}(v_1, v_2) = \int_{v_1}^{\infty} \frac{\frac{w_l\kappa_{il}\delta_{lk}}{x^2}F_{kj}\left(\left\{\frac{w_l\kappa_{il}\delta_{lk}}{x}\right\}_l\right)}{F_k\left(\left\{\frac{w_l\kappa_{il}\delta_{lk}}{x}\right\}_l\right)} \prod_{\tilde{k}} F_{\tilde{k}}\left(\left\{\frac{w_l\kappa_{il}\delta_{l\tilde{k}}}{\max\{x, v_2\}}\right\}_l\right) dx$$

With functional forms, $F_k(\mathbf{q}) = e^{-\lambda_k\left(\sum_{l=1}^n q_l^{-\theta/[1-\rho]}\right)^{1-\rho}}$, the second term in the integrand can be expressed as

$$\prod_{\tilde{k}} F_{\tilde{k}}\left(\left\{\frac{w_l\kappa_{il}\delta_{l\tilde{k}}}{\max\{x, v_2\}}\right\}_l\right) = \prod_{\tilde{k}} e^{-\lambda_{\tilde{k}}\left(\sum_{l=1}^n\left(\frac{w_l\kappa_{il}\delta_{l\tilde{k}}}{\max\{x,v_2\}}\right)^{-\theta/[1-\rho]}\right)^{1-\rho}} = e^{-\varphi_i\max\{x,v_2\}^{\theta}}$$

We also have $\frac{F_{kj}(\mathbf{q})}{F_k(\mathbf{q})} = \lambda_k\theta\left(\sum_{l=1}^n q_l^{-\theta/[1-\rho]}\right)^{1-\rho} q_j^{-\theta/[1-\rho]-1}$, so that the first term of the integrand is

$$\begin{aligned}
\frac{\frac{w_l\kappa_{il}\delta_{lk}}{x^2}F_{kj}\left(\left\{\frac{w_l\kappa_{il}\delta_{lk}}{x}\right\}_l\right)}{F_k\left(\left\{\frac{w_l\kappa_{il}\delta_{lk}}{x}\right\}_l\right)} &= \frac{w_l\kappa_{il}\delta_{lk}}{x^2}\lambda_k\theta\left(\sum_{l=1}^n\left(\frac{w_l\kappa_{il}\delta_{lk}}{x}\right)^{-\theta/[1-\rho]}\right)^{1-\rho-1}\left(\frac{w_j\kappa_{ij}\delta_{jk}}{x}\right)^{-\frac{\theta}{1-\rho}-1} \\
&= \theta x^{\theta-1}\lambda_k\left(\sum_{l=1}^n(w_l\kappa_{il}\delta_{lk})^{-\theta/[1-\rho]}\right)^{1-\rho-1}(w_j\kappa_{ij}\delta_{jk})^{-\frac{\theta}{1-\rho}} \\
&= \varphi_i\pi_{ijk}\theta x^{\theta-1}
\end{aligned}$$

The measure $\mathcal{V}_{ijk}$ can therefore be expressed as

$$\mathcal{V}_{ijk}(v_1, v_2) = \int_{v_1}^{\infty} e^{-\varphi_i\max\{x,v_2\}^{\theta}}\varphi_i\pi_{ijk}\theta x^{\theta-1}dx$$

with density

$$\frac{d^2\mathcal{V}_{ijk}(v_1, v_2)}{dv_1 dv_2} = \pi_{ijk}\varphi_i\theta v_1^{\theta-1}\varphi\theta v_2^{\theta-1}e^{-\varphi_i v_2^{\theta}}$$

With this density, we can derive the desired expression for the integral.

