The Global Diffusion of Ideas*

Francisco J. Buera
Federal Reserve Bank of Chicago

Ezra Oberfield
Princeton University

December 7, 2015

Preliminary and Incomplete

Abstract

We provide a tractable theory of innovation and diffusion of technologies to explore the role of international trade. 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 its trade shares and the stocks of knowledge of its trading partners. We use the framework to quantify the contribution of bilateral trade costs to cross-sectional TFP differences, long-run changes in TFP, and individual post-war growth miracles.

*We are grateful to Costas Arkolakis, Arnaud Costinot, Sam Kortum, Bob Lucas, Erzo Luttmer, Marc Melitz, Jesse Perla, Andres Rodriguez-Clare, Esteban Rossi-Hansberg, Chris Tonetti, Mike Waugh. The views expressed herein are those of the authors and not necessarily those of the Federal Reserve Bank of Chicago or the Federal Reserve System. All mistakes are our own.

Economic miracles are characterized by protracted growth in productivity, per-capita income, and increases in trade flows. The experiences of 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.[3]

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 producers. In our theory openness affects the creation and diffusion of ideas by determining the distribution from which 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. In our model, openness to tradeaffects the quality of the insights drawn by producers because it determines the set of sellers to a country and the set of technologies used domestically.

In this context, we provide conditions under which the distribution of productivity among producers within each country always converges to a Frechet distribution, no matter how trade barriers shape individual producers' local interactions. As a consequence, the state of knowledge within a country can be summarized by the level of this distribution, which we call the country's

[1] Sachs and Warner (1995), Dollar (1992), Ben-David (1993), Coe and Helpman (1995), and Frankel and Romer (1999) suggest a strong relationship between openness and growth, although Rodriguez and Rodrik (2001) subsequently argued that many estimates in the literature suffered from econometric issues including omitted variables, endogeneity, and lack of robustness. More recent contributions to the literature have developed strategies to overcome some of these issues. To estimate the impact of trade on growth, Feyrer (2009a,b) studies the natural experiments of the decade-long closing of the Suez Canal in the 1970's and the long run decline in the cost of shipping goods by air, each of which had larger impacts on some pairs of countries than others, and Pascali (2014) studies the introduction of the steamship which affected some trade routes more than others. See also Lucas (2009b) and Wacziarg and Welch (2008), and Donaldson (2015) for a review of the literature.

[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.

[3] Parente and Prescott (1994) and Klenow and Rodriguez-Clare (2005) argue that without some form of international spillovers or externalities, growth models have difficulty accounting for several facts about growth and development. Each argue that these facts can be explained by catchup growth to a world frontier of knowledge, an idea that goes back to at least Nelson and Phelps (1966). Comin and Hobijn (2010) document large cross-country differences in the speed with which frontier technologies are adopted and Comin et al. (2012) show that the speed of diffusion declines with distance.

stock of knowledge. The model is thus compatible with the Eaton and Kortum (2002) machinery which has been useful in studying trade flows in an environment with many asymmetric countries. We show that the change in a country's stock of knowledge can be characterized in terms of only its trade shares, its trading partners' stocks of knowledge, and parameters. This both yields qualitative insights and enables us to use actual trade flows to quantify the role of trade and geography in shaping idea flows and growth.

Our main focus is on how barriers to trade alter the learning process. Starting from autarky, opening to trade results in a higher temporary growth rate, and permanently higher level, of the stock of knowledge, as producers are exposed to more productive ideas. We separate the gains from trade into a static component and a dynamic component. The static component consists of the gains from increased specialization and comparative advantage, whereas the dynamic component are the gains that operate through the flow of ideas.

We first explore an environment in which producers in a country gain insights from those that sell goods to the country, following Alvarez et al. (2013). With this specification of learning, the dynamic gains from reducing trade barriers are qualitatively different from the static gains. The dynamic gains are largest for countries that are relatively closed, whereas the static gains from trade are largest for countries that are already relatively open. For a country with high trade barriers, the marginal import tends to be made by a foreign producer 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 have lower productivity and generate lower quality ideas.

Our model nests, at two extremes, a simple version of the Kortum (1997) model of pure innovation and one closely related to the Alvarez et al. (2008, 2013) model of pure diffusion. We span these two extremes by varying a single parameter, β, which we label the strength of diffusion. β measures the contribution of insights from others in the productivity of new ideas. One striking observation is that, for either of these two extremes, if a moderately open country lowers its trade costs, the resulting dynamic gains from trade are relatively small, whereas when β is in an intermediate range, the dynamic gains are larger. When β is small so that insights from others are relatively unimportant, it follows immediately that dynamic gains tend to be small. When β is larger, insights from others are more central. However, in the limiting model as β approaches the extreme of one, a country accrues almost all of the dynamic gains from trade as long as it is not in autarky.[4] A moderately open country is much better off than it would be in autarky, but further reductions in trade costs have little impact. As a consequence, it is only when β is in an intermediate range that the dynamic gains from trade are both sizable and would result from reductions in trade costs in the empirically relevant range.

We also explore a second channel, that individuals may draw insights from others that produce domestically, following Sampson (2014) and Perla et al. (2013). In this setting, lower trade barriers increase domestic competition and improve the distribution of productivity among those that continue to produce domestically, raising the quality of insights manager might draw from. Under this specification of learning, we show that the long-run dynamic gains from trade simply amplify the static gains from trade. When β is larger so that insights from others contribute more to the productivity of new ideas, the static gains from trade are amplified more and the dynamic gains are more sizable.

We next use our model to study the dynamics of a trade liberalization. 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.

We also study the how trade barriers affect incentives to innovate. Following Bernard et al.

[4] In the environment studied by Alvarez et al. (2013), both the steady state growth rate and the mass in the right tail of countries' productivity distributions are proportional to the number of countries not in autarky; trade costs have no other impact on these objects. The environment we study here with $\beta \nearrow 1$ has similar properties albeit in level differences rather than growth rate differences.

(2003) we focus on a decentralization in which producers engage in Bertrand competition, and each producer earns profit on sales to any destination to which that producer is the lowest-cost provider of a good. Motivated by the potential for profit, producers hire labor to generate new ideas. In this environment, we extend the result of Eaton and Kortum (2001) that on any balanced growth path, each country's research effort is independent of trade barriers.

To explore the ability of the theory to account for the evolution of the world distribution of productivity, we specify a quantitative version of the model that includes non-traded goods and intermediate inputs, and equipped labor with capital and education, and use it to study the ability of the theory to account for cross-country differences in TFP in 1962 and its evolution between 1962 and 2000. Following Waugh (2010), we use panel data on trade flows and relative prices to calibrate the evolution of bilateral trade costs, and take the evolution of population, physical and human capital, i.e., equipped labor, from the data. Given the evolution of trade costs and equipped labor, our model predicts the evolution of each country's TFP.

The predicted relationship between trade and the stock of knowledge depends on the value of β, the strength of diffusion, which indexes the contribution of insights drawn from others to the productivity of new ideas. We provide a simple strategy to calibrate this parameter, but also simulate the model for various alternative values and explore how well the model can quantitatively account for cross-country income differences and the evolution of countries' productivity over time.

We find that differences in trade costs can account for up to 45% of the cross-sectional dispersion of TFP in 1962, and up to 34% of the dispersion of TFP growth between 1962 and 2000 (44% and 33% for the calibrated value of $\beta = 0.75$). A majority of ability of the theory to account for TFP differences is given by the dynamic gains from trade, as lower trade costs lead to an improvement in the composition of insights drawn by domestic producers. The quantitative model is particularly capable of explaining much of the evolution of TFP in growth miracles, accounting for over a third of the TFP growth in China, South Korea and Taiwan. The ability of the model to account for both the dispersion of TFP and the dispersion of TFP growth is highest for intermediate values of the diffusion parameter, β.

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), Lucas (2009a) and Luttmer (2012).[5] We are particularly related to recent applications of these frameworks to study the connection between trade and the diffusion of ideas (Lucas, 2009b; Alvarez et al., 2013; Perla et al., 2013; Sampson, 2014).

In our model, the productivity of new ideas combines both insights from others and an original component.[6] As discussed earlier, our theory captures the models in Kortum (1997) and Alvarez et al. (2008, 2013) as special, and we argue, quantitatively less promising cases. In Kortum (1997) there is no diffusion of ideas and thus no dynamic gains from trade. In Alvarez et al. (2013) when 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. In addition, for the intermediate cases that are the focus of our analysis, $\beta \in [0, 1)$, the frontier of knowledge converges to a Frechet distribution.[7] This allows us use the machinery of Eaton and Kortum (2002), enabling us to quantify the role of both trade barriers and geography in the flow of ideas.

Eaton and Kortum (1999) also build a model of the diffusion of ideas across countries in which 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 do not affect the diffusion of ideas.

Our works relates to a large literature studying the connection between trade and growth, including the early contributions by Grossman and Helpman (1991) and Rivera-Batiz and Romer (1991). The one that is closest to ours is Grossman and Helpman (1991). They consider a small open economy in which technology is transferred from the rest of the world as an external effect,

[5] 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. Chiu et al. (2011) study information issues in the transfer of ideas.

[6] See König et al. (2012) and Benhabib et al. (2014) for models in which individuals can choose either to imitate or to innovate.

[7] Alvarez et al. (2013) study the case with $\beta = 1$ and deterministic arrival of ideas. In their model the limiting distribution of productivities is only Frechet in the extreme cases of autarky and costless trade among symmetric countries. With Poisson arrival of ideas the limiting distribution is log-logistic for these extreme cases. In our model the limiting distribution is Frechet for any $\beta \in [0, 1)$ and any configuration of trade costs.

and the pace of technology transfer is assumed to depend on the volume of trade. Our model incorporates this channel along with several others and embeds the mechanism in a quantitative framework. In addition, our paper relates to a large empirical literature providing evidence on the relationship between openness and diffusion of technologies. Our reduced form evidence is reminiscent of the early evidence discussed in Coe and Helpman (1995) and Coe et al. (1997) about the importance of knowledge spillovers through trade. See Keller (2009) for a recent review of this empirical literature, considering alternative channels, including trade and FDI.

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.[8]

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 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

[8]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 describe 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 an original contribution that is drawn from an exogenous distribution with CDF $H(z)$. We refer to $H(z)$ as the exogenous distribution of ideas.[9]

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.[10]

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, $M_t(q)$, 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

[9]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 countries' frontiers of knowledge.

[10]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 $\tilde{F}_t(q) = M_t(q)^m$, 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:

Assumption 1

i. The exogenous distribution has a Pareto right tail with exponent θ, so that $\lim_{z\to\infty}\frac{1-H(z)}{z^{-\theta}} = 1$.

ii. $\beta \in [0,1)$.

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

The first part of the assumption assumes that the right tail of the exogenous distribution of ideas is regularly varying.[11] We also assume that the strength of diffusion, β is strictly less than one, introducing diminishing returns into the quality of insights one draws. For this section we make one additional assumption: the source distribution $\tilde{G}_t$ has a sufficiently thin tail. 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, $\lim_{q\to\infty} q^{\beta\theta}[1-M_0(q)] = 0$. For example, a bounded initial distribution of knowledge would satisfy this assumption.

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 $G_t(q) \equiv \tilde{G}_t\left(m^{\frac{1}{(1-\beta)\theta}}q\right)$

[11] 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.

Proposition 1 *If Assumption 1 holds, 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 *If Assumption 1 holds and $\lim_{t\to\infty} \lambda_t = \infty$, 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 if the arrival rate of insight grows 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[12]

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

Since the frontier of knowledge at t satisfies $\tilde{F}_t(q) = M_t(q)^m$, the joint distribution can be written as

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

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,

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

2 International Trade

Consider a world in which n economies interact through trade and ideas diffuse through the contact of domestic managers 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

[12] 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.

and Lucas (2007), which we briefly introduce before deriving the equations which 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[13] 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.[14]

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 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

[13] Note that we have assumed for simplicity that neither consumers nor workers internalize that their consumption or production decisions may affect the insights they may draw, and thus prices do not reflect the possibility that idea flows may result from the production or consumption of the good. This assumption is not innocuous; in general, prices depend on how much each agent internalizes.

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

i, the producer in j uses $\kappa_{ij}/q_{j1}(s)$ units of labor. Thus the labor market clearing constraint for 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 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 and uniformly drawn from the distribution of productivity among all managers that sell goods to a country.[15] In this case, the source distribution is given by

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

As we show in Appendix C, after 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

$$\begin{aligned} \dot{\lambda}_{it} &= \alpha_{it} \int_0^\infty x^{\beta\theta} dG_i^S(q) \\ &= \Gamma(1-\beta)\alpha_{it} \sum_j \pi_{ij} \left(\frac{\lambda_j}{\pi_{ij}} \right)^\beta \end{aligned} \tag{4}$$

where $\Gamma(\cdot)$ is the Gamma function. That is, the evolution of the stock of ideas is close to a weighted sum of the stock of knowledge in all countries, where the weights are given by expenditure shares.

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 barriers affect which managers are able to sell goods to a country. Trade leads to tougher competition, so that there is more selection among the producers from which insights are drawn. Starting from autarky, lower trade barriers make it less likely that low productivity domestic producers can compete with high pro-

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

ductivity foreign producers. The subsequent insights drawn from these high productivity foreign producers will be better quality than those drawn from the low productivity domestic producers.[16] Higher trade barriers, on the other hand, lead to more selection among foreign managers into selling goods to country i. In fact, the less a foreign country sells to country i, the stronger selection is among its producers. The average quality of insights drawn from j is given by $(\lambda_j/\pi_{ij})^\beta$, where λ_j/π_{ij} is an average of productivity among sellers from j to i. Holding fixed j's stock of knowledge, a smaller π_{ij} reflects more selection into selling goods to i, which means that the insights drawn from sellers from j are likely to be higher quality insights.

Nevertheless, the overall 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.[17]

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

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. Suppose, for example, that trade costs are symmetric. If a country spends equally on imports from two trading partners, one with a high wage and one with a low wage, the country would improve the quality of its insights by tilting trade toward the trading partner with the higher wage. Intuitively, the marginal seller in the high wage country is more productive–and would generate higher quality insights–than the marginal seller in the low wage country, as the former must overcome the high wage.[18]

[16]This mechanism is emphasized by Alvarez et al. (2013).

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

[18]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

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

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

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

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

2.2 Learning from Producers

Another natural source of ideas is the interaction of technology managers with other domestic producers, or workers employed by these producers. In this section we consider the case in which the insights are drawn uniformly from the distribution of productivity among domestic managers that are actively producing.[19] 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 $i \neq j \neq k \neq i$. In this case, any manager that exports her good also sells domestically.[20] This greatly simplifies characterizing the learning process. The source distribution is

$$G_i(q) = G_i^P(q) = \frac{\int_{s \in S_{ii} | q_{i1} \leq q} ds}{\int_{s \in S_{ii}} ds}$$

countries. Of course, whether free trade is optimal depends on what individuals are able to internalize; we have assumed that consumers do not internalize that their consumption decisions affect the quality of insights drawn by managers.

[19]When insights are drawn from domestic producers, the assumption are drawn uniformly, instead of in proportion to the labor used in the production of each good, is more important. See Appendix C.1.1 for a characterization of the dynamics of the stock of ideas under alternative assumptions.

[20]To see this, suppose that there were 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.

