The Global Diffusion of Ideas
Supplementary Appendix

Francisco J. Buera
Washington University in St. Louis

Ezra Oberfield
Princeton University

September 18, 2019

C Omitted Proofs

C.1 Proof of Lemma 2

Lemma 3 *Suppose τ_1 and τ_2 satisfy $\tau_1 + \tau_2 < 1$. Suppose also that for each country, $F_i^{12}(q_1, q_2) = \left[1 + \lambda_i q_2^{-\theta} - \lambda_i q_1^{-\theta}\right] e^{-\lambda_i q_2^{-\theta}}$. Then*

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

where $B(\tau_1, \tau_2) \equiv \left\{1 - \frac{\tau_2}{1-\tau_1} + \frac{\tau_2}{1-\tau_1}\left(\frac{\varepsilon}{\varepsilon - 1}\right)^{-\theta(1-\tau_1)}\right\} \Gamma\left(1 - \tau_1 - \tau_2\right)$ and $\pi_{ij} \equiv \frac{\lambda_j (w_j \kappa_{ij})^{-\theta}}{\sum_k \lambda_k (w_k \kappa_{ik})^{-\theta}}$.

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

$$\begin{aligned} \mathcal{F}_{ij}(q_1, q_2) = & \int_0^{q_2} \prod_{k \neq j} F_k^{12}\left(\frac{w_k \kappa_{ik} x}{w_j \kappa_{ij}}, \frac{w_k \kappa_{ik} x}{w_j \kappa_{ij}}\right) F_j^{12}(dx, x) \\ & + \int_{q_2}^{q_1} \prod_{k \neq j} F_k^{12}\left(\frac{w_k \kappa_{ik} q_2}{w_j \kappa_{ij}}, \frac{w_k \kappa_{ik} q_2}{w_j \kappa_{ij}}\right) F_j^{12}(dx, q_2). \end{aligned} \quad (1)$$

The measure $\mathcal{F}_{ij}(q_1, q_2)$ gives 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 goods 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 (1) can be written as

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

The second term is

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

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

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

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

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

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

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

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

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

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

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

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

C.2 The $\beta \nearrow 1$ Limit

Claim 4 *Suppose trade shares are interior. For the learning-from-sellers specification,*

$$\lim_{\beta \nearrow 1} d\ln\hat{\lambda}_i = \frac{\sum_j \sum_k \hat{\lambda}_j \hat{\lambda}_k d\ln\pi_{jk}}{\sum_j \sum_k \hat{\lambda}_j \hat{\lambda}_k}$$

Proof. Define $\Omega_{ij}(\beta) = \frac{\pi_{ij}^{1-\beta}\hat{\lambda}_j^{\beta}}{\sum_k \pi_{ik}^{1-\beta}\hat{\lambda}_k^{\beta}}$. Note that $\Omega_{ij}(1) = \frac{\hat{\lambda}_j}{\sum_k \hat{\lambda}_k}$. Differentiating (10) yields

$$d\ln\hat{\lambda}_i = \sum_j \Omega_{ij}(\beta)\left[(1-\beta)\,d\ln\pi_{ij} + \beta d\ln\hat{\lambda}_j\right] \tag{2}$$

Taking the limit as $\beta \nearrow 1$ gives

$$\begin{aligned}\lim_{\beta \nearrow 1} d\ln\hat{\lambda}_i &= \sum_j \Omega_{ij}(1) \lim_{\beta \nearrow 1} d\ln\hat{\lambda}_j \\ &= \frac{\sum_j \hat{\lambda}_j \lim_{\beta \nearrow 1} d\ln\hat{\lambda}_j}{\sum_k \hat{\lambda}_k}\end{aligned}$$

Note that this does not depend on i. Thus for every i,j, $\lim_{\beta \nearrow 1} d\ln\hat{\lambda}_i = \lim_{\beta \nearrow 1} d\ln\hat{\lambda}_j$. Call this common change D. Next, we find D. To do this, we can multiply (2) by $\hat{\lambda}_i$ and sum over i to get

$$\sum_i \hat{\lambda}_i d\ln\hat{\lambda}_i = (1-\beta)\sum_i \hat{\lambda}_i \sum_j \Omega_{ij}(\beta) d\ln\pi_{ij} + \beta \sum_i \hat{\lambda}_i \sum_j \Omega_{ij}(\beta) d\ln\hat{\lambda}_j$$