$$\begin{aligned}\int_{s\in S_{ijk}} q_{jk1}(s)^{\tau_1\theta} p_i(s)^{-\tau_2\theta} ds &= \int_0^\infty \int_0^{v_2} \left(\frac{w_j\kappa_{ij}\delta_{jk}}{v_1}\right)^{\tau_1\theta} \min\left\{v_2, \frac{\varepsilon}{\varepsilon-1}v_1\right\}^{-\tau_2\theta} \frac{d^2\mathcal{V}_{ijk}(v_1,v_2)}{dv_1 dv_2} dv_1 dv_2 \\ &= \int_0^\infty \int_0^{v_2} \left(\frac{w_j\kappa_{ij}\delta_{jk}}{v_1}\right)^{\tau_1\theta} \min\left\{v_2, \frac{\varepsilon}{\varepsilon-1}v_1\right\}^{-\tau_2\theta} \pi_{ijk}\varphi_i\theta v_1^{\theta-1}\varphi\theta v_2^{\theta-1} e^{-\varphi_i v_2^\theta} dv_1 dv_2\end{aligned}$$

Using the change of variables $x_2 = \varphi_i v_2^\theta$ and $x_1 = \varphi_i v_1^\theta$

$$\begin{aligned}\int_{s\in S_{ijk}} q_{jk1}(s)^{\tau_1\theta} p_i(s)^{-\tau_2\theta} ds &= \pi_{ijk}(w_j\kappa_{ij}\delta_{jk})^{\tau_1\theta}\varphi_i^{\tau_1+\tau_2} \int_0^\infty \int_0^{x_2} x_1^{-\tau_1} \min\left\{x_2, \left(\frac{\varepsilon}{\varepsilon-1}\right)^\theta x_1\right\}^{-\tau_2} e^{-x_2} dx_1 dx_2 \\ &= \tilde{B}(\tau_1,\tau_2)\pi_{ijk}\varphi_i^{\tau_1+\tau_2}(w_j\kappa_{ij}\delta_{jk})^{\tau_1\theta}\end{aligned}$$

■

With this lemma in hand we can derive expressions for the trade share and price index. The share of i's expenditure on goods produced in j by multinationals based in k is

$$\frac{\int_{S_{ijk}} p_i(s)^{1-\varepsilon} ds}{\sum_{\tilde{j},\tilde{k}} \int_{S_{i\tilde{j}\tilde{k}}} p_i(s)^{1-\varepsilon} ds} = \frac{\tilde{B}\left(0,\frac{\varepsilon-1}{\theta}\right)\pi_{ijk}\varphi_i^{\frac{\varepsilon-1}{\theta}}}{\sum_{\tilde{j},\tilde{k}} \tilde{B}\left(0,\frac{\varepsilon-1}{\theta}\right)\pi_{i\tilde{j}\tilde{k}}\varphi_i^{\frac{\varepsilon-1}{\theta}}} = \pi_{ijk}$$

and the price level satisfies

$$P_i^{1-\varepsilon} = \sum_{\tilde{j},\tilde{k}} \int_{S_{i\tilde{j}\tilde{k}}} p_i(s)^{1-\varepsilon} ds = \sum_{\tilde{j},\tilde{k}} \tilde{B}\left(0,\frac{\varepsilon-1}{\theta}\right)\pi_{i\tilde{j}\tilde{k}}\varphi_i^{\frac{\varepsilon-1}{\theta}} = \tilde{B}\left(0,\frac{\varepsilon-1}{\theta}\right)\varphi_i^{\frac{\varepsilon-1}{\theta}}$$

The labor market clearing condition is

$$\begin{aligned}L_j &= \sum_{i,k}\int_{s\in S_{ijk}} \frac{\kappa_{ij}\delta_{jk}}{q_{jk1}(s)} c_i(s) ds = \frac{1}{w_j}\sum_{i,k} w_j\kappa_{ij}\delta_{jk}\int_{s\in S_{ijk}} \frac{1}{q_{jk1}(s)} p_i(s)^{-\varepsilon}\frac{X_i}{P_i^{1-\varepsilon}} ds \\ w_j L_j &= \sum_{i,k} w_j\kappa_{ij}\delta_{jk}\frac{X_i}{P_i^{1-\varepsilon}}\int_{s\in S_{ijk}} \frac{1}{q_{jk1}(s)} p_i(s)^{-\varepsilon} ds\end{aligned}$$