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 knowledge, is described by

$$\begin{aligned}\dot{\lambda}_{it} &= \alpha_{it}\int_0^{\infty} x^{\beta\theta} dG_i^P(q) \\ &= \Gamma(1-\beta)\alpha_{it}\left(\frac{\lambda_i}{\pi_{ii}}\right)^{\beta}\end{aligned}$$

Thus, the source distribution of country i is a function of the share of its expenditure on domestic goods 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.[21]

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

$$\dot{\hat{\lambda}}_{it} = \Gamma(1-\beta)\hat{\alpha}_{it}\left(\frac{\hat{\lambda}_{it}}{\pi_{ii}}\right)^{\beta} - \frac{\gamma}{(1-\beta)}\hat{\lambda}_{it}, \tag{9}$$

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

$$\hat{\lambda}_i = \Gamma(1-\beta)\frac{(1-\beta)\hat{\alpha}_i}{\gamma}\left(\frac{\hat{\lambda}_i}{\pi_{ii}}\right)^{\beta}. \tag{10}$$

2.3 Other Specifications of Learning

Variety

An implication of the learning from producers specification is that the rate of increase of a country's stock of ideas grows without bound as the share of its expenditure share shrinks to zero. In that case, only the most productive managers will be able to sell goods domestically, so the insights drawn from these firms will be very high quality. This causes the frontier of knowledge to increase at a faster rate. Because the arrival rate of ideas is independent of the mass of producers actively producing, low productivity firms simply crowd out high quality insights.

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

An alternative is that a manager would gain more and better insights if she were exposed to wider variety of production techniques. Suppose that ideas arrive in proportion to the mass of techniques a manager is exposed to. When insights are drawn from sellers, trade has no impact on the mass of good consumed, and hence on the variety of sellers one may draw insight from. On the other hand, when insights are drawn from producers, ideas arrive in proportion to the mass of domestic producers that are actively producing. This implies that a country's the stock of knowledge evolves as[22]

$$\dot{\lambda}_{it} = \Gamma(1-\beta)\alpha_{it}\pi_{ii}\left(\frac{\lambda_i}{\pi_{ii}}\right)^{\beta} \tag{11}$$

In contrast to the baseline specification, increased trade–a lower π_{ii}–would lower the growth rate of a country's stock of knowledge because of the loss of variety in learning. On a balanced growth path, the detrended stock of knowledge is

$$\hat{\lambda}_i \propto \alpha_i \pi_{ii}$$

Targeted Learning

If managers can glean better insights from more productive producers, one might think they might focus their attention so that insights are drawn disproportionately from those that are more productive.

Suppose that $\tilde{G}$ represents the distribution of productivity among those from whom a manager may draw insight. We assume now that the manager can target better insights by overweighting individual insights. Specifically, the individual can choose a schedule of arrival rates that accompany each potential insight. Let $\hat{\alpha}(x)$ be the arrival of insights from producers with productivity x. The manager chooses $\{\hat{\alpha}(x)\}$ subject to the constraint $\left[\int_0^\infty \hat{\alpha}(x)^{\frac{\phi}{\phi-1}} d\tilde{G}(x)\right]^{\frac{\phi-1}{\phi}} \leq \alpha$ for some $\phi > 1$.[23]

In this case, a country's stock of knowledge evolves as[24] $\dot{\lambda}_t = \int_0^\infty \hat{\alpha}_t(x) x^{\beta\theta} dG_t(x)$. An individual

[22] In an Eaton-Kortum framework, π_{ii} is both the share of i's spending on domestic goods *and* the fraction of varieties produced domestically. Thus in the baseline, the arrival rate is α_{it}, whereas in (11) the arrival rate is $\alpha_{it}\pi_{ii}$.

[23] The case of $\phi = 1$ would correspond to the baseline model, in which case the constraint could be written as $\sup_x \hat{\alpha}(x) \leq \alpha$.

[24] More formally, with the functional form assumptions, in the limit as $m \to \infty$, the law of motion for F will be $d\ln F_t(q) = \int_0^\infty \hat{\alpha}_t(x) x^{\beta\theta} dG(x)$ so that as $t \to \infty$, $F_t\left(\lambda_t^{1/\theta} q\right) \to e^{-q^{-\theta}}$ with $\dot{\lambda}_t = \int_0^\infty \hat{\alpha}_t(x) x^{\beta\theta} dG_t(x)$.

that wants to learn as quickly as possible will thus choose the schedule $\{\hat{\alpha}(x)\}$ to solve

$$\max_{\{\hat{\alpha}(x)\}} \int_0^\infty \hat{\alpha}(x) x^{\beta\theta} dG(x) \qquad \text{subject to} \qquad \left[\int \hat{\alpha}(x)^{\frac{\phi}{\phi-1}} dG(x)\right]^{\frac{\phi-1}{\phi}} \leq \alpha$$

Optimal behavior implies that the change in a country's stock of knowledge is

$$\dot{\lambda} = \alpha \left[\int \left(x^{\beta\theta}\right)^{\phi} dG(x)\right]^{\frac{1}{\phi}}$$

With learning from sellers, the change in a country's stock of knowledge is[25]

$$\dot{\lambda}_i = \Gamma(1-\beta\phi)\alpha_i \left[\sum_j \pi_{ij} \left(\frac{\lambda_j}{\pi_{ij}}\right)^{\beta\phi}\right]^{1/\phi}$$

Similarly, with learning from producers, the change in a country's stock of knowledge is

$$\dot{\lambda}_i = \Gamma(1-\beta\phi)\alpha_i \left[\left(\frac{\lambda_i}{\pi_{ii}}\right)^{\beta\phi}\right]^{\frac{1}{\phi}} = \Gamma(1-\beta\phi)\alpha_i \left(\frac{\lambda_i}{\pi_{ii}}\right)^{\beta}$$

For each specification, learning is faster when learning is more targeted ($\Gamma(1-\beta\phi)$ is increasing in ϕ), although this constant plays no role in how the economy responds to changes in trade costs. As we show in the appendix, an environment with a higher ϕ is quantitatively very similar to an environment with higher β.

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:

$$y_i \equiv \frac{w_i}{P_i} \propto \left(\frac{\lambda_i}{\pi_{ii}}\right)^{1/\theta} \tag{12}$$

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

[25] Note that for either specification, a finite growth rate of the stock of knowledge requires $\phi < 1/\beta$. This limits how directly a manager can target the highest producers with the highest productivity.

models, e.g., Eaton and Kortum (2002).[26] 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. In a symmetric world, the share of a country's expenditure on domestic goods is $\pi_{ii} = \frac{1}{1+(n-1)\kappa^{-\theta}}$, while the share of its expenditure on imports from each trading partner is $\frac{1-\pi_{ii}}{n-1}$. Specializing either equation (8) or equation (10), each country's de-trended stock of knowledge on a balanced growth path is

$$\text{Sellers} \quad : \quad \hat{\lambda}_i = \left[\frac{(1-\beta)\hat{\alpha}}{\gamma}\Gamma(1-\beta)\right]^{\frac{1}{1-\beta}} \left[\pi_{ii}^{1-\beta} + (n-1)\left(\frac{1-\pi_{ii}}{n-1}\right)^{1-\beta}\right]^{\frac{1}{1-\beta}}. \tag{13}$$

$$\text{Producers} \quad : \quad \hat{\lambda}_i = \left[\frac{(1-\beta)\hat{\alpha}}{\gamma}\Gamma(1-\beta)\right]^{\frac{1}{1-\beta}} \pi_{ii}^{-\frac{\beta}{1-\beta}} \tag{14}$$

The de-trended real per-capita income is obtained by substituting these into equation (12)

$$\text{Sellers} \quad : \quad \hat{y}_i \propto \left[1 + (n-1)^{\beta}\left(\frac{1-\pi_{ii}}{\pi_{ii}}\right)^{1-\beta}\right]^{\frac{1}{1-\beta}\frac{1}{\theta}} \tag{15}$$

$$\text{Producers} \quad : \quad \hat{y}_i \propto \pi_{ii}^{-\frac{1}{1-\beta}\frac{1}{\theta}} \tag{16}$$

Using these equations, we can ask how a change in trade costs would impact countries' real incomes. It is instructive to compare three cases: the static case in which $\beta = 0$, learning from sellers, and learning from producers. For each, we can summarize how countries' real incomes

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

change with trade patterns by looking at the elasticity of real income with respect to the share of expenditures spent on domestic goods.[27]

Holding a country's stock of knowledge fixed, the change in real income arising from a changing trade barriers depends only on the trade elasticity:

$$\left.\frac{d\ln y_i}{d\ln \pi_{ii}}\right|_{\lambda_i \text{ fixed}} = -\frac{1}{\theta}$$

For each of the two specifications of learning, we can summarize the elasticity of real income to the domestic expenditure share:

$$\text{Sellers} \quad : \quad \frac{d\ln \hat{y}_i}{d\pi_{ii}} = -\frac{1}{\theta}\frac{(n-1)^{\beta}\left(\frac{1-\pi_{ii}}{\pi_{ii}}\right)^{1-\beta}\frac{1}{1-\pi_{ii}}}{1+(n-1)^{\beta}\left(\frac{1-\pi_{ii}}{\pi_{ii}}\right)^{1-\beta}} \tag{17}$$

$$\text{Producers} \quad : \quad \frac{d\ln \hat{y}_i}{d\ln \pi_{ii}} = -\frac{1}{1-\beta}\frac{1}{\theta} \tag{18}$$

Both the static and dynamic gains depend on the curvature of the productivity distribution, θ; a higher θ corresponds to thinner right tails. With higher θ, there are fewer highly productive producers abroad whose goods can be imported, and there are fewer highly productive producers from whom insights may be drawn. 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. This can be seen most clearly be comparing autarky to costless trade. Equations (15) and (16) reveal that for either specification of learning, the ratio of real income under costless trade ($\pi_{ii} = 1/n$) to real income under autarky ($\pi_{ii} = 1$) is $n^{\frac{1}{1-\beta}\frac{1}{\theta}}$. In the limit as β goes to 1, the gains from trade relative to autarky grow arbitrarily large.[28]

[27] A convenient feature of the symmetric example is that, since every country has the same stock of knowledge and the same wage, the share of a country's expenditure on domestic goods is $\frac{1}{1+(n-1)\kappa^{-\theta}}$. Thus the change in κ causes the same change in trade shares whether stocks of knowledge are held fixed, insights are drawn from sellers, or insights are drawn from producers. In a world with asymmetric countries, the a change in trade barriers (κ) would cause different changes in trade shares in each version of the model. However, as we will show below, the overarching message—that when insights are drawn from producers the dynamic gains amplify the static gains whereas when insights are drawn from sellers the dynamics gains are largest when countries are close to autarky—will remain.

[28] These limiting cases are close to the models analyzed by Alvarez et al. (2013), Sampson (2014), and Perla et al. (2013). 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

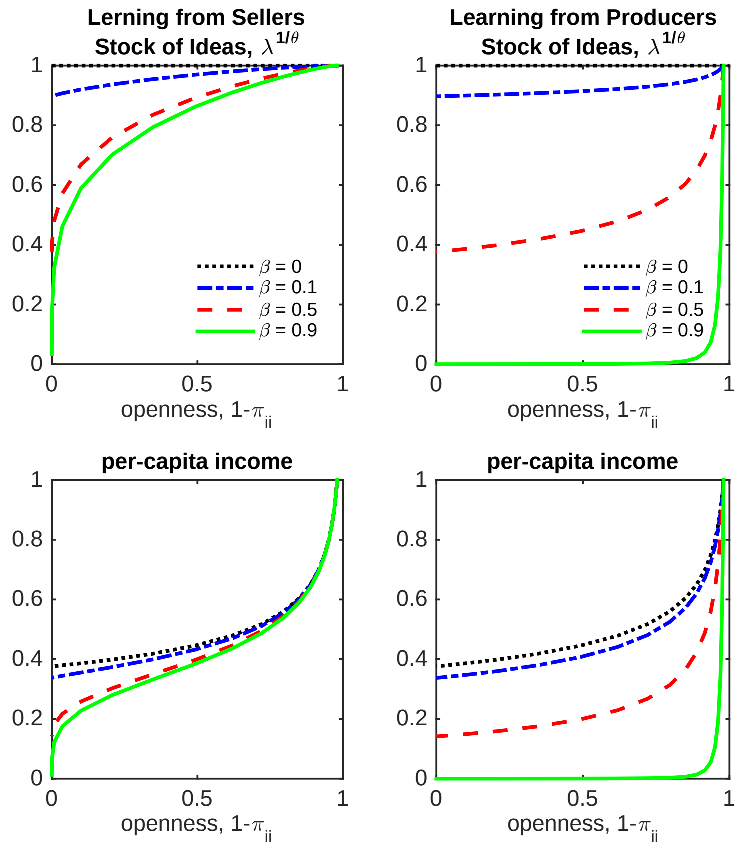

Figure 1: Gain from Reducing common trade barriers

Note: This figure shows each country's stock of ideas and per capita income relative to their values under costless trade.

For several values of β, the top panels of Figure 1 illustrate the common value of each country's stock of knowledge relative to its level under free trade. The bottom panels show the corresponding real income per capita. The left (right) panels focus on the specification of when insights are drawn from sellers (domestic producers). As a benchmark, the dotted line represents $\beta = 0$, which corresponds to the static trade model of Eaton and Kortum (2002). As trade costs rise, countries become more closed and their stocks of knowledge decline. When β is larger, the dynamic gains from trade are larger.

When insights are drawn from domestic producers, the gains from trade simply amplify the static gains. The diffusion parameter β determines the strength of the is amplification. One way of interpreting equation (18) is that the diffusion of ideas causes the static gains from trade to

the growth rate unchanged.

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.[29]

When insights are drawn from sellers, the dynamic gains from reducing trade barriers are qualitatively different from the static gains. The dynamic gains are largest when the world is relatively closed, whereas the static gains are largest when the world is relatively open. This can be seen from the left panels of Figure 1 but also by inspecting the limiting values of (17). As the world becomes more open, the total gains from reducing trade barriers corresponds to the static gains, $\lim_{\pi_{ii}\to 1/n} \frac{d\ln \hat{y}_i}{d\ln \pi_{ii}} = -\frac{1}{\theta}$. In contrast, as the world becomes more closed, the marginal dynamic gains grow arbitrarily large, $\lim_{\pi_{ii}\to 1} \frac{d\ln \hat{y}_i}{d\ln \pi_{ii}} = -\infty$. Put differently, when the economy is relatively open, the total gains from reducing trade barriers are composed mostly of the static gains, whereas when the world is relatively closed, the total gains are composed mostly of the dynamic gains.

To understand this, consider a country close to autarky. If trade costs decline, the marginal import tends to be made by a foreign producer with high productivity. While the high trade costs imply that the static gains from trade remain relatively small, the insights drawn from this marginal producer tends 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.

In contrast, when insights are drawn from domestic producers, the dynamic gains from reducing trade barriers are largest when the world is already relatively open, as shown clearly in the bottom right panel.

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.

[29] In a world with roundabout production, 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.