Subtracting $\beta \sum_i \lambda_i \sum_j \Omega_{ij}(\beta) d\ln\hat{\lambda}_j$ from each side and dividing by $(1-\beta)\sum_i \hat{\lambda}_i$ yields

$$\frac{1}{\sum_i \hat{\lambda}_i} \frac{\sum_i \hat{\lambda}_i d\ln\hat{\lambda}_i - \beta \sum_i \hat{\lambda}_i \sum_j \Omega_{ij}(\beta) d\ln\hat{\lambda}_j}{1-\beta} = \frac{\sum_i \hat{\lambda}_i \sum_j \Omega_{ij}(\beta) d\ln\pi_{ij}}{\sum_i \hat{\lambda}_i}$$

The limit as $\beta \nearrow 1$ of the right hand side can be expressed as

$$\lim_{\beta \nearrow 1} \frac{\sum_i \hat{\lambda}_i \sum_j \Omega_{ij}(\beta) d\ln\pi_{ij}}{\sum_i \hat{\lambda}_i} = \frac{\sum_i \hat{\lambda}_i \sum_j \Omega_{ij}(1) d\ln\pi_{ij}}{\sum_i \hat{\lambda}_i} = \frac{\sum_i \sum_j \hat{\lambda}_i \hat{\lambda}_j d\ln\pi_{ij}}{\sum_i \hat{\lambda}_i \sum_k \hat{\lambda}_k}$$

We complete the proof by showing that the limit of the left hand side is D. Using L'Hospital's rule, this limit can be expressed as

$$\lim_{\beta\nearrow 1} LHS = \frac{1}{\sum_i \hat{\lambda}_i} \lim_{\beta\nearrow 1} \frac{\sum_i \hat{\lambda}_i d\ln\hat{\lambda}_i(\beta) - \beta \sum_i \hat{\lambda}_i \sum_j \Omega_{ij}(\beta) d\ln\hat{\lambda}_j(\beta)}{1-\beta}$$

$$= \frac{1}{\sum_i \hat{\lambda}_i} \lim_{\beta\nearrow 1} \left\{ \begin{array}{c} -\sum_i \hat{\lambda}_i \frac{d\left[d\ln\hat{\lambda}_i(\beta)\right]}{d\beta} + \sum_i \hat{\lambda}_i \sum_j \Omega_{ij}(\beta) d\ln\hat{\lambda}_j(\beta) \\ +\beta \sum_i \hat{\lambda}_i \sum_j \frac{d\Omega_{ij}(\beta)}{d\beta} d\ln\hat{\lambda}_j(\beta) + \beta \sum_i \hat{\lambda}_i \sum_j \Omega_{ij}(\beta) \frac{d\left[d\ln\hat{\lambda}_j(\beta)\right]}{d\beta} \end{array} \right\}$$

$$= \frac{1}{\sum_i \hat{\lambda}_i} \left\{ \begin{array}{c} -\sum_i \hat{\lambda}_i \lim_{\beta\nearrow 1} \frac{d\left[d\ln\hat{\lambda}_i(\beta)\right]}{d\beta} + \sum_i \hat{\lambda}_i \sum_j \Omega_{ij}(1) \lim_{\beta\nearrow 1} d\ln\hat{\lambda}_j(\beta) \\ +\sum_i \hat{\lambda}_i \sum_j \lim_{\beta\nearrow 1} \frac{d\Omega_{ij}(\beta)}{d\beta} \lim_{\beta\nearrow 1} d\ln\hat{\lambda}_j(\beta) + \sum_i \hat{\lambda}_i \sum_j \Omega_{ij}(1) \lim_{\beta\nearrow 1} \frac{d\left[d\ln\hat{\lambda}_j(\beta)\right]}{d\beta} \end{array} \right\}$$