Using Lemma 10, this is

$$
\begin{aligned}
w_j L_j &= \sum_{i,k} w_j \kappa_{ij} \delta_{jk} X_i \frac{\tilde{B}\left(-\frac{1}{\theta}, \frac{\varepsilon}{\theta}\right) \pi_{ijk} \varphi_i^{-\frac{1}{\theta}+\frac{\varepsilon}{\theta}} \left(w_j \kappa_{ij} \delta_{jk}\right)^{-1}}{\tilde{B}\left(0, \frac{\varepsilon-1}{\theta}\right) \varphi_i^{\frac{\varepsilon-1}{\theta}}} \\
&= \frac{\tilde{B}\left(-\frac{1}{\theta}, \frac{\varepsilon}{\theta}\right)}{\tilde{B}\left(0, \frac{\varepsilon-1}{\theta}\right)} \sum_{i,k} X_i \pi_{ijk} \\
&= \frac{\theta}{\theta+1} \sum_{i,k} X_i \pi_{ijk}
\end{aligned}
$$

Source Distributions Finally, we derive expressions for the source distributions. Before doing that, it is useful to note the following relationship

$$
\begin{aligned}
\pi_{ijk}^{1-\rho} \left(\sum_l \pi_{ilk}\right)^{\rho} &= \left[\frac{1}{\varphi_i} \lambda_k \left(\sum_{l=1}^{n} \left(w_l \kappa_{il} \delta_{lk}\right)^{-\theta/[1-\rho]}\right)^{1-\rho} \frac{\left(w_j \kappa_{ij} \delta_{jk}\right)^{-\theta/[1-\rho]}}{\sum_{l=1}^{n} \left(w_l \kappa_{il} \delta_{lk}\right)^{-\theta/[1-\rho]}}\right]^{1-\rho} \\
&\quad \times \left[\frac{1}{\varphi_i} \lambda_k \left(\sum_{l=1}^{n} \left(w_l \kappa_{il} \delta_{lk}\right)^{-\theta/[1-\rho]}\right)^{1-\rho}\right]^{\rho} \\
&= \frac{\lambda_k \left(w_j \kappa_{ij} \delta_{jk}\right)^{-\theta}}{\varphi_i}
\end{aligned}
$$

We first study learning from sellers, where learning is in proportion to expenditure. The source distribution is $G_i^S(q) = \sum_{j,k} \int_{\left\{s \in S_{ijk} | q_{jk1} < q\right\}} \left[\frac{p_i(s)}{P_i}\right]^{1-\varepsilon} ds$, so the important term in the learning equations is

$$
\begin{aligned}
\int_0^{\infty} q^{\beta\theta} dG_i^S(q) &= \sum_{j,k} \int_{\left\{s \in S_{ijk} | q_{jk1} < q\right\}} q_{jk1}^{\beta\theta}(s) \left[\frac{p_i(s)}{P_i}\right]^{1-\varepsilon} ds \\
&= \frac{\sum_{j,k} \tilde{B}\left(\beta, \frac{\varepsilon-1}{\theta}\right) \frac{\pi_{ijk}}{\varphi_i^{-\left(\beta \dot{-} \frac{\varepsilon-1}{\theta}\right)}} \left(w_j \kappa_{ij} \delta_{jk}\right)^{\beta\theta}}{P_i^{1-\varepsilon}} \\
&= \frac{\tilde{B}\left(\beta, \frac{\varepsilon-1}{\theta}\right)}{\tilde{B}\left(0, \frac{\varepsilon-1}{\theta}\right)} \sum_{j,k} \frac{\pi_{ijk}}{\varphi_i^{-\beta}} \left(w_j \kappa_{ij} \delta_{jk}\right)^{\beta\theta} \\
&= \frac{\tilde{B}\left(\beta, \frac{\varepsilon-1}{\theta}\right)}{\tilde{B}\left(0, \frac{\varepsilon-1}{\theta}\right)} \sum_{j,k} \pi_{ijk} \left[\frac{\lambda_k}{\pi_{ijk}^{1-\rho} \left(\sum_l \pi_{ilk}\right)^{\rho}}\right]^{\beta}
\end{aligned}
$$