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 $j \neq i$. To a first order the long-run impact of a change in κ on country i's real income is

$$\begin{aligned}
\text{Sellers:} \quad \frac{d\log y_i}{d\log\kappa} &= -\frac{1+2\theta}{\frac{1-\Omega_i\beta}{1-\Omega_i}\frac{1-\pi_{ii}}{(1-\beta)+\beta(1-\pi_{ii})}\frac{\pi_{ii}+\theta(1+\pi_{ii})}{1-\pi_{ii}}+1} \\
\text{Producers:} \quad \frac{d\log y_i}{d\log\kappa} &= -\frac{1+2\theta}{(1-\beta)\frac{\pi_{ii}+\theta(1+\pi_{ii})}{1-\pi_{ii}}+1}
\end{aligned}$$

where $\Omega_i \equiv \frac{\pi_{ii}^{1-\beta}\lambda_i^{\beta}}{\sum_j \pi_{ij}^{1-\beta}\lambda_j^{\beta}}$.

When insights are drawn from sellers, the term Ω_i is the share of the growth in i's stock of knowledge that is associated with purchasing goods from i. One implication is that, holding fixed π_{ii}, the response of real income to a decline in trade costs is larger when Ω_i 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.

A second implication of the claim is that for both specifications of learning, the dynamic gains from trade are always weakly positive. The static gains can be can be found by evaluating either expression at $\beta = 0$.

Finally, when insights are drawn from sellers, in the limiting model as β approaches one, the dynamic gains from lower trade barriers approaches zero.[30] This may seem puzzling; as the contribution of insights from others in the productivity of new ideas becomes larger and the model approaches one of pure diffusion, the dynamic gains from trade become relatively unimportant.

To resolve this, it will be useful to plot the the income of the single deviant country for various values of β. Consider now 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

[30]To see this, simply evaluate the expression at $\beta = 0$ and $\beta = 1$ and note that $\beta = 0$ implies $\Omega_i = \pi_{ii}$.

cost, i.e., $\kappa_{nj} = \kappa_{jn} = \kappa_n \geq 1$, $j < n$.[31]

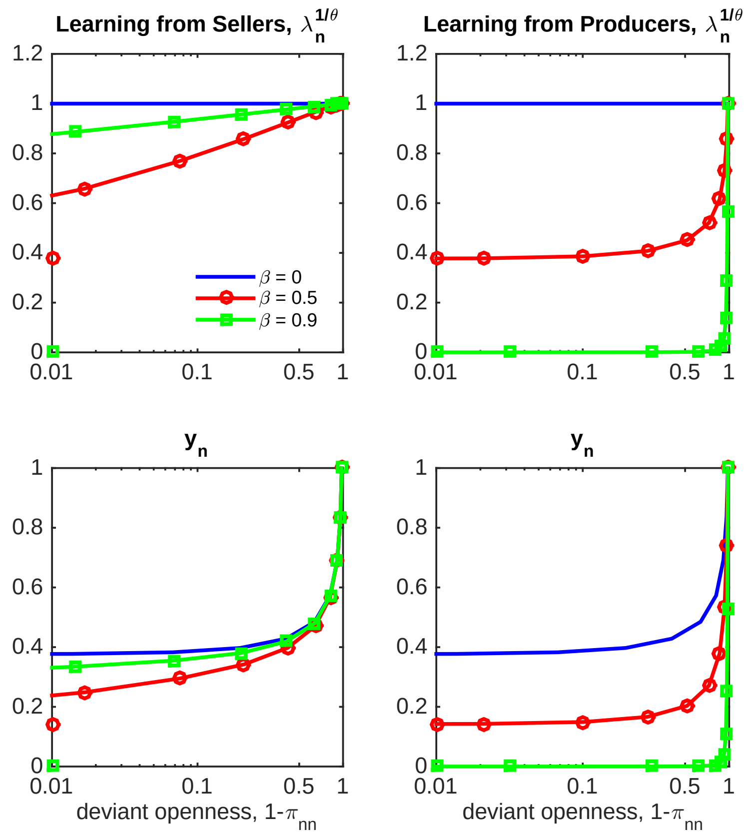

Figure 2: The Stock of Ideas and Per Capita Income of the Deviant Economy.

Note: The figure plots a deviant country's stock of knowledge (top panels) and per-capita income (bottom panels) relative to what it would be under costless trade for the cases when insights are drawn from sellers (left panels) and domestic producers (right panels). Each curve is continuous, and the dot along the left axis is the value in autarky, the limit as $\pi_{ii} \to 1$.

The top panels of Figure 2 show how the deviant country's stock of knowledge changes with the degree of openness. The x-axis measures openness as the fraction of a country's spending on imports, $1 - \pi_{ii}$. On the y-axis we report the stock of ideas relative to the case with costless trade ($\kappa_n = 1$). The bottom panels show the corresponding real incomes. The different lines correspond to alternative values of β, which controls the importance of insights from others. The solid line

[31] 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 $\theta = 4$, the curvature of the Frechet distribution, which also 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.

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 the right panels show learning from domestic producers.

Across balanced growth paths, as the 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, trade costs have effects on per-capita income beyond the static gains from trade.

Figure 2 replicates the curious features that if the deviant economy is moderately open, the gains from lowering trade costs are small if the model is close to one of pure innovation ($\beta = 0$) or close to one of pure diffusion ($\beta = 1$). Those two models differ however, in the gains relative to autarky. For $\beta = 0.9$, when a country moves from autarky to only slightly open, the dynamic gains from trade are quite large. But subsequent lowering of trade costs have relatively small impact on the country's stock of knowledge. It is only for intermediate values of β that lowering trade barriers would have a larger dynamic impact a country's stock of knowledge for a wide range of trade shares.

Why are the dynamic gains from trade concentrated near autarky when β is close to one? The reason is the concavity generated by β in combining insights from others with the exogenous components of ideas. When β is large, the difference between a high and low quality insight is magnified, and a country's growth depends much more heavily on insights from the most productive producers. When a country is only slightly open, it is already importing goods from most of the highest productivity foreign producers. Indeed, as $\beta \to 1$ as long as the deviant country is even slightly open, its stock of knowledge is as high as it would be under costless trade.[32] This feature of the model will be especially important for understanding our quantitative results when we study the implications of actual changes in trade volumes.

3.3 A Core-Periphery Economy

The interaction between geography and the diffusion of knowledge can be easily seen with the example of an economy that takes a core-periphery structure. Suppose there are n core countries

[32] 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.

and n periphery countries. Trade between a core country and any other country incurs an iceberg trade cost of κ. Trade among any two periphery countries must pass through the core, and thus incurs an iceberg cost of κ^2. All countries are otherwise symmetric.

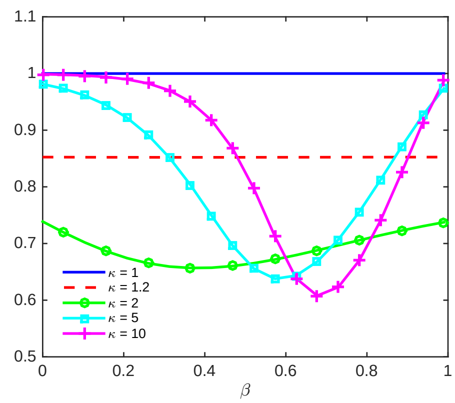

Figure 3: A Core-Periphery Economy

Note: For various values of iceberg trade costs, this figure plots the ratio of real income in periphery countries to real income in core countries.

Figure 3 shows the real income of periphery countries relative to that of the core countries. Each curve corresponds a level of κ, and shows the ratio of real incomes for various values of β. Note that for each level of trade barriers, the relative income of periphery countries as β approaches one is the the same as it would be in a static trade model. Consistent with the discussion in the previous section, the income gap is wider when β takes an intermediate value.

The income of core and periphery countries are similar when trade costs are either very low ($\kappa \approx 1$) or very high ($\kappa \nearrow \infty$); in either case, core countries effectively have no advantage. Thus if trade costs fall steadily, income differences will initially grow and eventually shrink.

3.4 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?

Consider a world economy that starts with $n-1$ open economies and a single deviant economy that are on a balanced growth path. Figure 4 shows the evolution of the real income in the initially deviant economy following a trade liberalization. The left panel shows an example in which the deviant country is initially in autarky and the $n-1$ open economies trade costlessly. For each of the two learning specifications, the figure traces out the real income in the deviant country.[33] The paths of the (de-trended) stocks of knowledge solve the differential equations in (7) and (9), depending on whether insights are drawn from sellers or producers.[34] On impact real income jumps as it would in a static model. Over time, the deviant country's stock of knowledge improves. When insights are drawn from sellers, real income converges more quickly to the steady state; 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.[35]

The right panel of Figure 4 shows a more empirically relevant example of a world where trade costs are such that imports comprise 5% of expenditures for the deviant country and 20% of expenditures for $n-1$ open countries. At time zero, trade costs for the deviant country fall enough so that its import share rises to 20%. Each curve shows real income relative to the new symmetric balanced growth path when insights are drawn from sellers. In line with previous results, for intermediate values of β, the total increase in income is larger.

In addition, the change in the deviant country's stock of knowledge leads to a protracted transition as the dynamic gains from trade are slowly realized. Ten years after the liberalization, only XXX% of the dynamic gains from trade have been realized.

[33] We use this extreme example of a liberalization from autarky to costless trade because these are two special cases in both specifications of learning predict the same stocks of knowledge. This makes it easier to contrast the speed of convergence across the two specifications.

[34] We set $\beta = 0.5$. The rest of the parameters follow the calibration in footnote 31.

[35] We can get a more general version of this result for a small open economy in a world with arbitrary trade barriers. Log-linearizing around a balanced growth path, let $\check{y}_i$ denote the log deviation of i's detrended real income from its long run value and let variables with no decoration denote their long run values. The speed of convergence of a small open economy is

$$\begin{aligned} \text{Sellers} \quad &: \quad \frac{d}{dt}\log \check{y}_i = -\gamma\left\{1 - \frac{\Omega_i - \pi_{ii}}{1+\theta\left(1+\pi_{ii}\right)} + \frac{\beta}{1-\beta}\left(1-\Omega_i\right)\right\} \\ \text{Producers} \quad &: \quad \frac{d}{dt}\log \check{y}_i = -\gamma\left\{1 + \frac{\beta}{1-\beta}\frac{1-\pi_{ii}}{1+\theta\left(1+\pi_{ii}\right)}\right\} \end{aligned}$$

where $\Omega_i \equiv \frac{\pi_{ii}^{1-\beta}\lambda_i^{\beta}}{\sum_j \pi_{ij}^{1-\beta}\lambda_j^{\beta}}$ is the share of i's insights drawn from i. From these expressions, one can infer both that convergence is faster when diffusion is more important (β is larger) and that the speed of convergence does not depend on α_i. Convergence is faster with learning from sellers unless Ω_i is significantly larger than than π_{ii}, a case in which i's stock of knowledge is much larger than that of its trading partners.

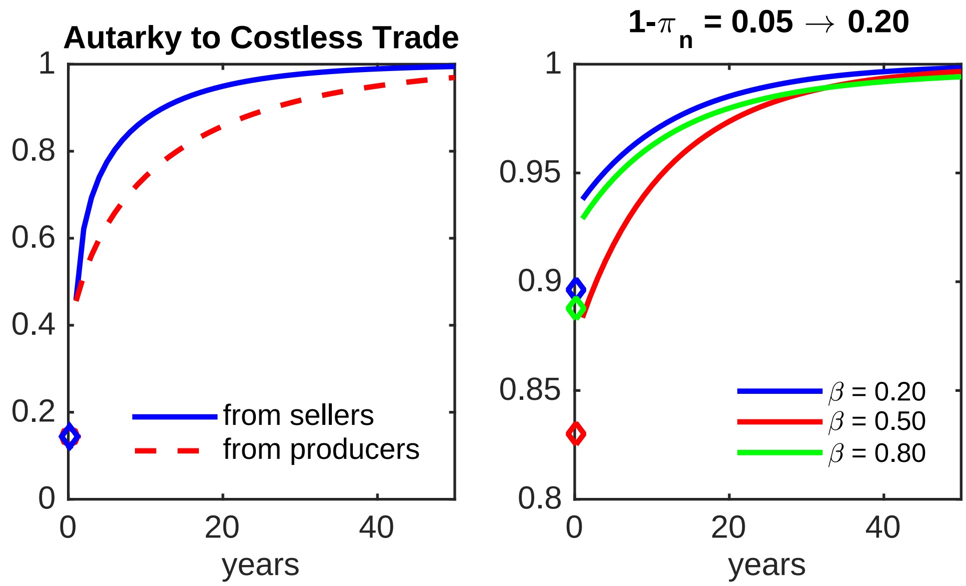

Figure 4: Dynamics Following a Trade Liberalization.

Note: This figure shows the evolution of real income for a deviant country whose trade barriers suddenly fall. The left panel compares the predictions of each specification of learning when the deviant country moves from autarky costless trade. The right panel studies learning from sellers and compares the predictions for several values of β when trade costs fall enough to shift the deviant country's import share from 5% to 20%. Each curve shows real income relative to the new symmetric balanced growth path.

4 Research

This section endogenizes the arrival rate of ideas, broadly following Rivera-Batiz and Romer (1991) and Eaton and Kortum (2001). Labor can engage in two types of activities, production and research. The production sector is described by Section 2. In the research sector, entrepreneurs generate ideas by hiring labor. The labor resource constraint in country i at t is thus

$$L_{it}^P + L_{it}^R = L_{it}$$

where L_{it}^P is the labor used in production and L_{it}^R is the labor used in research. We will show that on a balanced growth path, the fraction of labor engaged in research is independent of trade barriers.

We assume that there is a mass of managers and each manager is, in principle, capable of producing all varieties $s \in [0, 1]$. Each manager is characterized by a profile of productivities $q(s)$ with which she can produce the various goods. We assume that if an individual manager employs

l units of labor in research, then ideas arrive independently for each variety at rate $\tilde{\alpha} l$. Thus the arrival of goods is uniform across goods; research effort is not directed at particular goods.[36]

Suppose also that the entrepreneurs behave as if there is a tax T_i on profit. This may be an actual tax, or it may stand in for other distortions (as in Parente and Prescott (1994)). Let V_{it} is the expected pretax value of a single idea generated in i at t. Each manager chooses a research intensity l to maximize $\tilde{\alpha} l(1-T_i)V_{it} - w_{it} l$. For research to be interior, it must be that

$$\tilde{\alpha} V_{it} = w_{it}$$

We next compute the expected pretax value of an idea, V_{it}. In Appendix E we prove the following intermediate step: if $\Pi_{i\tau}$ is total flow of profit earned by entrepreneurs in i at time τ, then the flow of profit earned in i at time τ from ideas generated between t and t' (with $t < t' < \tau$) is $\frac{\lambda_{it'} - \lambda_{it}}{\lambda_{i\tau}} \Pi_{i\tau}$. The basic idea is that, among ideas on the frontier at time τ, knowing the time at which the idea was generated does not provide any additional information about the quality of the idea.[37]

Taking the limit as $t' \to t$ implies that the flow of profit at τ from ideas generated at the instant t is $\frac{\dot{\lambda}_{it}}{\lambda_{i\tau}} \Pi_{i\tau}$. As a consequence, the present value of revenue from ideas generated in i at instant t is

$$\int_t^\infty e^{-\rho(\tau - t)} \frac{P_{it}}{P_{i\tau}} \frac{\dot{\lambda}_{it}}{\lambda_{i\tau}} \Pi_{i\tau} d\tau$$

where $e^{-\rho(\tau-t)} \frac{P_{it}}{P_{i\tau}}$ is the real discount factor between t and τ. The cumulative arrival of ideas at t is $\tilde{\alpha} L_{it}^R$, so that the total pretax value of the ideas generated at t is $\tilde{\alpha} L_{it}^R V_{it}$. We thus have

$$\tilde{\alpha} L_{it}^R V_{it} = \int_t^\infty e^{-\rho(\tau - t)} \frac{P_{it}}{P_{i\tau}} \frac{\dot{\lambda}_{it}}{\lambda_{i\tau}} \Pi_{i\tau} d\tau$$

The optimal choice of research intensity implies $\tilde{\alpha}(1-T_i)V_{it} = w_{it}$. In addition, Section 2 showed that profit among all entrepreneurs is proportional to the wage bill in production, $\Pi_{i\tau} = \frac{w_{i\tau} L_{i\tau}^P}{\theta}$.