Using $D = \lim_{\beta\nearrow 1} d\ln\hat{\lambda}_j(\beta)$ and $\Omega_{ij}(1) = \frac{\hat{\lambda}_j}{\sum_k \hat{\lambda}_k}$, this is

$$\lim_{\beta\nearrow 1} LHS = \frac{1}{\sum_i \hat{\lambda}_i} \left\{ \begin{array}{c} -\sum_i \hat{\lambda}_i \lim_{\beta\nearrow 1} \frac{d\left[d\ln\hat{\lambda}_i(\beta)\right]}{d\beta} + \sum_i \hat{\lambda}_i \sum_j \frac{\hat{\lambda}_j}{\sum_k \hat{\lambda}_k} D \\ +\sum_i \hat{\lambda}_i \sum_j \lim_{\beta\nearrow 1} \frac{d\Omega_{ij}(\beta)}{d\beta} D + \sum_i \hat{\lambda}_i \sum_j \frac{\hat{\lambda}_j}{\sum_k \hat{\lambda}_k} \lim_{\beta\nearrow 1} \frac{d\left[d\ln\hat{\lambda}_j(\beta)\right]}{d\beta} \end{array} \right\}$$

$$= \frac{1}{\sum_i \hat{\lambda}_i} \left\{ \sum_i \hat{\lambda}_i + \sum_i \hat{\lambda}_i \sum_j \lim_{\beta\nearrow 1} \frac{d\Omega_{ij}(\beta)}{d\beta} \right\} D$$

$$= \left\{ 1 + \frac{\sum_i \hat{\lambda}_i \sum_j \lim_{\beta\nearrow 1} \frac{d\Omega_{ij}(\beta)}{d\beta}}{\sum_i \hat{\lambda}_i} \right\} D$$

Finally, using $\sum_j \Omega_{ij}(\beta) = 1$, we have $\sum_j \lim_{\beta\nearrow 1} \frac{d\Omega_{ij}(\beta)}{d\beta} = \lim_{\beta\nearrow 1} \frac{d}{d\beta} \left(\sum_j \Omega_{ij}(\beta) \right) = 0$. As a result, $\lim_{\beta\nearrow 1} LHS = D$.

■

D 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, labor generates ideas. The labor resource constraint in country i at t is thus

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

where L_{it}^P is labor used in production and L_{it}^R is 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 new ideas arrive in proportion to labor engaged in research. Specifically, if a measure of L_{it}^R of labor is engaged in research then ideas with original component greater than z arrive independently across goods at Poisson rate $\tilde{A}(z)L_{it}^R = \tilde{\alpha}z^{-\theta}L_{it}^R$. Thus the arrival of goods is uniform across goods; research effort is not directed at particular goods.[1] If a new idea arrives for a particular good, the researcher matches with a random potential producer of the good and makes a take-it-or-leave-it offer to the potential producer.[2]

Suppose also that potential producers 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} be the expected pretax value of all ideas generated in i at t. Since the payoff to research is linear, an interior equilibrium requires that

$$w_{it}L_{it}^R \geq (1 - T_i)V_{it}.$$

We next compute the expected pretax value of research at t, V_{it}. In the next subsection, 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 $\frac{\lambda_{it'} - \lambda_{it}}{\lambda_{i\tau}}\Pi_{i\tau}$. Put differently, 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.[3]

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 profit from ideas generated in i at instant t is

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

where $e^{-\rho(\tau-t)}\frac{P_{it}}{P_{i\tau}}$ is the real discount factor between t and τ. As we show in Claim 3 in Appendix A.2, profit among all entrepreneurs is proportional to the wage bill in production, $\Pi_{i\tau} = \frac{w_{i\tau}L_{i\tau}^P}{\theta}$. Together these imply that equilibrium research intensity satisfies

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

[1] It is not crucial that each researcher might generate ideas for *all* goods. We could just as easily have assumed that each researcher can generate ideas for 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.

[2] It would be equivalent to assume that the researcher now has the possibility of producing the good herself.

[3] While 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).

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 a 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:

$$\frac{r_i}{1-r_i} = \frac{1-T_i}{\theta\left[(1-\beta)\frac{\rho}{\gamma}+\beta\right]-1}. \tag{3}$$

Equation (3) 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 of 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.[4]

Finally, we define

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

where r_{it} is defined in (3) and depends on the country-specific distortion to R&D effort.

D.1 Research and Profit

In this section we prove the following claim:

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

Claim 5 *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 $\mathcal{V}_{ji\tau}^{(t,t']}(v_1, v_2)$ be the probability that at time τ, the lowest cost technique to provide a variety to j was discovered by a researcher 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.