With learning from producers, the source distribution is

$$
\begin{aligned}
G_j^P(q) &= \sum_{i,k} \int_{\left\{s\in S_{ijk} | q_{jk1}<q\right\}} \frac{1}{L_j}\frac{\kappa_{ij}\delta_{jk}}{q_{jk1}(s)} c_i(s)\, ds = \frac{1}{w_jL_j}\sum_{i,k} w_j\kappa_{ij}\delta_{jk}\int_{\left\{s\in S_{ijk} | q_{jk1}<q\right\}} \frac{1}{q_{jk1}(s)} p_i(s)^{-\varepsilon}\frac{X_i}{P_i^{1-\varepsilon}}ds \\
&= \frac{1}{w_jL_j}\sum_{i,k} w_j\kappa_{ij}\delta_{jk}\frac{X_i}{P_i^{1-\varepsilon}}\int_{\left\{s\in S_{ijk} | q_{jk1}<q\right\}} \frac{1}{q_{jk1}(s)} p_i(s)^{-\varepsilon} ds
\end{aligned}
$$

The important term in the learning equation is then

$$
\begin{aligned}
\int_0^\infty q^{\beta\theta} dG_j^P(q) &= \frac{1}{w_jL_j}\sum_{i,k} w_j\kappa_{ij}\delta_{jk}\frac{X_i}{P_i^{1-\varepsilon}}\int_0^\infty q_{jk1}(s)^{\beta\theta-1} p_i(s)^{-\varepsilon} ds \\
&= \frac{1}{w_jL_j}\sum_{i,k} w_j\kappa_{ij}\delta_{jk}X_i \frac{\tilde{B}\left(\beta-\frac{1}{\theta},\frac{\varepsilon}{\theta}\right)\frac{\pi_{ijk}}{\varphi_i^{-\left(\beta-\frac{1}{\theta}+\frac{\varepsilon}{\theta}\right)}}\left(w_j\kappa_{ij}\delta_{jk}\right)^{\beta\theta-1}}{\tilde{B}\left(0,\frac{\varepsilon-1}{\theta}\right)\varphi_i^{-\frac{1-\varepsilon}{\theta}}} \\
&= \frac{\tilde{B}\left(\beta-\frac{1}{\theta},\frac{\varepsilon}{\theta}\right)}{\tilde{B}\left(0,\frac{\varepsilon-1}{\theta}\right)}\frac{1}{w_jL_j}\sum_{i,k} X_i\pi_{ijk}\left[\varphi_i\left(w_j\kappa_{ij}\delta_{jk}\right)^{\theta}\right]^{\beta} \\
&= \frac{\tilde{B}\left(\beta-\frac{1}{\theta},\frac{\varepsilon}{\theta}\right)}{\tilde{B}\left(-\frac{1}{\theta},\frac{\varepsilon}{\theta}\right)}\sum_{i,k} r_{ijk}\left[\varphi_i\left(w_j\kappa_{ij}\delta_{jk}\right)^{\theta}\right]^{\beta} \\
&= \frac{\tilde{B}\left(\beta-\frac{1}{\theta},\frac{\varepsilon}{\theta}\right)}{\tilde{B}\left(-\frac{1}{\theta},\frac{\varepsilon}{\theta}\right)}\sum_{i,k} r_{ijk}\left[\frac{\lambda_k}{\pi_{ijk}^{1-\rho}\left(\sum_l \pi_{ilk}\right)^{\rho}}\right]^{\beta}
\end{aligned}
$$

where $r_{ijk} = \frac{\frac{\theta}{\theta+1}\pi_{ijk}X_i}{w_iL_i} = \frac{\frac{\tilde{B}\left(-\frac{1}{\theta},\frac{\varepsilon}{\theta}\right)}{\tilde{B}\left(0,\frac{\varepsilon-1}{\theta}\right)}\pi_{ijk}X_i}{w_iL_i}$