[36] We could just as easily have assumed that each manager is capable of producing a subset of the goods with positive measure, and call this subset an industry. The part of the assumption that is crucial is that the research effort is uniform across varieties.

[37] Both the arrival rate of ideas and the source distribution at time t affect the probability that the idea is on the frontier at $\tau \geq t$ and the unconditional distribution of the idea's productivity, these have no impact on the conditional distribution of productivity conditioning on being on the frontier. This is a useful and well known property of extreme value distributions, see Eaton and Kortum (1999).

Together these imply that the optimal research intensity satisfies

$$w_{it}L_{it}^{R} = (1-T_i)\int_t^{\infty} e^{-\rho(\tau-t)}\frac{P_{it}}{P_{i\tau}}\frac{\dot{\lambda}_{it}}{\lambda_{i\tau}}\frac{w_{i\tau}L_{i\tau}^{P}}{\theta}d\tau$$

Letting $r_{it} = \frac{L_{it}^{R}}{L_{it}}$ be the fraction of labor engaged in research, this can be rearranged as

$$r_{it} = \frac{1-T_i}{\theta}\int_t^{\infty} e^{-\rho(\tau-t)}\frac{P_{it}}{P_{i\tau}}\frac{\dot{\lambda}_{it}}{\lambda_{i\tau}}\frac{(1-r_{i\tau})\,w_{i\tau}L_{i\tau}}{w_{it}L_{it}}d\tau$$

Finally, using $w_{it}/P_{it} \propto (\lambda_{it}/\pi_{iit})^{1/\theta}$, this can be written as

$$r_{it} = \frac{1-T_i}{\theta}\int_t^{\infty} e^{-\rho(\tau-t)}\,(1-r_{i\tau})\,\frac{\dot{\lambda}_{it}}{\lambda_{i\tau}}\frac{(\lambda_{i\tau}/\pi_{ii\tau})^{1/\theta}\,L_{i\tau}}{(\lambda_{it}/\pi_{iit})^{1/\theta}\,L_{it}}d\tau$$

If labor grows at rate γ so that $L_{i\tau} = L_{it}e^{\gamma(\tau-t)}$, then there is balanced growth path with $r_{it} = r_i$, $\lambda_{it} = e^{\frac{\gamma}{1-\beta}t}\lambda_i$, $\pi_{iit} = \pi_{ii}$. Plugging these in gives

$$r_i = \frac{1-T_i}{\theta}\int_t^{\infty} e^{-\rho(\tau-t)}\,(1-r_i)\,\frac{\frac{\gamma}{1-\beta}}{e^{\frac{\gamma}{1-\beta}(\tau-t)}}\left(e^{\frac{\gamma}{1-\beta}(\tau-t)}\right)^{1/\theta}e^{\gamma(\tau-t)}d\tau$$

Integrating and rearranging gives a simple characterization of the fraction of the labor force engaged in research:

$$\frac{r_i}{1-r_i} = \frac{1-T_i}{\theta\left[(1-\beta)\frac{\rho}{\gamma}+\beta\right]-1} \tag{19}$$

Equation 19 implies that on a balanced growth path, the fraction of labor engaged in research is independent of both trade barriers and the cross-country distribution of knowledge. The only thing that alters research effort are distortions on the payoff to innovation. This aligns with results of Eaton and Kortum (2001), Atkeson and Burstein (2010), and the knowledge specification of Rivera-Batiz and Romer (1991) with flows of only goods, all of which imply that integration has little impact on R&D effort.

It is important to keep in mind that, in this context, integration still has an impact on a country's stock of knowledge. Even if a country's R&D effort does not change, integration could lead to larger increases in a country's stock of knowledge if new ideas are based on better insights.[38]

[38] See also Rivera-Batiz and Romer (1991) and Baldwin and Robert-Nicoud (2008).

Finally, we define

$$\alpha_{it} = \frac{\tilde{\alpha}}{m} r_{it} L_{it}$$

where r_{it} is defined in equation (19) and depends on country-specific distortion to R&D effort.

5 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. With this in mind, 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 addition, to simplify the exposition, we focus on the case in which insights are drawn from sellers to a market. This version of the model has particularly rich testable implications and, as we show in this section, it provides a promising quantitative theory of dynamic gains from trade.[39]

5.1 Extended Trade Model

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 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

[39] In the Appendix we present results for two alternative cases in which insights are drawn from domestic producers.

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.[40]

In addition to the iceberg transportation costs κ_{ij}, we assume that a fraction $1-\mu$ 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 Appendix F we present the expressions for price indices, trade share and evolution of the stock of ideas of this version of the model.

5.2 Calibration

We need to calibrate seven common parameters, $(\theta, \eta, \zeta, \mu, \gamma, \varepsilon, \beta)$, and two sets of parameters that are country and time specific, the matrix of transportation costs $\mathbf{K}_t = [\kappa_{int}]$ and the vector of arrival rates $\alpha_t = (\alpha_{1t}, ..., \alpha_{nt})$.

We set $\theta = 4$. 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 tradable goods $\mu = 0.5$. We set $\varepsilon = 1$, but note that alternative values do not affect the results significantly.

Following the strategy in Waugh (2010), we show in Appendix F that given values for θ, μ, and ε as well as data on bilateral trade shares over time, the iceberg cost of shipping a good to country i from country j at time t is

$$\kappa_{ijt} = \frac{p_{it}}{p_{jt}} \left(\frac{1-\pi_{iit}}{\pi_{ijt}} \frac{Z_{it}}{1-Z_{it}} \right)^{\frac{1}{\theta}} \left[\frac{(1-\mu) + \mu Z_{it}^{-\frac{\varepsilon-1}{\theta}}}{(1-\mu) + \mu Z_{jt}^{-\frac{\varepsilon-1}{\theta}}} \right]^{\frac{1}{\varepsilon-1}}$$

[40]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 technology as the consumption and intermediate input aggregates.

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 data on real GDP and the price index from PWT 8.0.[41]

To assign values to the vector of arrival rates $\hat{\alpha}_{\mathbf{t}} = (\hat{\alpha}_{1t}, ..., \hat{\alpha}_{nt})$ we proceed in two steps. Given the evolution of trade flows summarized by Z_{it}, we compute, in each year, the stocks of knowledge needed to match each country's TFP using

$$\hat{\lambda}_{it} \propto \left[(1-\mu) + \mu Z_{it}^{-\frac{\varepsilon-1}{\theta}}\right]^{\frac{-\theta}{\varepsilon-1}} \left(\frac{w_{it}}{P_{it}}\right)^{(1-\eta)\theta}$$

This is a generalization of equation (12) for the model with intermediate inputs and non-traded goods. We measure TFP in the data as a standard Solow residual using real GDP, physical capital (K), employment (emp) and average human capital (h) from the PWT 8.0, i.e., $TFP =$ real GDP$/[K^{1/3} \cdot (emp \cdot h)^{2/3}]$.[42]

Given the evolution of trade flows and the stock of knowledge, and a values for β and γ, we back out sequences of the arrival rate of ideas using the law of motion of the stock of ideas

$$\begin{aligned}\hat{\lambda}_{it+1} \propto & \hat{\alpha}_{it} \left[\frac{1-\mu}{(1-\mu) + \mu Z_{it}^{-\frac{\varepsilon-1}{\theta}}} \hat{\lambda}_{it}^{\beta} \right. \\ & \left. + \frac{\mu Z_{it}^{-\frac{\epsilon-1}{\theta}}}{1-\mu+\mu Z_{it}^{-\frac{\epsilon-1}{\theta}}} \left[Z_{it}^{1-\beta} \hat{\lambda}_{it}^{\beta} + (1-Z_{it})^{1-\beta} \sum_{j \neq i} \left(\frac{\pi_{ij}}{1-\pi_{ii}}\right)^{1-\beta} \hat{\lambda}_{jt}^{\beta} \right] \right] - \frac{\gamma}{1-\beta} \hat{\lambda}_{it}\end{aligned}$$

This is a discrete time generalization of equation (4) for the model with non-trade goods. The

[41] In particular, we measure real GDP using real GDP at constant national prices (rgdpna). We scale the real GDP series for each country so that its value in 1962 coincides with the real GDP measure given by the output-side real GDP at chained PPPs (rgdpo). We measure the price index using the price level of cgdpo (pl_gdpo), where cgdpo is the output-side real GDP at current PPPs.

[42] In any calibration of the model, we must take a stand on how to apportion a country's TFP into a stock of knowledge, which may generate idea flows, and other factors, such as allocational efficiency, that are unlikely to diffuse across borders. Our baseline calibration assumes that physical and human capital differences are unlikely to diffuse across borders, but that after controlling for those, all residual TFP differences are due to differences in the stocks of knowledge and trade barriers. In Appendix XX we consider an alternative calibration strategy. We project $\log TFP$ onto on R&D intensity, the log of the human capital stock and the log of an import-weighted average of trading partners' TFP. We assign the residual TFP from this regression to a neutral productivity terms affecting the units of equipped labor and not the stock of knowledge, and choose stocks of knowledge to match predicted TFP from the regression. RESULTS ARE....

sequence of arrival rates of ideas are the residuals that explained the evolution of TFP between 1962 and 2000 that is not accounted by the dynamics of trade costs. When doing counterfactuals where the arrival rate of ideas is held constant, we assume that countries were on a balanced growth path in 1962, and assign to each country the arrival rate of ideas $\hat{\alpha}_{\mathbf{i,0}}$ to exactly match the stocks of knowledge in 1962 when specializing equation (**??**) to a steady state, i.e., $\hat{\lambda}_{it+1} = \hat{\lambda}_{it} = \hat{\lambda}_{i1962}$. We also use this as one benchmark to evaluate the cross-sectional implications of the theory in Section 5.7.

We are left with two parameters to calibrate: the strength of the diffusion of ideas, β, and the growth rate of the arrival rate of ideas, γ. Rather than taking a strong stand on the value of the diffusion parameter, β, we explore how well the model can quantitatively account for cross-country income differences and the evolution of countries' productivity over time for alternative values of $\beta \in (0, 1)$. That being said, it is useful to discuss a simple (albeit heroic) strategy to calibrate this parameter. On a balanced growth path the growth rate of productivity is $(1/\theta)(\dot{\lambda}/\lambda) = \gamma/(\theta(1-\beta))$. Identifying the growth rate of the arrival of ideas with the average growth rate of population in the US between 1962 and 2000, $\gamma = 0.01$, assuming that the growth rate of productivity on a balanced growth path equals the average growth rate for the US between 1962 and 2000 (0.8% per year), the value of $\theta = 4$ implies an approximate value for the diffusion parameter $\beta = 0.7$.[43] When we consider alternative value of the diffusion parameter β, we recalibrate γ to match an average growth rate of TFP in the US of 0.8 percent.

5.3 Sample Selection

The sample of countries in our quantitative analysis consists of a balanced panel of countries that is obtained by merging the PWT 8.0 with the NBER-UN dataset on bilateral trade flows from 1962 to 2000. We further restrict this sample to those countries with a population above 1 million in 1962 and oil rents that are smaller than 20 % of GDP in 2000. We exclude Hong Kong, Panama and Singapore, as these are countries where re-exports play a very large role. The final sample consists of 65 countries.[44]

[43] This discussion ignores the impact that changes in world openness had on the TFP growth of the US. If we take this into account, the diffusion parameter that is consistent with the observed growth of TFP in the US between 1962 and 2000 is approximately $\beta = 0.6$.

[44] Argentina, Australia, Austria, Belgium-Luxemburg (we consider the sum of the two countries, as the UN-NBER trade data is reported only for the sum), Bolivia, Brazil, Cameroon, Canada, Chile, China, Colombia, Costa Rica, Cote

5.4 Reduced Form Evidence

Before discussing the results from the calibrated model we present suggestive reduced form evidence of the mechanisms emphasize by the theory. We start by discussing cross-sectional evidence in 1962, the first year of our sample.

Given the arrival rate of ideas in a country, the theory predicts that the main drivers of a country's TFP are its openness and the TFP of their trading partners. The first panel of Figure 5 shows that countries that are less open (high π_{ii}) tend to have lower TFP, although this relationship is not statistically significant. The second panel shows the relationship between the TFP of a country's trading partners and its own TFP. In particular, for each country we compute an import weighted average of a country's trading partners' TFP: $\frac{\sum_{j \neq i} \pi_{ij} TFP_j}{1-\pi_{ii}}$. The figure shows that countries with more productive trading partners tend to be (statistically significantly) more productive.

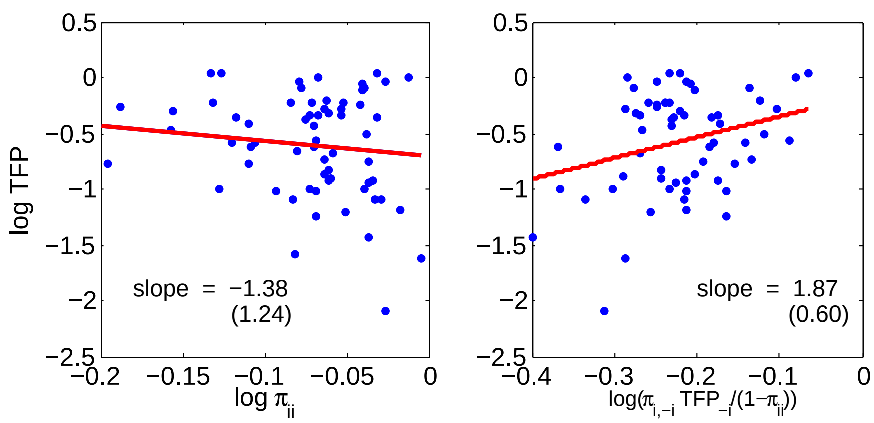

Figure 5: Cross-sectional TFP differences in 1962

Note: The first panel shows the cross-sectional relationship between (lack of) openness, as measured by countries' expenditure shares on domestic goods, and TFP. The right panel shows the cross-sectional relationship between the each country's TFP and its exposure to other high TFP trading partners, as measured by an import-weighted average of trading partners' TFP. In each panel we report the slope of the regression line and its standard error in parenthesis.