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 from sales to each destination j:

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

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 follows immediately that

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

We now compute $\mathcal{V}_{ji\tau}^{(t,t']}(v_1, v_2)$. Let $F_i^{(t,t']}(q)$ be the probability that none of the researchers in i drew a technique with productivity better than q between t and t'. Similarly, let $F_i^{12,(t,t']}(q_1, q_2)$ be the probability that between t and t', no researcher in i drew an idea better than q_1 and at most one idea was better than q_2. $\mathcal{V}_{ji\tau}^{(t,t']}(v_1, v_2)$ can be expressed as

$$\begin{aligned}
\mathcal{V}_{ji\tau}^{(t,t']}(v_1, v_2) = &\int_{v_2}^\infty \left[\prod_{k \neq i} F_k^{(-\infty,\tau]}\left(\frac{w_k \kappa_{jk}}{x}\right)\right] F_i^{(-\infty,t]}\left(\frac{w_i \kappa_{ji}}{x}\right) F_i^{(t',\tau]}\left(\frac{w_i \kappa_{ji}}{x}\right) \frac{dF_i^{(t,t']}\left(\frac{w_i \kappa_{ji}}{x}\right)}{dx} dx \\
&+ \int_{v_1}^{v_2} \left[\prod_{k \neq i} F_k^{(-\infty,\tau]}\left(\frac{w_k \kappa_{jk}}{v_2}\right)\right] F_i^{(-\infty,t]}\left(\frac{w_i \kappa_{ji}}{v_2}\right) F_i^{(t',\tau]}\left(\frac{w_i \kappa_{ji}}{v_2}\right) \\
&\qquad \frac{dF_i^{12,(t,t']}\left(\frac{w_i \kappa_{ji}}{x}, \frac{w_i \kappa_{ji}}{v_2}\right)}{dx} dx.
\end{aligned}$$

The expression for $\mathcal{V}_{ji\tau}^{(t,t']}(v_1, v_2)$ 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 researcher in i between t and t'. $\frac{d}{dx}F_i^{(t,t']}\left(\frac{w_i\kappa_{ji}}{x}\right)$ measures the likelihood that the best idea between t and t' delivered marginal cost $x \in [v_2, \infty)$, $F_i^{(-\infty,t]}\left(\frac{w_i\kappa_{ji}}{x}\right) F_i^{(t',\tau]}\left(\frac{w_i\kappa_{ji}}{x}\right)$ is the probability that no ideas drawn at other times delivered a cost lower than x, and $\left[\prod_{k\neq i} F_k^{(-\infty,\tau]}\left(\frac{w_k\kappa_{jk}}{x}\right)\right]$ is the probability that no researcher from any other country drew an idea that would provide the good to j at marginal cost lower than x. The second term represents the probability that the best technique to serve j delivers marginal cost $x \in [v_1, v_2)$ and was drawn by a researcher in i between t and t', and that no other idea can deliver the variety to j with marginal cost lower than v_2.

This can be rearranged as

$$\mathcal{V}_{ji\tau}^{(t,t']}(v_1, v_2) = \int_{v_2}^{\infty} \left[\prod_k F_k^{(-\infty,\tau]}\left(\frac{w_k\kappa_{jk}}{x}\right)\right] \frac{\frac{d}{dx}F_i^{(t,t']}\left(\frac{w_i\kappa_{ji}}{x}\right)}{F_i^{(t,t']}\left(\frac{w_i\kappa_{ji}}{x}\right)} dx + \int_{v_1}^{v_2} \left[\prod_k F_k^{(-\infty,\tau]}\left(\frac{w_k\kappa_{jk}}{v_2}\right)\right] \frac{\frac{d}{dx}F_i^{12,(t,t']}\left(\frac{w_i\kappa_{ji}}{x}, \frac{w_i\kappa_{ji}}{v_2}\right)}{F_i^{(t,t']}\left(\frac{w_i\kappa_{ji}}{v_2}\right)} dx.$$