E.1 Symmetric Countries

For symmetric countries with symmetric trade cost κ and FDI cost δ, the trade FDI shares can be written as

$$\pi_{iii} = \frac{(1+(n-1)(\kappa\delta)^{-\frac{\theta}{1-\rho}})^{-\rho}}{(1+(n-1)(\kappa\delta)^{-\frac{\theta}{1-\rho}})^{1-\rho}+(n-1)(\delta^{-\frac{\theta}{1-\rho}}+\kappa^{-\frac{\theta}{1-\rho}}+(n-2)(\kappa\delta)^{-\frac{\theta}{1-\rho}})^{1-\rho}}$$

$$\pi_{iik} = \frac{(\delta^{-\frac{\theta}{1-\rho}}+\kappa^{-\frac{\theta}{1-\rho}}+(n-2)(\kappa\delta)^{-\frac{\theta}{1-\rho}})^{-\rho}\delta^{-\frac{\theta}{1-\rho}}}{(1+(n-1)(\kappa\delta)^{-\frac{\theta}{1-\rho}})^{1-\rho}+(n-1)(\delta^{-\frac{\theta}{1-\rho}}+\kappa^{-\frac{\theta}{1-\rho}}+(n-2)(\kappa\delta)^{-\frac{\theta}{1-\rho}})^{1-\rho}}, \qquad i \neq k$$

$$\pi_{ijk} = \frac{(\delta^{-\frac{\theta}{1-\rho}}+\kappa^{-\frac{\theta}{1-\rho}}+(n-2)(\kappa\delta)^{-\frac{\theta}{1-\rho}})^{-\rho}(\kappa\delta)^{-\frac{\theta}{1-\rho}}}{(1+(n-1)(\kappa\delta)^{-\frac{\theta}{1-\rho}})^{1-\rho}+(n-1)(\delta^{-\frac{\theta}{1-\rho}}+\kappa^{-\frac{\theta}{1-\rho}}+(n-2)(\kappa\delta)^{-\frac{\theta}{1-\rho}})^{1-\rho}}, \qquad i \neq j \neq k \neq i$$

$$\pi_{ijj} = \frac{(\delta^{-\frac{\theta}{1-\rho}}+\kappa^{-\frac{\theta}{1-\rho}}+(n-2)(\kappa\delta)^{-\frac{\theta}{1-\rho}})^{-\rho}\kappa^{-\frac{\theta}{1-\rho}}}{(1+(n-1)(\kappa\delta)^{-\frac{\theta}{1-\rho}})^{1-\rho}+(n-1)(\delta^{-\frac{\theta}{1-\rho}}+\kappa^{-\frac{\theta}{1-\rho}}+(n-2)(\kappa\delta)^{-\frac{\theta}{1-\rho}})^{1-\rho}}, \qquad i \neq j$$

$$\pi_{iji} = \frac{(1+(n-1)(\kappa\delta)^{-\frac{\theta}{1-\rho}})^{-\rho}(\kappa\delta)^{-\frac{\theta}{1-\rho}}}{(1+(n-1)(\kappa\delta)^{-\frac{\theta}{1-\rho}})^{1-\rho}+(n-1)(\delta^{-\frac{\theta}{1-\rho}}+\kappa^{-\frac{\theta}{1-\rho}}+(n-2)(\kappa\delta)^{-\frac{\theta}{1-\rho}})^{1-\rho}}, \qquad i \neq j$$

The price index is

$$P = \Gamma\left(1-\frac{\varepsilon-1}{\theta}\right)^{\frac{1}{1-\varepsilon}}\lambda^{-\frac{1}{\theta}}\Bigg(\left(1+(n-1)(\kappa\delta)^{-\frac{\theta}{1-\rho}}\right)^{1-\rho} + (n-1)\left(\kappa^{-\frac{\theta}{1-\rho}}+\delta^{-\frac{\theta}{1-\rho}}+(n-2)(\kappa\delta)^{-\frac{\theta}{1-\rho}}\right)^{1-\rho}\Bigg)^{-\frac{1}{\theta}}$$