Over time, among the many factors that would alter a country's productivity, the model emphasizes changes in openness, changing exposure to trading partners, and changes in trading partners' productivity. Figure 6 shows some simple reduced form patterns in the data.

d'Ivoire, Denmark, Dominican Republic, Ecuador, Egypt, Finland, France, Germany, Ghana, Greece, Guatemala, Honduras, India, Indonesia, Ireland, Israel, Italy, Jamaica, Japan, Jordan, Kenya, South Korea, Malaysia, Mali, Mexico, Morocco, Mozambique, Netherlands, New Zealand, Niger, Norway, Pakistan, Paraguay, Peru, Philippines, Portugal, Senegal, South Africa, Spain, Sri Lanka, Sweden, Switzerland, Syria, Taiwan, Tanzania, Thailand, Tunisia, Turkey, Uganda, United Kingdom, United States, Uruguay, and Zambia.

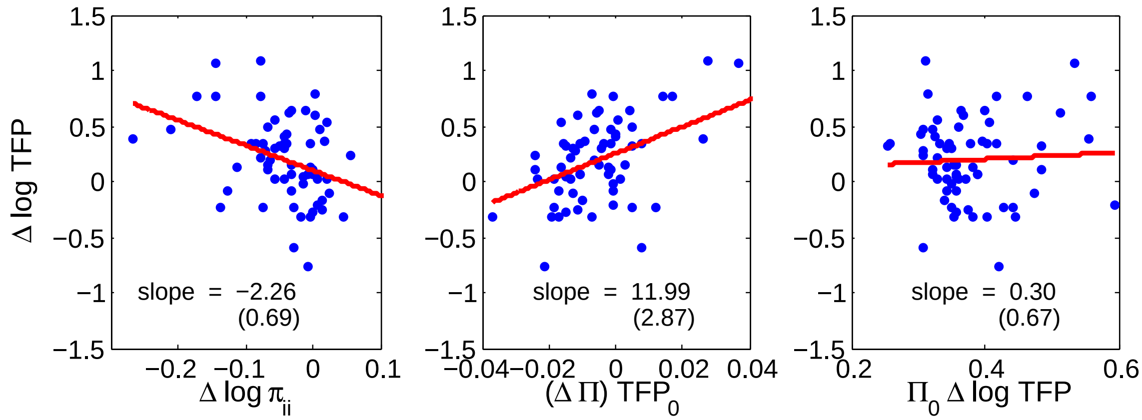

Figure 6: Openness and Changes in TFP, 1962-2000

Note: The first panel shows the cross-sectional relationship between changes in countries' TFP and changes in (lack of) openness, as measured by the change in expenditure share on domestic goods. The second panel shows the cross-sectional relationship between changes in countries' TFP and changes in countries exposure to trading partners who had high TFP in 1962, where exposure is an import-weighted average. The third panel shows the cross-sectional relationship between changes in countries' TFP and changes in trading partners' TFPs, weighted by expenditure shares in 1962. In each panel we report the slope of the regression line and its standard error in parenthesis.

The first panel shows the relationship between changes in openness and changes in TFP. Consistent with the model, countries that increased expenditures on imports tended to have (statistically significantly) larger increases in TFP.

The second panel shows the association between the change in countries composition of expenditures and TFP growth. For each country, we compute the changes in exposure to trading partners with high initial TFP. Specifically, for country i we compute $\sum_j (\pi_{ij}^{2000} - \pi_{ij}^{1962}) TFP_j^{1962}$. Consistent with the theory, there is a clear pattern that countries that increased import exposure to trading partners with high initial productivity saw (statistically significantly) larger increases in TFP.

The third panel shows that countries whose trading partners became more productive tended to see increases in TFP. While this relationship is consistent with the model, it is fairly weak and statistically insignificant.

5.5 Explaining the Dynamics of TFP

Motivated by the reduced form evidence, this section studies the ability of the model to account for the evolution of productivity over time. We begin our quantitative analysis by comparing the static and dynamic gains from changes in trade costs. As discussed in Section 5.2, we use expenditure shares to back out the evolution of bilateral iceberg trade costs over time. We make the stark assumption that, given trade costs, each country was on its balanced growth path in 1962. Under this assumption, we can find the set of country-specific arrival rates $\{\alpha_i\}$ that match the cross-section of productivity perfectly. Finally, we ask: if the arrival rates of ideas had remained constant over time and only trade costs changed, what would each country's TFP be in 2000?

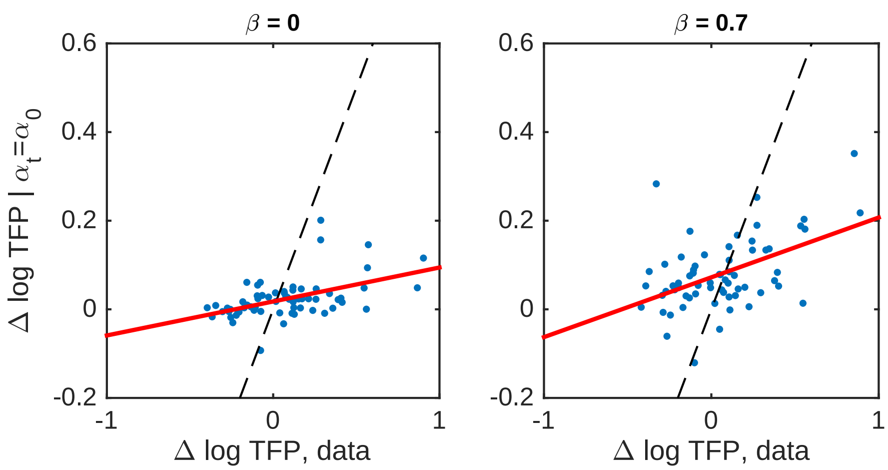

Figure 7: Trade and the TFP Dynamics, 1962-2000

Note: Each panel plots countries' actual changes in TFP against the predicted change in TFP of the model if only trade costs change. We compute this counterfactual under the assumptions that the arrival rates are heterogenous across countries, that each country was on its balanced growth path in 1962, and that arrival rates have remained constant since 1962. The first panel assumes that $\beta = 0$. The second panel assumes $\beta = 0.7$. In addition, each figure plots a dashed 45-degree line and a red regression line.

Figure 7 compares the predicted change in TFP from the model to that of the data for alternative calibration of the diffusion parameter. Each point represents a country, and each panel contains a regression line through the observations and a dashed 45-degree line. The first panel shows the predicted changes in TFP when $\beta = 0$ so that there are no dynamic gains from trade. The model predicts only small changes in TFP, consistent with small static gains from trade. In the second panel β is set to 0.7, the value implied by the simple calibration discussed at the end of Section 5.2.

In this panel the regression line is more upward sloping, indicating a stronger relationship between the predicted and actual changes in TFP.

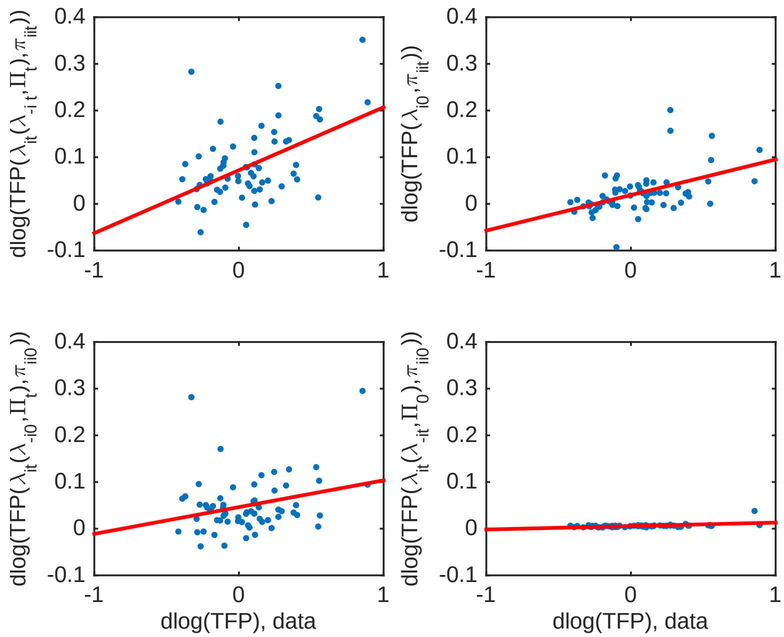

Figure 8: Decomposition of Changes in TFP, 1962-2000

Note: The top left panel plots changes in TFP against predicted changes in TFP under the assumption that learning is from sellers, $\beta = 0.7$, and the arrival rate of ideas is kept at its 1962 value, $\alpha_{it} = \alpha_{i0}$. The remaining three panels plot actual changes in TFP against the various components of predicted changes in TFP. The top right panel hold fixed all countries stocks of knowledge. The bottom panels allow each country's stock of knowledge to evolve, but hold fixed initial trade shares in computing TFP. In the bottom left panel, learning is such that each country's trading partners' stocks of knowledge are held fixed at their initial levels, but trade shares evolve. In the bottom right panel, learning is such that trading partners' stocks of knowledge evolve but initial trade shares are held fixed at their initial levels.

To get a sense of what is driving the model's predictions, we can decompose the predicted TFP changes into various components. Each panel in Figure 8 displays the changes in countries' TFP on the x-axis and some measure of the model's predicted changes in TFP when $\beta = 0.7$ and insights are drawn from sellers on the y-axis. The top-left panel contains predicted TFP of the model where trade costs are allowed to change, but the arrival rate of ideas is kept equal to the 1962 value. In the model, each country's TFP is a function of its stock of knowledge and its expenditure on domestic goods, $TFP(\lambda_i, \pi_{ii})$. In turn, each country's stock of knowledge is a function of others' stocks of knowledge and import shares, $\lambda\left(\{\lambda_{jt}\}_{j\neq i,t\geq 0},\{\Pi_t\}_{t\geq 0}\right)$, where $\Pi = \{\pi_{ij}\}_{i,j=1,\ldots,N}$ is the matrix of trade shares. The top right panel shows the static effects of changes in trade costs, $d\ln TFP(\lambda_{it}(\lambda_{-i0},\Pi_0),\pi_{iit}) = \ln TFP(\lambda_{i0},\pi_{iit}) - \ln TFP(\lambda_{i0},\pi_{ii0})$, where each country's stock of knowledge is held fixed at its initial level. Countries that saw an increase in the trade share tended to increase TFP more. The two lower panels show the contribution of the two drivers of dynamic gains from trade. The lower left panel holds fixed trading partner's stocks of knowledge, but allows the dynamic gains from trade through changing exposure to different trading partners. The lower right panel holds fixed trade shares but allows the dynamic gains from trade through changes in trading partners' productivities. Consistent with Figure 6, changes in exposure to trading partners plays an important role, but changes in trading partners productivities plays almost no role.

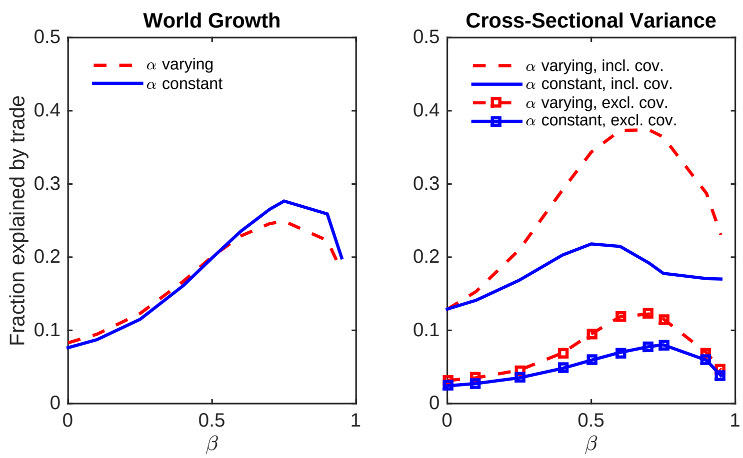

Figure 9: The contribution of changes in trade costs to changes in TFP

Note: This figure reports the fraction of TFP growth accounted for by trade costs, for various values of β, according to two decompositions. In both panels, the solid lines correspond to (20) in which the contribution of trade is evaluated holding the arrival rates constant; the dashed lines to correspond to (21) in which the contribution of trade is evaluated at the evolving arrival rates that are consistent with data. The left panel reports the fraction of total growth in TFP accounted for by changes in trade costs. The right panel reports the fraction of variance in TFP growth rates accounted for by changes in trade costs. The lines with square markers report exclude the covariance between the contribution from trade and the contribution from changing arrival rates of ideas; the lines without markers include the covariance. In all cases, insights are drawn from sellers.

Figure 9 shows a more systematic assessment of how the strength of diffusion alters the explanatory power of trade in the model. For each β we compute two counterfactuals to assess the contribution of trade to changes in TFP. Each provides a different way of dividing changes in each country's TFP into a contribution from changes in trade barriers and a contribution from changes in the arrival rates of ideas.[45] First, we compute how countries' TFP would have evolved if trade costs evolved as they do in data but each country's arrival rate of ideas remained fixed at its 1962 level. The predicted change in TFP is the contribution of trade and the residual is the contribution from changes in arrival rates. The second counterfactual computes the changes in TFP if the arrival rates of ideas evolved as they do in the data but the trade costs remained fixed at their 1962 levels. The predicted changes in TFP are the contribution of changes in the arrival rates and the contribution of trade are the residuals.[46]

$$\ln \frac{TFP_i(\alpha_t, \kappa_t)}{TFP_i(\alpha_0, \kappa_0)} = \underbrace{\ln \frac{TFP_i(\alpha_0, \kappa_t)}{TFP_i(\alpha_0, \kappa_0)}}_{\text{cont. from trade}} + \underbrace{\ln \frac{TFP_i(\alpha_t, \kappa_t)}{TFP_i(\alpha_0, \kappa_t)}}_{\text{cont. from arrival rates}} \tag{20}$$

$$\ln \frac{TFP_i(\alpha_t, \kappa_t)}{TFP_i(\alpha_0, \kappa_0)} = \underbrace{\ln \frac{TFP_i(\alpha_t, \kappa_0)}{TFP_i(\alpha_0, \kappa_0)}}_{\text{cont. from arrival rates}} + \underbrace{\ln \frac{TFP_i(\alpha_t, \kappa_t)}{TFP_i(\alpha_t, \kappa_0)}}_{\text{cont. from trade}} \tag{21}$$

We can summarize the role of trade in a few different ways. We first compute the fraction of changes in TFP growth accounted for by contributions from trade and from contributions from changes in the arrival rate of ideas.[47] The solid line corresponds to the first counterfactual in which the contribution of trade is evaluated at the initial arrival rates, and the dashed line corresponds to the second counterfactual in which contribution of trade is evaluated at the actual arrival rates.

According to this decomposition, both counterfactuals indicate that the static gains from trade ($\beta = 0$) account for roughly eight percent of the growth in TFP from 1962-2000. With $\beta > 0$, changes in trade costs are more important. The contribution trade is highest if $\beta = 0.7$, a setting in which a quarter of the increases in TFP are accounted for by changes in trade costs.

While the model predicts that changes in trade cost can account for a significant fraction of TFP growth, it is possible that model assigns growth to the wrong countries. To address this, the

[45] This is in some ways analogous to dividing changes in nominal GDP into changes in a price index and changes in a quantity index. If the price index is a Lespeyres index then the quantity index is a Paasche index and vice versa.

[46] We use here the shorthand that α_t and κ_t represent a sequence of the vector of arrival rates and matrices of trade costs that are required to match the data, and α_0 and κ_0 are the initial values. Thus $TFP_i(\alpha_0, \kappa_t)$ represents the TFP of country i in a counterfactual in which the arrival rates of technologies are held fixed at their initial levels but trade costs evolve as they do in the data. By construction $TFP_i(\alpha_t, \kappa_t)$ equals country i's TFP in 2000 and $TFP_i(\alpha_0, \kappa_0)$ is what i's TFP would have been had the world remained on the balanced growth path since 1962 – the vector of TFP in 2000 would be a scalar multiple of the vector of 1962 TFPs.