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

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

and following the logic of Appendix A.1, we have

$$F_i^{12,(t,t']}(q_1, q_2) = \left(1 + (\lambda_{t'} - \lambda_t)(q_2^{-\theta} - q_1^{-\theta})\right) e^{-(\lambda_{it'}-\lambda_{it})q_2^{-\theta}}$$

so that

$$\frac{\frac{d}{dx}F_i^{12,(t,t']}\left(\frac{w_i\kappa_{ji}}{x}, \frac{w_i\kappa_{ji}}{v_2}\right)}{F_i^{(t,t']}\left(\frac{w_i\kappa_{ji}}{v_2}\right)} = \frac{\lambda_{it'} - \lambda_{it}}{\lambda_{i\tau}} \frac{\frac{d}{dx}F_i^{12,(-\infty,\tau]}\left(\frac{w_i\kappa_{ji}}{x}, \frac{w_i\kappa_{ji}}{v_2}\right)}{F_i^{(-\infty,\tau]}\left(\frac{w_i\kappa_{ji}}{v_2}\right)}.$$

Consequently

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

which completes the proof. ■

E Quantitative Exploration: Additional Exercises

This section presents several additional exercises: (i) We consider an alternative calibration strategy for the arrival rate of ideas, using data on R&D intensity; (ii) we consider an alternative model of non-tradable goods in which the productivity of these goods is exogenously given and, therefore, trade only impact the productivity of traded goods; (iii) we present results for learning from producers, under the baseline assumptions that insights are drawn uniformly; (iv) we explore variations of both specifications of learning using different weights; for learning from sellers, we allow insights to be weighted in proportion to consumption or in proportion to expenditure, while for learning from producers we allow insights to be weighted by those producers' employment; (v) we explore the extension with targeted learning; (vi) we explore the extension in which individuals draw more insights when exposed to a wider variety of ideas; (vii) we present counterfactual evolutions of the stock of knowledge and, therefore, counterfactual TFP series for each country to illustrate the role of changes in the trade exposure towards more productive trading partners; (viii) we calibrate the model more directly to measures of PPP.

E.1 Calibrating α with R&D Data

In our main results we calibrated the evolution of the arrival rate of ideas $\hat{\alpha}_t$ to match the evolution of TFP given the dynamics of trade costs. In doing this, we took the strong stand that all residual TFP differences are due to differences in stocks of knowledge and trade barriers. In this section we consider a calibration using information on R&D and human capital stocks to provide an alternative measure of stocks of knowledge.

In particular, for the initial period, we project log TFP onto R&D intensity, the log of the human capital stock, the log of the own trade share, and the log of an imported-weighted average of trading partners' TFP. We assign the residual initial TFP from this regression to a neutral productivity terms affecting the units of equipped labor and not the stock of knowledge, and choose initial stocks of knowledge to match predicted TFP from these regressions. The knowledge-driven TFP accounts for 31 percent of the variation of log TFP (as measured by R-squared).

To calibrate the evolution of the arrival rate of ideas in the subsequent periods we proceed in three steps. First, initial arrival rates of ideas are chosen to exactly match initial stocks of knowledge, taking as given the initial trade flows, and assuming that the world economy was on a balanced growth path in 1962. Second, we project the initial calibrated value for the arrival rate of ideas on initial R&D intensity and the human capital stock. Third, we use the resulting coefficients and the data on the evolution of R&D intensity and the human capital stock to predict the evolution of arrival rates of ideas. The evolution of stocks of knowledge then depends on the evolution of the arrival rates of ideas and trade costs. As before, we assign the residual evolution of TFP to a neutral productivity terms affecting the units of equipped labor and not the stock of knowledge. We denote the part of TFP accounted for by the evolution of stocks of knowledge the *knowledge-driven TFP growth*.[5]

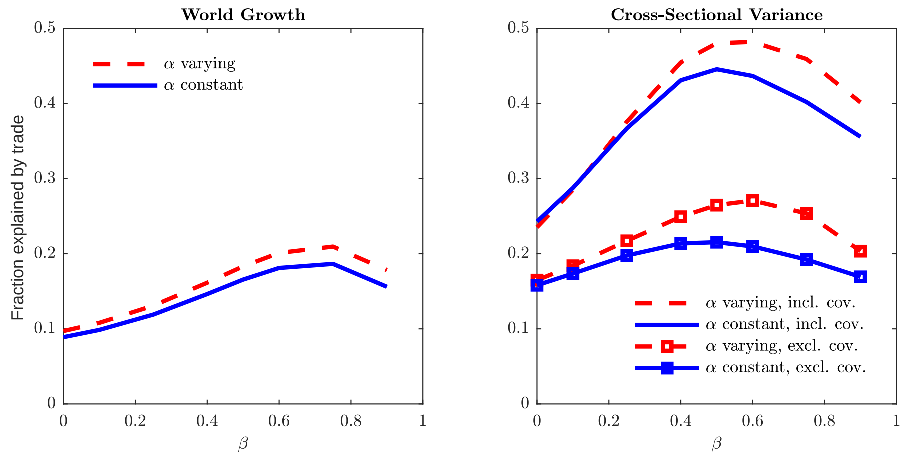

Figure OA.1: The contribution of changes in trade costs to changes in knowledge-driven TFP.