References

Alvarez, F., F. Buera, and R. Lucas Jr (2008): "Models of Idea Flows," *NBER Working Paper 14135.*

——— (2013): "Idea Flows, Economic Growth, and Trade," *NBER Working Paper 19667.*

Alvarez, F. and R. J. Lucas (2007): "General equilibrium analysis of the Eaton-Kortum model of international trade," *Journal of Monetary Economics*, 54, 1726–1768.

Arkolakis, C., A. Costinot, and A. Rodríguez-Clare (2012): "New Trade Models, Same Old Gains?" *American Economic Review*, 102, 94–130.

Atkeson, A. and A. T. Burstein (2010): "Innovation, Firm Dynamics, and International Trade," *Journal of Political Economy*, 118, 433–484.

Bernard, A., J. Eaton, J. Jensen, and S. Kortum (2003): "Plants and Productivity in International Trade," *The American Economic Review*, 93, 1268–1290.

Connolly, M. and K.-M. Yi (2009): "How much of South Korea's growth miracle can be explained by trade policy?" Tech. rep.

Eaton, J. and S. Kortum (1999): "International Technology Diffusion: Theory and Measurement," *International Economic Review*, 40, 537–70.

——— (2002): "Technology, geography, and trade," *Econometrica*, 70, 1741–1779.

Feenstra, R. C., R. E. Lipsey, H. Deng, A. C. Ma, and H. Mo (2005): "World Trade Flows: 1962-2000," NBER Working Papers 11040, National Bureau of Economic Research, Inc.

Feyrer, J. (2009a): "Distance, Trade, and Income The 1967 to 1975 Closing of the Suez Canal as a Natural Experiment," Working Paper 15557, National Bureau of Economic Research.

——— (2009b): "Trade and Income – Exploiting Time Series in Geography," Working Paper 14910, National Bureau of Economic Research.

Jovanovic, B. and G. M. MacDonald (1994): "Competitive Diffusion," *Journal of Political Economy*, 102, 24–52.

Jovanovic, B. and R. Rob (1989): "The Growth and Diffusion of Knowledge," *Review of Economic Studies*, 56, 569–82.

Kortum, S. (1997): "Research, patenting, and technological change," *Econometrica: Journal of the Econometric Society*, 65, 1389–1419.

Lucas, R. E. (2009a): "Ideas and Growth," *Economica*, 76, 1–19.

——— (2009b): "Trade and the Diffusion of the Industrial Revolution," *American Economic Journal: Macroeconomics*, 1, 1–25.

Lucas, R. E. and B. Moll (2014): "Knowledge Growth and the Allocation of Time," *Journal of Political Economy*, 122, pp. 1–51.

Luttmer, E. (2012): "Eventually, Noise and Imitation Implies Balanced Growth," Working Papers 699, Federal Reserve Bank of Minneapolis.

Oberfield, E. (2013): "Business Networks, Production Chains, and Productivity: A Theory of Input-Output Architecture," Tech. rep.

Perla, J. and C. Tonetti (2014): "Equilibrium Imitation and Growth," *Journal of Political Economy*, 122, 52 – 76.

Perla, J., C. Tonetti, and M. Waugh (2013): "Equilibrium Technology Diffusion, Trade, and Growth," Tech. rep.

Ramondo, N. and A. Rodriguez-Clare (2013): "Trade, Multinational Production, and the Gains from Openness," *Journal of Political Economy*, 121, 273 – 322.

Sampson, T. (2014): "Dynamic Selection: An Idea Flows Theory of Entry, Trade and Growth," .

Simonovska, I. and M. E. Waugh (2014): "The elasticity of trade: Estimates and evidence," *Journal of International Economics*, 92, 34–50.