[47] For each of the two counterfactuals, this is $\frac{\sum_i \text{contribution from trade}}{\sum_i \ln \frac{TFP_i(\alpha_t, \kappa_t)}{TFP_i(\alpha_0, \kappa_0)}}$

right panel of Figure 9 shows the fraction of variation in TFP growth rates accounted for trade costs. The variance of TFP growth can be decomposed into three components, the variance of contributions of changes in trade costs, the variance of the contributions of changes in arrival rates, and the covariance of the two. The figure plots four lines. The two solid lines correspond to the decomposition in (20) in which the contribution of trade is evaluated holding the arrival rates of ideas fixed at their initial levels. The two dashed lines correspond to (21) in which the contribution of trade is evaluated allowing the arrival rates to evolve as they must to explain the data. The lines that are marked with squares represent the fraction of variance of TFP growth rates accounted for by the variance of the contributions from trade. The lines without markers add in the covariance between the two contributions.

Three lessons emerge. First, the ability of the model to account for TFP changes is greatest for intermediate values of the diffusion parameter, β. Second, the covariance terms are also quite large; countries whose TFP rose most saw increase stemming from trade but also from increasing the arrival ideas. This is consistent with the notion that some countries reformed along many margins, which both increased trade and increased R&D. Including this covariance, changes in trade cost can account for more than a third of the variation of changes in TFP (when β is roughly 0.6).

Third, when the contribution of trade is evaluated at the arrival rates drawn from data, trade accounts for more of the variance of TFP changes (either including or excluding the covariance). This happens because changes in trade costs and in the arrival rates of ideas are complementary. Intuitively, improvements in the quality of insights matter more when the arrival rate of these insights is greater.

5.6 Growth Miracles

To illustrate the fit of the model more concretely, we show how the model predicts changes in TFP during several growth miracles. We begin 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.

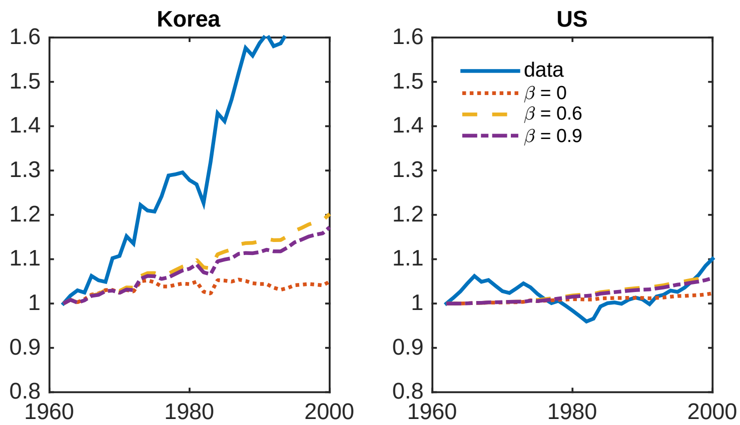

Figure 10: Openness and the Evolution of TFP: South Korea and the US

Note: This figure plots the changes in TFP for South Korea (top panels) and the US (bottom panels) under the specification of learning insights from sellers (left panels) and learning from producers (right panels). In each panel, we plot the actual change in TFP and changes in TFP generated by the model for various values of β. In all cases, TFP is detrended by the average growth rate of TFP in the US.

Figure 10 explores the implied dynamics of TFP under various assumptions. As in Figure 7, we assume that arrival rates of ideas are heterogenous, that, given trade costs, each country was on its balanced growth path in 1962, and that arrival rates have remained constant since 1962. Figure 10 shows the evolution of TFP for South Korea (left panel) and the US (right panel) for this case. 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 two lines illustrate the dynamic gains from trade implied by the model.

Two 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 right panel shows that changes in the dynamic gains from trade identified by the model are less relevant for understanding the growth experience of a developed country close to its balanced growth path.

Figure 11 shows the evolution of TFP for a larger set of Asian countries that experienced high

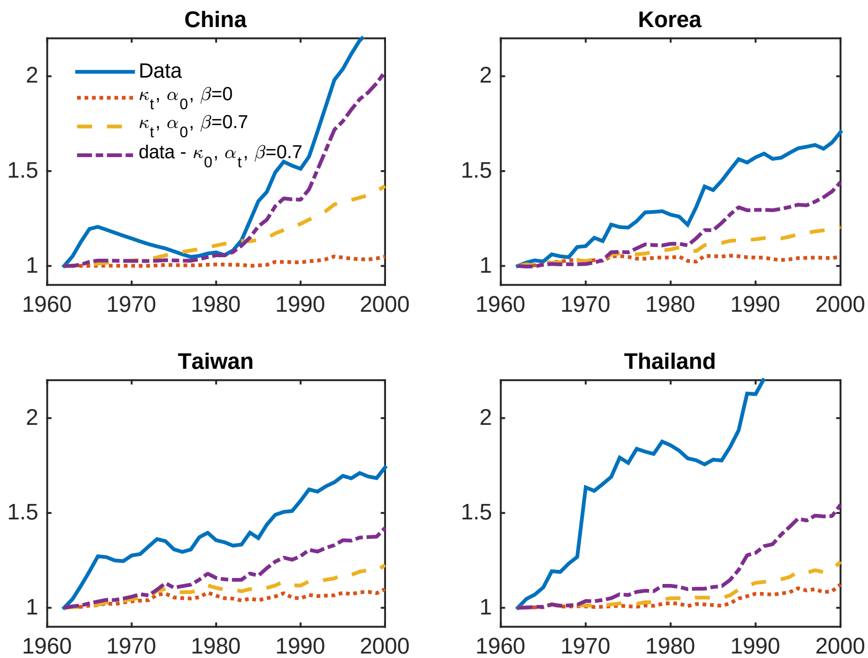

Figure 11: Growth Miracles, $\beta = 0.5$

Note: In all cases, TFP is detrended by the average growth rate of TFP in the US.

growth in the post-war period. For each country, the solid line is the data, while the dotted line is the model with $\beta = 0$ when trade costs are adjusted, but the arrival rates of ideas is held as in the 1962 values. The dashed line shows the evolution of TFP for the simple calibration of the diffusion parameter, $\beta = 0.7$. For some countries such as South Korea and China, the diffusion of ideas due to trade explains a substantial fraction of TFP growth. For others, such as Thailand changes in trade costs account for a smaller, but significant, fraction of TFP growth. Finally, the dashed dotted line shows the evolution of TFP netted of the contribution of changes in the arrival rate of ideas. The fact that this second measure gives a larger contribution of trade suggests strong complementarities between changes in trade costs and in the arrival rates of ideas, consistent with the results in Figure 9.

5.7 Explaining the Initial Distribution of TFP

We next assess the role of trade barriers in accounting for the initial cross-country TFP differences. To this end, we first make the extreme assumption that the arrival rate of ideas is the same in each country, $\hat{\alpha}_i = \hat{\alpha}$. Given the calibrated trade costs and a value of β, we solve for the balanced growth path of the model. In this case, trade is the only force driving TFP differences.

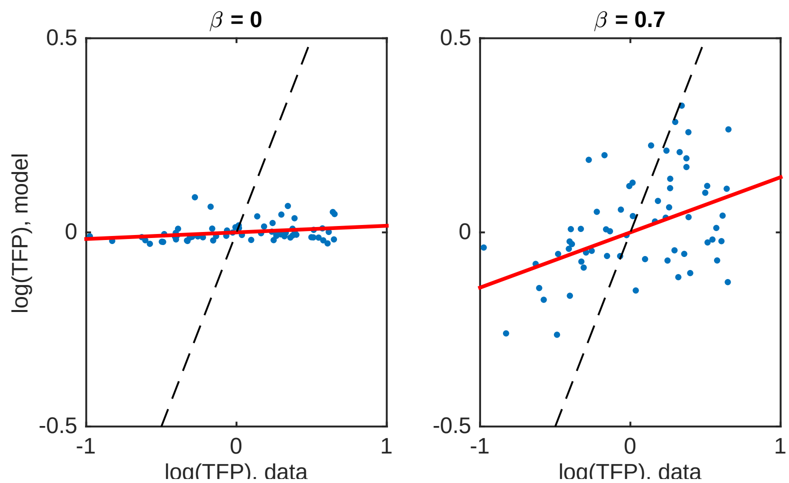

Figure 12: Openness and the Distribution of TFP in 1962

Note: Each panel plots countries actual TFP in the 1962 against the predicted TFP of the model under the assumption that the arrival rate of ideas is uniform across countries. The first panel assumes that $\beta = 0$. The second considers $\beta = 0.7$ and the case in which insights are drawn from sellers. In addition, each figure plots a dashed 45-degree line and a red regression line.

Figure 12 compares the implied distribution of TFP in the balance growth path of the model with the observed distribution of TFP in 1962, the first year in our sample. Each dot represents a country. If the model perfectly predicted each country's TFP, each dot would be on the (dashed) 45 degree line. The first panel shows the the case when $\beta = 0$, so that there is no cross-country diffusion of ideas and differences in countries' TFP represent only the static Ricardian gains from trade. As the panel shows, differences in openness generate only a small amount of cross-country differences in productivity.

The second panel assumes that $\beta = 0.7$ so that the cross-country TFP differences represent both the static and dynamic gains from trade. The model generates more variation in TFP across countries. The red regression line in each panel provides a simple measure of the average ability of the theory to account for the initial cross country differences in TFP. The positive slope implies a positive correlation between the model's predictions and the data.

We next assess more systematically how the strength of diffusion affects the ability of the model to account for cross-country TFP differences in Figure 13. For each model, we divide variation in TFP into a contribution from trade and a contribution from arrival rates of ideas. Given trade costs, we compute for the vector of arrival rates $\{\alpha_i\}$ so that the world would be on a balanced growth path. Given these, we can compute the $\bar{\kappa}$, a number so that if each bilateral iceberg trade cost were $\bar{\kappa}$, the volume of world trade would be unchanged. The contribution of trade to cross-sectional TFP differences is the variation that comes from the counterfactual of changing trade costs from κ_{ij} to $\bar{\kappa}$.

Similar to Section 5.5, we can do this in two ways. In (22) the contribution of trade is evaluated at common arrival rates $\bar{\alpha}$, while in (23) it is evaluated at the country specific arrival rates.

$$\ln \frac{TFP_i(\alpha_i, \kappa_{ij})}{TFP_i(\bar{\alpha}, \bar{\kappa})} = \underbrace{\ln \frac{TFP_i(\bar{\alpha}, \kappa_{ij})}{TFP_i(\bar{\alpha}, \bar{\kappa})}}_{\text{cont. from trade}} + \underbrace{\ln \frac{TFP_i(\alpha_i, \kappa_{ij})}{TFP_i(\bar{\alpha}, \kappa_{ij})}}_{\text{cont. from arrival rates}} \tag{22}$$

$$\ln \frac{TFP_i(\alpha_i, \kappa_{ij})}{TFP_i(\bar{\alpha}, \kappa)} = \underbrace{\ln \frac{TFP_i(\alpha_i, \bar{\kappa})}{TFP_i(\bar{\alpha}, \bar{\kappa})}}_{\text{cont. from arrival rates}} + \underbrace{\ln \frac{TFP_i(\alpha_i, \kappa_{ij})}{TFP_i(\alpha_i, \kappa_{ij})}}_{\text{cont. from trade}} \tag{23}$$

Each panel of Figure 13 has four lines. The two solid lines correspond to the decomposition in (22) in which the contribution of trade is evaluated holding the arrival rates of ideas fixed at a common level $\bar{\alpha}$. The two dashed lines correspond to (23) in which the contribution of trade is evaluated using country-specific arrival rates. The lines that are marked with squares represent the fraction of cross-sectional variance of TFP accounted for by the variance of the contributions from trade. The lines without markers add in the covariance between the two types of contributions. The left and right panels illustrate the ability of the theory to account for the cross section variance in 1962 and 2000, respectively.

When $\beta = 0$ so that there is no diffusion of ideas, the model accounts from roughly 2% to 8% of the variation, consistent with the first panel of Figure 12. As we consider specifications for which the strength of diffusion is larger, the model accounts for more of the variation in TFP. Interestingly, while the ability of the model to account for the initial differences in TFP initially

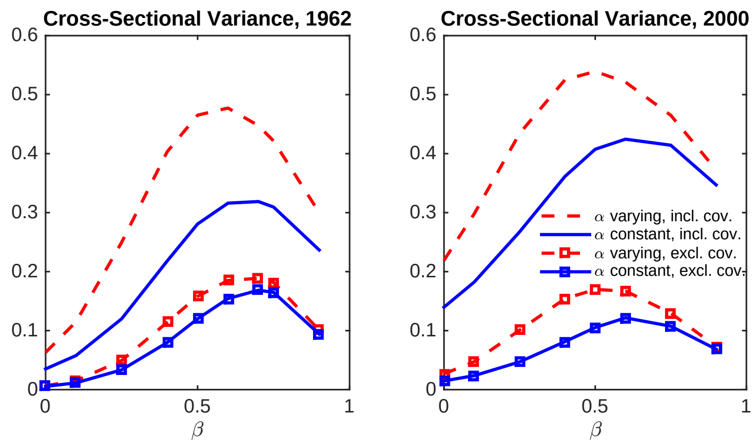

Figure 13: Mean squared error of predicted TFP

Note: This figure reports the fraction of the cross-sectional variance in log TFP in 1962 accounted for by trade, for various β, according to two decompositions. The solid lines correspond to (22) in which the contribution of trade is evaluated at a common arrival rate; the dashed lines to correspond to (23) in which the contribution of trade is evaluated at country-specific arrival rates. The lines with square markers report exclude the covariance between the contribution from trade and the contribution from variation in arrival rates of ideas; the lines without markers include the covariance. In all cases, insights are drawn from sellers.

rises with the strength of diffusion, it is greatest for cases with intermediate values of the diffusion parameter, β. As highlighted when discussing Figure 2, for β close to 1 a country's stock of ideas depends much more heavily on insights from the most productive producers, so that even countries close to autarky have accrued most of the dynamic gains from trade. Consequently when β is close to 1, the model does not predict much dispersion in TFP among countries that are moderately open. For a large range of values of β, however, (the lack of) openness accounts on average for up to 45% of the cross-sectional dispersion in TFP.

5.8 Counterfactual Trade Liberalizations: Niger in Belgium or Switzerland

To further illustrate the role of geography in determining productivity differences and the (potential) dynamics of the model, we consider two counterfactual experiments where we assign to Niger, one of the poorest countries in our sample, the trade costs faced by Belgium and Switzerland, two rich countries with populations of comparable sizes.[48] This exercise is presented in Figure 14.

On the right panel we plot the evolution of TFP, normalized by the TFP of the US. On impact

[48] While Niger's population is comparable to that of Belgium and Switzerland, 4% vs. 4% and 3% the US population, respectively, its endowment of equipped labor is an order of magnitude lower.