Note: This figure reports the contribution of trade to knowledge-driven TFP for various values of β, according to two decompositions described in Section 4.4. 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. See the note to Figure 6 for details.

Figure OA.1 reports the contribution of trade to knowledge-driven TFP. This figure is the counterpart of Figure 6. The message in the main body of the paper is preserved, with trade explaining a bit less of the level and a bit more of the variance of knowledge-driven TFP growth.

E.2 Alternative Modeling of Non-Tradable Goods

In the main text we modeled non-traded goods simply as a subset of the differentiated goods that is subject to infinite trade costs. As such, these goods are only produced domestically. Nevertheless, the productivity of these goods is affected by the insights obtained from domestic and foreign producers of other goods. In this section we consider an alternative model of non-traded goods in which the productivity of these goods is exogenously given and, therefore, trade can only impact the productivity of traded goods.

In particular, we consider a version of the model along the lines of Alvarez and Lucas (2007) where all the differentiated goods are tradable but final consumption is produced with a combination of non-tradable consumption C_i^N and the differentiated tradable aggregate

[5] We use the data on R&D expenditure to GDP in Lederman and Saenz (2005).

C_i^T,

$$U\left(C_i^N, C_i^T\right) = \left(C_i^N\right)^{\upsilon}\left(C_i^T\right)^{1-\upsilon}$$

where non-tradable output is produced with a labor only technology

$$C_i^N = A_i^N L_i^N$$

and the tradable consumption aggregate is produced using the differentiated intermediate inputs as in the benchmark model.

An equilibrium with trade balance is given by wages that solve

$$(1-\alpha)\, w_i L_i = \sum_j \pi_{ji}\,(1-\alpha)\, w_j L_j,\ \ i = 1, ..., n, \tag{4}$$

where the share of j's tradable spending on goods from country i is

$$\pi_{ji} = \frac{\left((p_i^T)^{\eta} w_i^{1-\eta}\kappa_{ji}\right)^{-\theta}\lambda_i}{\sum_k \left((p_k^T)^{\eta} w_k^{1-\eta}\kappa_{jk}\right)^{-\theta}\lambda_k}, \tag{5}$$

the price of the intermediate aggregate p_i^T in country i is

$$p_i^T \propto \left[\sum_k \left((p_k^T)^{\eta} w_k^{1-\eta}\kappa_{ik}\right)^{-\theta}\lambda_k\right]^{-\frac{1}{\theta}}, \tag{6}$$

and the price of aggregate consumption in each country i is

$$p_i = \frac{\left(p_i^T\right)^{1-\alpha}\left(p_i^N\right)^{\alpha}}{(1-\alpha)^{1-\alpha}\alpha^{\alpha}}, \tag{7}$$

where the price of non-tradable consumption is

$$p_i^N = \frac{w_i}{A_i^N}. \tag{8}$$

The evolution of the stock of knowledge is as in the simple model in Section 2.1,

$$\dot{\lambda}_i \propto \sum_j \pi_{ij}\left(\frac{\lambda_j}{\pi_{ij}}\right)^{\beta}. \tag{9}$$

In this version of the model both static and dynamics gains from trade accrue solely to the tradable sector, and the analysis of the equilibrium can be decoupled in a convenient way. Equations (4), (5), (6) and (9) can be used to solve for the vector of world wages, trade patterns, the price of the tradable aggregate, and the evolution of the stock of knowledge. Given wages, the price of the tradable aggregate, and the non-tradable productivity, equations (7) and (8) can be used to solve for the price of aggregate consumption and real income.

To quantify the role of trade costs in explaining (tradable) productivity, a similar strategy can be used to calibrate the evolution of trade costs and the arrival rate of ideas. In particular, the stock of knowledge of country i satisfies

$$\lambda_i \propto \left(\frac{w_i}{p_i^T}\right)^{(1-\eta)\theta} \pi_{ii},$$

and the iceberg costs of shipping goods from j to i is given by

$$\kappa_{ij} = \frac{p_i^T}{p_j^T}\left(\frac{\pi_{jj}}{\pi_{ij}}\right)^{\frac{1}{\theta}}.$$

To operationalize these equations we need information about the price of the tradable aggregate p_i^T in each country i. In our model the price of the tradable aggregate equals the price index of imports, $p_i^T = p_i^I$, provided that the later is a simple quantity weighted average of the price of imports. We use this fact and Penn World Tables 8.0 data on the price index for imports (pl_m) to calibrate this version of the model.

Figure OA.2 shows measures of the fraction of world TFP growth (left panel) and the cross-sectional variance of TFP growth (right panel) explained by trade costs. The lessons from Section 4.4 are robust to this alternative modeling of non-tradable goods. The contribution of trade in accounting for TFP changes is greatest for intermediate values of the diffusion parameter, β. For these values, gains from trade are up to three times as large when the model allows for dynamic gains from trade.

E.3 Learning from Producers

In this section we present results when insights are drawn from producers. Figure OA.3 reports the fraction of TFP growth explained by trade costs in the specification where insights are drawn uniformly from domestic producers (see Section 2.2). In this version of the model, diffusion simply amplifies the static gains from trade. Changes in the composition of trading partners, which was an important correlate of TFP gains in the reduced form evidence in Figure OA.12, and an important driver of TFP gains in the benchmark model with learning from sellers is absent. Therefore, it is not surprising that the ability of trade costs to account

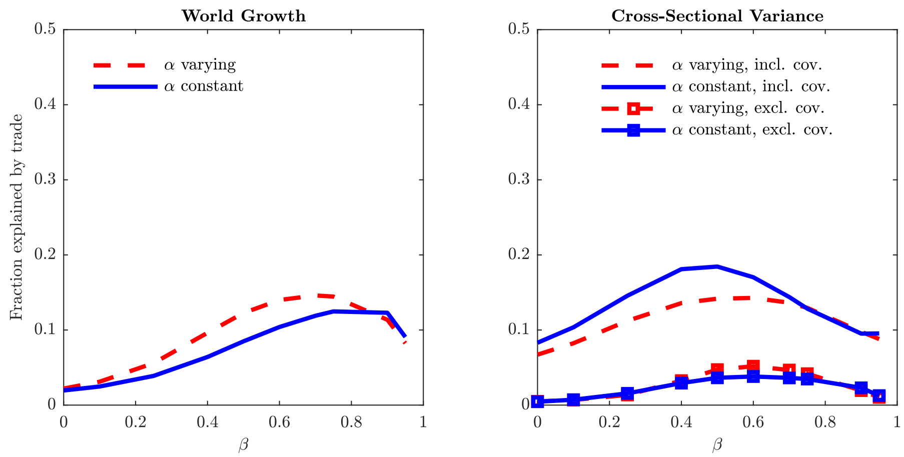

Figure OA.2: The contribution of changes in trade costs to changes in TFP: Alternative Modeling of Non-tradable Goods.

Note: This figure reports the fraction of TFP growth accounted for by trade costs, for various values of β, according to two decompositions described in Section 4.4. 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. See the note to Figure 6 for details.

for TFP gains is significantly muted.