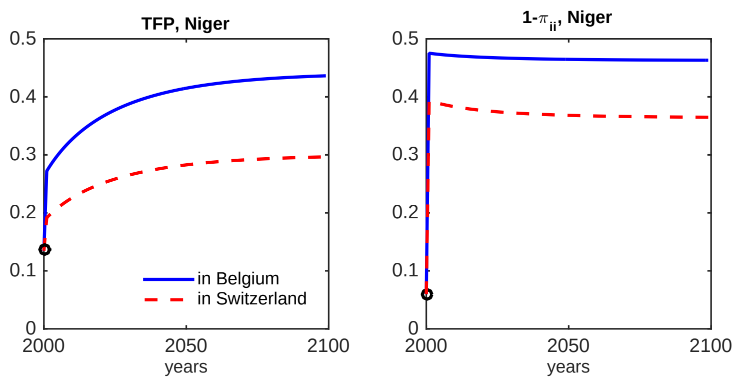

Figure 14: Transitional Dynamics after Assigning Belgium and Switzerland's Trade Costs to Niger.

the TFP jumps due to the static gains from trade, reflecting increased specialization and comparative advantage. Given that Niger is a very isolated economy, compared to the more integrated developed countries, the static gains lead to between 5 to 10 percentage points increases. Over time, as firms in Niger get to interact and getting insights from more productive foreign firms, TFP continues to grow. The second phase is more gradual, as it is mediated by the random arrival of insights. The dynamic gains a large, more than doubling the static gains. Overall, the theory predicts that over a century the productivity of Niger would increase from 15 percent of that of the US to between 30 and 45 percent.

On the right panel we show the evolution of Niger's foreign expenditure share in the counterfactual experiments. The sharp increase in the foreign expenditure share illustrate the large differences in trade cost characterizing poor economies.

6 Conclusion

In this paper we have provided a tractable theory of the cross-country diffusion of ideas across countries and provided a quantitative assessment of the role of trade in the transmission of knowledge. We found that when the model is specified so that the strength of diffusion of ideas is at an intermediate level, the model predicts a stronger response of TFP to changes in trade barriers than if the model were specified at either extreme of one of pure innovation or of pure diffusion. We show quantitatively that ...

The analysis points to critical importance of the parameter beta. While we provided one crude/heroic strategy to calibrate β, a more robust strategy would make better use of the variation in trade costs identified by Feyrer (2009b,a).

Of course we omitted many channels that may have complement or offset the role of trade in the diffusion of ideas. Chief among these are FDI and purposeful imitation. The spillovers from are modeled as an external effect, which likely reflects how some but not all ideas diffuse. In addition we have abstracted from variation across industries. Knowledge from one industry may be more useful in generating ideas to be used in the same industry than in other industries. In light of this, our quantitative results assessing the role of openness should be viewed as a first step than the final word.

A Technology Diffusion

Lemma 3 *Under Assumption* **??**, *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 $z^\theta [1 - H(z)] \leq (z^*)^\theta [1 - H(z)] \leq (z^*)^\theta$. Thus for any z, $\frac{1-H(z)}{z^{-\theta}} \leq K \equiv \max\left\{1 + \delta, (z^*)^\theta\right\}$ ■

Claim 4 *Suppose that Assumption* **??** *and Assumption* **??** *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 3, there is a $K < \infty$ such that for any z, $\frac{1-H(z)}{z^{-\theta}} \leq K$. 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 5 *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

$$\mathcal{F}_{ij}(q_1,q_2) = \int_0^{q_2}\prod_{k\neq j} F_k^{12}\left(\frac{w_k\kappa_{ik}x}{w_j\kappa_{ij}},\frac{w_k\kappa_{ik}x}{w_j\kappa_{ij}}\right)F_j^{12}(dx,x) + \int_{q_2}^{q_1}\prod_{k\neq j} F_k^{12}\left(\frac{w_k\kappa_{ik}q_2}{w_j\kappa_{ij}},\frac{w_k\kappa_{ik}q_2}{w_j\kappa_{ij}}\right)F_j^{12}(dx,q_2) \tag{24}$$

$\mathcal{F}_{ij}(q_1,q_2)$ 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 (24) 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

$$\begin{aligned}
&\int_0^\infty\int_{q_2}^\infty q_1^{\theta\tau_1}\min\left\{\frac{w_j\kappa_{ij}}{q_2},\frac{\varepsilon}{\varepsilon-1}\frac{w_j\kappa_{ij}}{q_1}\right\}^{-\theta\tau_2}\frac{\partial^2\mathcal{F}_{ij}\left(q_1,q_2\right)}{\partial q_1\partial q_2}dq_1dq_2 \\
=\ &\int_0^\infty\int_{q_2}^\infty q_1^{\theta\tau_1}\min\left\{\frac{w_j\kappa_{ij}}{q_2},\frac{\varepsilon}{\varepsilon-1}\frac{w_j\kappa_{ij}}{q_1}\right\}^{-\theta\tau_2}\frac{1}{\pi_{ij}}\left(\theta\lambda_j q_1^{-\theta-1}\right)\left(\theta\lambda_j q_2^{-\theta-1}\right)e^{-\frac{1}{\pi_{ij}}\lambda_j q_2^{-\theta}}dq_1dq_2
\end{aligned}$$

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

$$\left(w_j\kappa_{ij}\right)^{-\theta\tau_2}\pi_{ij}\left(\frac{\lambda_j}{\pi_{ij}}\right)^{\tau_1+\tau_2}\int_0^\infty\int_0^{x_2}x_1^{-\tau_1}\min\left\{x_2,\left(\frac{\varepsilon}{\varepsilon-1}\right)^\theta x_1\right\}^{-\tau_2}e^{-x_2}dx_1dx_2$$

Define $\tilde{B}(\tau_1,\tau_2)\equiv\int_0^\infty\int_0^{x_2}x_1^{-\tau_1}\min\left\{x_2,\left(\frac{\varepsilon}{\varepsilon-1}\right)^\theta x_1\right\}^{-\tau_2}e^{-x_2}dx_1dx_2$, 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 6 *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 5, 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 7 *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 5, 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 5, this is

$$\begin{aligned} \int_0^\infty q^{\beta\theta} dG_i(q) &= \sum_j \frac{1}{P_i^{1-\varepsilon}} \tilde{B}\left(\beta, \frac{\varepsilon-1}{\theta}\right) \left[\sum_k \lambda_k \left(w_k \kappa_{ik}\right)^{-\theta}\right]^{\frac{\varepsilon-1}{\theta}} \pi_{ij} \left(\frac{\lambda_j}{\pi_{ij}}\right)^\beta \\ &= \frac{\tilde{B}\left(\beta, \frac{\varepsilon-1}{\theta}\right)}{\tilde{B}\left(0, \frac{\varepsilon-1}{\theta}\right)} \sum_j \pi_{ij} \left(\frac{\lambda_j}{\pi_{ij}}\right)^\beta \end{aligned} \quad (25)$$

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 (25)) 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 5, this is

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

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 5, this is

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

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 5, 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 5, this is

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

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

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

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

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

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

Log-linearizing around the steady state gives

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

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

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

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

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

Learning from Producers Let $\Omega^P_{ij} \equiv \frac{r_{ji}\left(\hat{\lambda}_i/\pi_{ji}\right)^\beta}{\sum_k r_{ki}\left(\hat{\lambda}_i/\pi_{ki}\right)^\beta}$. 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 Research

This section proves the following claim:

Claim 8 *If $\Pi_{i\tau}$ is total flow of profit earned by entrepreneurs in i at time τ, then the flow of profit earned in i at time τ from ideas generated between t and t' (with $t < t' < \tau$) is:*

$$
\frac{\lambda_{it'} - \lambda_{it}}{\lambda_{i\tau}} \Pi_{i\tau}
$$

Proof. For $v_1 \leq v_2$, let $\tilde{\mathcal{V}}_{ji\tau}^{(t,t']}(v_1, v_2)$ be the probability that at time τ, the lowest cost technique to provide a variety to j was discovered by a manager in i between times t and t', that the marginal cost of that lowest-cost technique is no lower than v_1, and that marginal cost of the next-lowest-cost supplier is not lower than v_2. Just as in Appendix B, we will define $\mathcal{V}_{ji\tau}^{(t,t']}(v_1, v_2) = \lim_{m\to\infty} \tilde{\mathcal{V}}_{ji\tau}^{(t,t']}\left(m^{-\frac{1}{(1-\beta)\theta}} v_1, m^{-\frac{1}{(1-\beta)\theta}} v_2\right)$.

Let $\Pi_{i\tau}^{(t,t']}$ be profit from all techniques drawn in i between t and t'. Thus total profit in i at τ is $\Pi_{i\tau}^{(-\infty,\tau]}$. Defining $p(v_1,v_2) \equiv \min\left\{v_2, \frac{\varepsilon}{\varepsilon-1}v_1\right\}^{-\varepsilon}$, we can compute each of these by summing over profit form sales to each destiantion j:

$$\begin{aligned}
\Pi_{i\tau}^{(t,t']} &= \sum_j \int_0^\infty \int_{v_1}^\infty \left[p(v_1,v_2) - v_1\right] p(v_1,v_2)^{-\varepsilon} P_j^{\varepsilon-1} X_j \mathcal{V}_{ji\tau}^{(t,t']}(dv_1, dv_2) \\
\Pi_{i\tau}^{(-\infty,\tau]} &= \sum_j \int_0^\infty \int_{v_1}^\infty \left[p(v_1,v_2) - v_1\right] p(v_1,v_2)^{-\varepsilon} P_j^{\varepsilon-1} X_j \mathcal{V}_{ji\tau}^{(-\infty,\tau]}(dv_1, dv_2)
\end{aligned}$$

We will show below that $\mathcal{V}_{ji\tau}^{(t,t']}(v_1,v_2) = \frac{\lambda_{it'}-\lambda_{it}}{\lambda_{i\tau}}\mathcal{V}_{ji\tau}^{(-\infty,\tau]}(v_1,v_2)$. It will follow immediately that

$$\Pi_{i\tau}^{(t,t']} = \frac{\lambda_{it'} - \lambda_{it}}{\lambda_{i\tau}} \Pi_{i\tau}^{(-\infty,\tau]}$$

We now compute $\mathcal{V}_{ji\tau}^{(t,t']}(v_1,v_2)$. For each of the m managers in i, let $M_i^{(t,t']}(q)$ be the probability that the no technique drawn between t and t' delivers productivity better than q. Similarly, define $\tilde{F}_i^{(t,t']}(q) \equiv M_i^{(t,t']}(q)^m$ be the probability that none of the m managers drew a technique with productivity better than q between t and t'. Finally, let $F_i^{(t,t']}(q) = \lim_{m\to\infty} \tilde{F}_i^{(t,t']}\left(m^{\frac{1}{1-\beta}\frac{1}{\theta}}q\right)$. We have

$$\tilde{\mathcal{V}}_{ji\tau}^{(t,t']}(v_1,v_2) = \left\{\begin{array}{c}
m\int_{v_2}^\infty \left[\prod_{k\neq i}\tilde{F}_k^{(-\infty,\tau]}\left(\frac{w_k\kappa_{jk}}{x}\right)\right] M_i^{(-\infty,\tau]}\left(\frac{w_i\kappa_{ji}}{x}\right)^{m-1} M_i^{(-\infty,t]}\left(\frac{w_i\kappa_{ji}}{x}\right) \\
M_i^{(t',\tau]}\left(\frac{w_i\kappa_{ji}}{x}\right) dM_i^{(t,t']}\left(\frac{w_i\kappa_{ji}}{x}\right) \\
+m\int_{v_1}^{v_2} \left[\prod_{k\neq i}\tilde{F}_k^{(-\infty,\tau]}\left(\frac{w_k\kappa_{jk}}{v_2}\right)\right] M_i^{(-\infty,\tau]}\left(\frac{w_i\kappa_{ji}}{v_2}\right)^{m-1} M_i^{(-\infty,t]}\left(\frac{w_i\kappa_{ji}}{x}\right) \\
M_i^{(t',\tau]}\left(\frac{w_i\kappa_{ji}}{x}\right) dM_i^{(t,t']}\left(\frac{w_i\kappa_{ji}}{x}\right)
\end{array}\right\}$$

The expression for $\tilde{\mathcal{V}}_{ji\tau}^{(t,t']}(v_1,v_2)$ contains two terms. The first represents the probability that the best technique to serve j delivers marginal cost greater than v_2 and was drawn by a manager in i between t and t'. For each of the m managers in i, $dM_i^{(t,t']}\left(\frac{w_i\kappa_{ji}}{x}\right)$ measures the likelihood that the managers best draw between t and t' delivered marginal cost $x \in [v_2,\infty]$, $M_i^{(-\infty,t]}\left(\frac{w_i\kappa_{ji}}{x}\right) M_i^{(t',\tau]}\left(\frac{w_i\kappa_{ji}}{x}\right)$ is the probabiltiy that the manager had no better draws, and $\left[\prod_{k\neq i}\tilde{F}_k^{(-\infty,\tau]}\left(\frac{w_k\kappa_{jk}}{x}\right)\right] M_i^{(-\infty,\tau]}\left(\frac{w_i\kappa_{ji}}{x}\right)^{m-1}$ is the probability that no other manager from any destination would be able to provide the good at marginal cost lower than x. The second terms represents the proability that the best technique to serve j delivers marginal cost between v_1 and

v_2 and was drawn by a manager in i between t and t', and that no othe manager can delvier the variety with marginal cost lower than v_2.

Using $M_i^{(-\infty,t]}(q)\,M_i^{(t',\tau]}(q)\,M_i^{(t,t']}(q) = \tilde{F}_i^{(-\infty,\tau]}(q)$, $\tilde{F}_i^{(-\infty,\tau]}(q) = M_i^{(-\infty,\tau]}(q)^m$ and $\tilde{F}_i^{(t,t']}(q) = M_i^{(t,t']}(q)^m$, this can be written as

$$\tilde{\mathcal{V}}_{ji\tau}^{(t,t']}(v_1, v_2) = \left\{ \begin{array}{c} \int_{v_2}^{\infty} \left[\prod_k \tilde{F}_k^{(-\infty,\tau]}\left(\frac{w_k\kappa_{jk}}{x}\right)\right] \frac{d\tilde{F}_i^{(t,t']}\left(\frac{w_i\kappa_{ji}}{x}\right)}{\tilde{F}_i^{(t,t']}\left(\frac{w_i\kappa_{ji}}{x}\right)} \\ + \int_{v_1}^{v_2} \left[\prod_k \tilde{F}_k^{(-\infty,\tau]}\left(\frac{w_k\kappa_{jk}}{v_2}\right)\right] \left(\frac{\tilde{F}_i^{(-\infty,\tau]}\left(\frac{w_i\kappa_{ji}}{x}\right)}{\tilde{F}_i^{(-\infty,\tau]}\left(\frac{w_i\kappa_{ji}}{v_2}\right)}\right)^{\frac{1}{m}} \frac{d\tilde{F}_i^{(t,t']}\left(\frac{w_i\kappa_{ji}}{x}\right)}{\tilde{F}_i^{(t,t']}\left(\frac{w_i\kappa_{ji}}{x}\right)} \end{array} \right\}$$

Evaluating this at $m^{-\frac{1}{1-\beta}\frac{1}{\theta}}v_1$ and $m^{-\frac{1}{1-\beta}\frac{1}{\theta}}v_2$ and taking a limit as $m \to \infty$ gives

$$\mathcal{V}_{ji\tau}^{(t,t']}(v_1, v_2) = \left\{ \begin{array}{c} \int_{v_2}^{\infty} \left[\prod_k F_k^{(-\infty,\tau]}\left(\frac{w_k\kappa_{jk}}{x}\right)\right] \frac{dF_i^{(t,t']}\left(\frac{w_i\kappa_{ji}}{x}\right)}{F_i^{(t,t']}\left(\frac{w_i\kappa_{ji}}{x}\right)} \\ + \int_{v_1}^{v_2} \left[\prod_k F_k^{(-\infty,\tau]}\left(\frac{w_k\kappa_{jk}}{v_2}\right)\right] \frac{dF_i^{(t,t']}\left(\frac{w_i\kappa_{ji}}{x}\right)}{F_i^{(t,t']}\left(\frac{w_i\kappa_{ji}}{x}\right)} \end{array} \right\}$$

Finally, following the logic of Section 1, we have $F_i^{(t,t']}(q) = e^{-(\lambda_{it'}-\lambda_{it})q^{-\theta}}$, so that

$$\frac{dF_i^{(t,t']}\left(\frac{w_i\kappa_{ji}}{x}\right)}{F_i^{(t,t']}\left(\frac{w_i\kappa_{ji}}{x}\right)} = \frac{\lambda_{it'} - \lambda_{it}}{\lambda_{i\tau}} \frac{dF_i^{(-\infty,\tau]}\left(\frac{w_i\kappa_{ji}}{x}\right)}{F_i^{(-\infty,\tau]}\left(\frac{w_i\kappa_{ji}}{x}\right)}$$

We thus have

$$\mathcal{V}_{ji\tau}^{(t,t']}(v_1, v_2) = \frac{\lambda_{it'} - \lambda_{it}}{\lambda_{i\tau}} \mathcal{V}_{ji\tau}^{(-\infty,\tau]}(v_1, v_2)$$

which completes the proof. ■

F Quantitative Model

This appendix present expressions for the price index, expenditure shares and the law of motion of the stock of ideas for the extension of the model discussed in Section 5, incorporating non-tradable goods, intermediate inputs and equipped labor. The price index

$$p_i^{1-\varepsilon} =\propto (1-\mu)\left[\left(p_i^{\eta} w_i^{1-\eta}\right)^{-\theta}\lambda_i\right]^{-\frac{1-\varepsilon}{\theta}} + \mu\left[\sum_{j=1}^{n}\left(p_j^{\eta} w_j^{1-\eta}\kappa_{ij}\right)^{-\theta}\lambda_j\right]^{-\frac{1-\varepsilon}{\theta}}.$$

The fraction of income that country i expends in goods from country $j \neq i$

$$\pi_{ij} = \frac{\mu\left(p_j^{\eta} w_j^{1-\eta}\kappa_{ij}\right)^{-\theta}\lambda_j\left[\sum_{k=1}^{n}\left(p_k^{\eta} w_k^{1-\eta}\kappa_{ik}\right)^{-\theta}\lambda_k\right]^{-\frac{1-\varepsilon}{\theta}-1}}{(1-\mu)\left[\left(p_i^{\eta} w_i^{1-\eta}\right)^{-\theta}\lambda_i\right]^{-\frac{1-\varepsilon}{\theta}} + \mu\left[\sum_k\left(p_k^{\eta} w_k^{1-\eta}\kappa_{ik}\right)^{-\theta}\lambda_k\right]^{-\frac{1-\varepsilon}{\theta}}}.$$

The fraction of income that country i expends in its own goods is given by the sum of the non tradable and tradable shares

$$\pi_{ii} = \pi_i^{NT} + \pi_i^{T}$$

where the non tradable share

$$\pi_i^{NT} = \frac{(1-\mu)\left[\left(p_i^{\eta} w_i^{1-\eta}\right)^{-\theta}\lambda_i\right]^{-\frac{1-\varepsilon}{\theta}}}{(1-\mu)\left[\left(p_i^{\eta} w_i^{1-\eta}\right)^{-\theta}\lambda_i\right]^{-\frac{1-\varepsilon}{\theta}} + \mu\left[\sum_k\left(p_k^{\eta} w_k^{1-\eta}\kappa_{ik}\right)^{-\theta}\lambda_k\right]^{-\frac{1-\varepsilon}{\theta}}}$$

and the tradable (own) share

$$\pi_i^{T} = \frac{\mu\left(p_i^{\eta} w_i^{1-\eta}\right)^{-\theta}\lambda_i\left[\sum_k\left(p_k^{\eta} w_k^{1-\eta}\kappa_{ik}\right)^{-\theta}\lambda_k\right]^{-\frac{1-\varepsilon}{\theta}-1}}{(1-\mu)\left[\left(p_i^{\eta} w_i^{1-\eta}\right)^{-\theta}\lambda_i\right]^{-\frac{1-\varepsilon}{\theta}} + \mu\left[\sum_k\left(p_k^{\eta} w_k^{1-\eta}\kappa_{ik}\right)^{-\theta}\lambda_k\right]^{-\frac{1-\varepsilon}{\theta}}}.$$

The evolution of the stock of ideas when learning is from sellers

$$\dot{\lambda}_i \propto \pi_i^{NT}\lambda_i^{\beta} + \pi_i^{T}\left(\frac{\lambda_i}{\frac{\pi_i^{T}}{\pi_i^{T}+\sum_{k\neq i}\pi_{ik}}}\right)^{\beta} + \sum_{j\neq i}\pi_{ij}\left(\frac{\lambda_j}{\frac{\pi_{ij}}{\pi_i^{T}+\sum_{k\neq i}\pi_{ik}}}\right)^{\beta}.$$

The evolution of the stock of ideas when learning is from domestic producers

$$\dot{\lambda}_i \propto \pi_i^{NT} \lambda_i^{\beta} + \pi_i^{T} \left(\frac{\lambda_i}{\frac{\pi_i^T}{\pi_i^T + \sum_{k \neq i} \pi_{ik}}} \right)^{\beta} + \sum_{j \neq i} \frac{L_j w_j \pi_{ji}}{L_i w_i} \left(\frac{\lambda_i}{\frac{\pi_{ji}}{\pi_j^T + \sum_{k \neq j} \pi_{jk}}} \right)^{\beta}.$$

The market clearing conditions are the same as in the baseline model once labor is reinterpreted as equipped labor.

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.

BALDWIN, R. E. AND F. ROBERT-NICOUD (2008): "Trade and growth with heterogeneous firms," *Journal of International Economics*, 74, 21–34.

BEN-DAVID, D. (1993): "Equalizing exchange: trade liberalization and income convergence," *The Quarterly Journal of Economics*, 108, 653.

BENHABIB, J., J. PERLA, AND C. TONETTI (2014): "The Growth Dynamics of Innovation, Diffusion, and the Technology Frontier," Tech. rep., working paper.

BERNARD, A., J. EATON, J. JENSEN, AND S. KORTUM (2003): "Plants and Productivity in International Trade," *The American Economic Review*, 93, 1268–1290.

CHIU, J., C. MEH, AND R. WRIGHT (2011): "Innovation and Growth with Financial, and other, Frictions," Tech. rep., National Bureau of Economic Research.

COE, D. AND E. HELPMAN (1995): "International r&d spillovers," *European Economic Review*, 39, 859–887.

COE, D., E. HELPMAN, AND A. HOFFMAISTER (1997): "North-South spillovers," *Economic Journal*, 107, 134–149.

Comin, D. and B. Hobijn (2010): "An Exploration of Technology Diffusion," *American Economic Review*, 100, 2031–2059.

Comin, D. A., M. Dmitriev, and E. Rossi-Hansberg (2012): "The spatial diffusion of technology," Tech. rep., National Bureau of Economic Research.

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

Dollar, D. (1992): "Outward-oriented developing economies really do grow more rapidly: evidence from 95 LDCs, 1976-1985," *Economic development and cultural change*, 523–544.

Donaldson, D. (2015): "The Gains from Market Integration," *Annual Review of Economics*.

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

——— (2001): "Technology, trade, and growth: A unified framework," *European economic review*, 45, 742–755.

——— (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.

Frankel, J. A. and D. Romer (1999): "Does trade cause growth?" *American economic review*, 379–399.

Grossman, G. and E. Helpman (1991): "Trade, Knowledge Spillovers, and Growth," *European Economic Review*, 35, 517–526.

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.

Keller, W. (2009): "International Trade, Foreign Direct Investment, and Technology Spillovers," Working Paper 15442, National Bureau of Economic Research.

Klenow, P. J. and A. Rodriguez-Clare (2005): "Externalities and growth," *Handbook of economic growth*, 1, 817–861.

König, M., J. Lorenz, and F. Zilibotti (2012): "Innovation vs imitation and the evolution of productivity distributions," .

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.

Nelson, R. R. and E. S. Phelps (1966): "Investment in humans, technological diffusion, and economic growth," *The American Economic Review*, 69–75.

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

Parente, S. L. and E. C. Prescott (1994): "Barriers to technology adoption and development," *Journal of political Economy*, 298–321.

Pascali, L. (2014): "The wind of change: Maritime technology, trade and economic development," .

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.

Rivera-Batiz, L. A. and P. M. Romer (1991): "Economic Integration and Endogenous Growth," *The Quarterly Journal of Economics*, 106, 531–555.

Rodriguez, F. and D. Rodrik (2001): "Trade policy and economic growth: a skeptic's guide to the cross-national evidence," in *NBER Macroeconomics Annual 2000, Volume 15*, MIT Press, 261–338.

Sachs, J. and A. Warner (1995): "Economic reform and the process of global integration," *Brookings papers on economic activity*, 1995, 1–118.

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.

Wacziarg, R. and K. Welch (2008): "Trade liberalization and growth: New evidence," *The World Bank Economic Review*, 22, 187.

Waugh, M. E. (2010): "International Trade and Income Differences," *American Economic Review*, 100, 2093–2124.

Table 1: TFP Growth by Country, 1962-2000. Data vs. Models with $\beta = 0.5$.

	Data	Learning from Sellers			Learning
		Total Effect	Static Effect [a]	Trade Exposure[b]	from Producers
China	0.767	0.313	0.037	0.268	0.261
Thailand	0.745	0.204	0.090	0.107	0.249
Japan	0.465	0.008	0.001	0.002	-0.006
Taiwan	0.455	0.187	0.074	0.107	0.061
Ireland	0.453	0.164	0.117	0.043	0.292
South Korea	0.446	0.184	0.038	0.140	0.231
Israel	0.322	0.051	0.014	0.033	0.098
Greece	0.316	0.079	0.021	0.053	0.088
Sri Lanka	0.306	0.063	0.019	0.038	0.229
Egypt	0.280	0.126	0.001	0.119	0.202
Finland	0.238	0.094	0.028	0.062	0.104
Pakistan	0.208	0.022	-0.007	0.024	0.084
Tunisia	0.163	0.115	0.036	0.074	0.144
Belgium+Lux.	0.157	0.168	0.126	0.038	0.201
Norway	0.146	0.006	-0.002	0.004	0.031
Italy	0.114	0.047	0.020	0.024	0.053
UK	0.076	0.048	0.020	0.024	0.076
Malaysia	0.062	0.206	0.155	0.045	0.119
India	0.039	0.056	-0.007	0.058	0.067
Mozambique	0.029	0.183	0.035	0.140	0.416
Denmark	0.024	-0.002	0.004	-0.010	-0.001
Australia	0.022	0.051	0.014	0.033	0.027
Turkey	0.022	0.054	0.019	0.032	0.014
Portugal	0.021	0.100	0.041	0.056	0.078
Austria	0.020	0.090	0.034	0.053	0.078
Indonesia	0.002	0.112	0.017	0.089	0.074
USA	0.000	0.057	0.019	0.035	0.038
Netherlands	-0.024	0.033	0.023	0.006	-0.016
Sweden	-0.026	0.043	0.024	0.015	0.084
France	-0.035	0.071	0.030	0.038	0.074
Mali	-0.085	-0.053	-0.025	-0.033	0.127
Spain	-0.112	0.066	0.031	0.032	0.040
New Zealand	-0.136	0.087	0.022	0.062	0.087

[a] The static effect is given by the change in TFP caused by the change in the own trade share, i.e., $\Delta \log TFP(\lambda_0, \pi_{iit})$.

[b] The effect of the trade exposure is given by the change in TFP caused by the change in the trade exposure, holding fixed the stock of ideas of all the other countries, i.e., $\Delta \log TFP(\lambda_t(\lambda^0_{-i}, \Pi^t), \pi_{ii0})$.

TFP Growth by Country, 1962-2000. Data vs. Models with $\beta = 0.5$ (cont'd).

	Data	Learning from Sellers			Learning
		Total Effect	Static Effect [a]	Trade Exposure[b]	from Producers
Germany	-0.170	0.052	0.025	0.023	0.045
Ecuador	-0.180	0.041	0.014	0.024	-0.015
Canada	-0.189	0.073	0.045	0.025	0.065
Brazil	-0.197	0.007	0.001	0.002	-0.026
Tanzania	-0.203	0.128	-0.002	0.124	-0.025
Chile	-0.206	0.067	0.025	0.039	0.155
Switzerland	-0.258	0.003	0.009	-0.009	-0.017
Argentina	-0.259	0.009	0.002	0.003	0.024
Dom. Rep.	-0.262	0.034	-0.002	0.033	-0.012
Syria	-0.265	0.013	-0.004	0.013	-0.023
Guatemala	-0.271	0.049	0.004	0.042	0.078
Costa Rica	-0.288	0.027	-0.004	0.028	0.061
Morocco	-0.297	0.061	0.017	0.040	0.066
Cote d'Ivoire	-0.298	0.015	-0.013	0.024	-0.017
Uruguay	-0.302	0.041	0.014	0.025	0.285
Colombia	-0.337	0.019	0.004	0.012	0.054
Mexico	-0.400	0.066	0.043	0.021	0.055
Senegal	-0.400	0.034	0.002	0.027	-0.020
Kenya	-0.419	0.015	-0.015	0.025	-0.073
Cameroon	-0.499	0.002	-0.009	0.006	-0.020
Uganda	-0.532	0.122	-0.005	0.124	0.123
Paraguay	-0.549	0.045	0.020	0.021	-0.039
Ghana	-0.550	-0.007	0.003	-0.014	0.053
Philippines	-0.555	0.085	0.042	0.037	0.065
Jamaica	-0.565	-0.002	-0.014	0.008	0.053
Bolivia	-0.604	-0.008	-0.006	-0.005	0.087
South Africa	-0.632	0.019	-0.001	0.017	-0.024
Zambia	-0.634	-0.155	-0.080	-0.078	-0.127
Peru	-0.639	-0.082	-0.026	-0.060	-0.153
Honduras	-0.644	0.075	-0.002	0.072	0.122
Niger	-0.911	0.064	0.004	0.056	0.164
Jordan	-1.089	0.000	-0.026	0.023	0.069

[a] The static effect is given by the change in TFP caused by the change in the own trade share, i.e., $\Delta \log TFP(\lambda_0, \pi_{iit})$.

[b] The effect of the trade exposure is given by the change in TFP caused by the change in the trade exposure, holding fixed the stock of ideas of all the other countries, i.e., $\Delta \log TFP(\lambda_t(\lambda^0_{-i}, \Pi^t), \pi_{ii0})$.