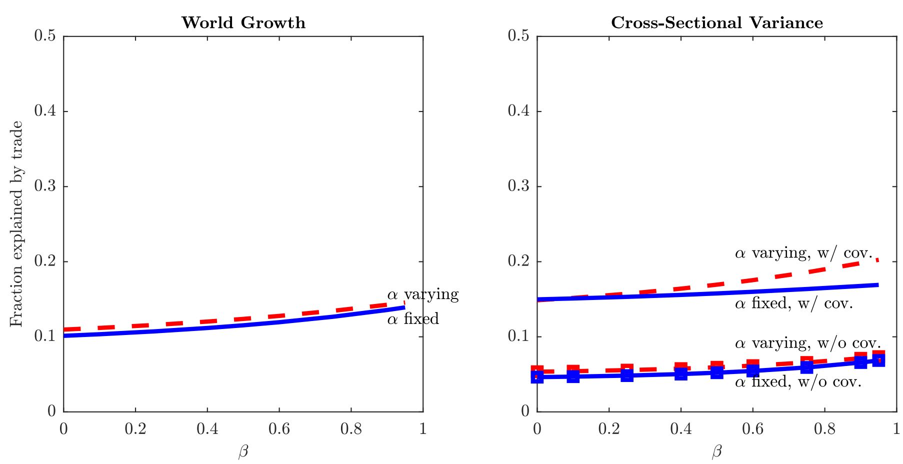

Figure OA.3: The contribution of changes in trade costs to changes in TFP: Learning from Producers, Uniformly.

Note: This figure reports the fraction of TFP growth accounted for by trade costs, for various values of β, according to two decompositions described in Section 4.4. 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. See the note to Figure 6 for details.

E.4 Alternative Weights

Here we explore three alternative processes by which individuals can learn. In the baseline, individuals are equally likely to learn from all active sellers or all active domestic producers, independently of how much of the seller's variety they consume or the size of that producer. In the first alternative, insights are drawn from sellers in proportion to the expenditure on each seller's good. In the second, insights are drawn in proportion to consumption of each seller's goods. In both cases, the speed of learning is the same as our baseline up to a constant. In the third, insights are drawn from domestic producers in proportion to those producers' employment.

E.4.1 Learning from Sellers in Proportion to Expenditure

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 3, this is

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

E.4.2 Learning from Sellers in Proportion to Consumption

i's consumption of good 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 3, this is

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

E.4.3 Learning from Producers in Proportion to Employment

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

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

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 3, this is

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

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

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

Figure OA.4 reports the fraction of TFP growth explained by trade costs in the specification where insights are drawn from domestic producers in proportion to employment.

This version of the model has some qualitative features similar to the benchmark model with learning from sellers, and some important differences when considering values of β close to 1. In the limit as $\beta \to 1$, a balanced growth path requires an joint restriction on arrival rates and trade shares, $\alpha_i = \sum_j \pi_{ji}\alpha_j$. Since the counterfactuals (which do not impose BGP) violate this restriction, they produce extreme (but uninteresting) results.

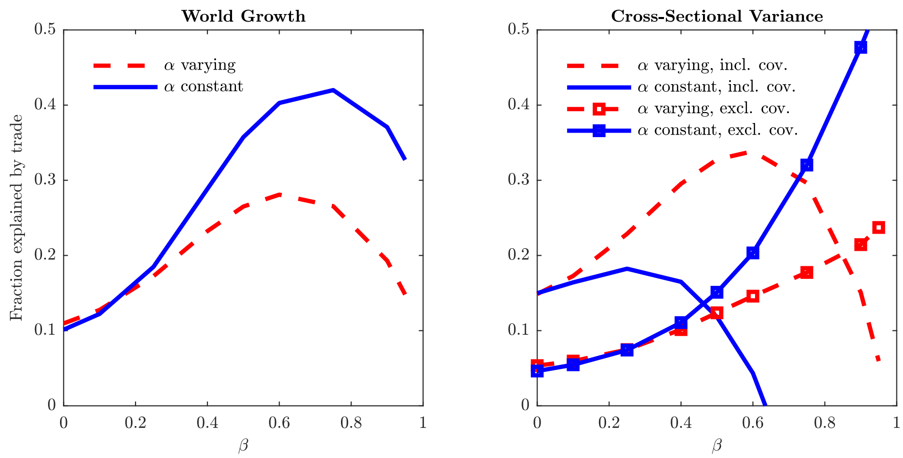

Figure OA.4: The contribution of changes in trade costs to changes in TFP: Learning from Producers, in proportion to labor used.

Note: This figure reports the fraction of TFP growth accounted for by trade costs, for various values of β, according to two decompositions described in Section 4.4. 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. See the note to Figure 6 for details.

E.5 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 G represents the distribution of productivity among those from whom one may draw insights. We assume now that producers 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 $\left[\int_0^\infty \hat{\alpha}(x)^{\frac{\phi}{\phi-1}} dG(x)\right]^{\frac{\phi-1}{\phi}} \leq \alpha$ for some $\phi > 1$.[6]

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

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

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

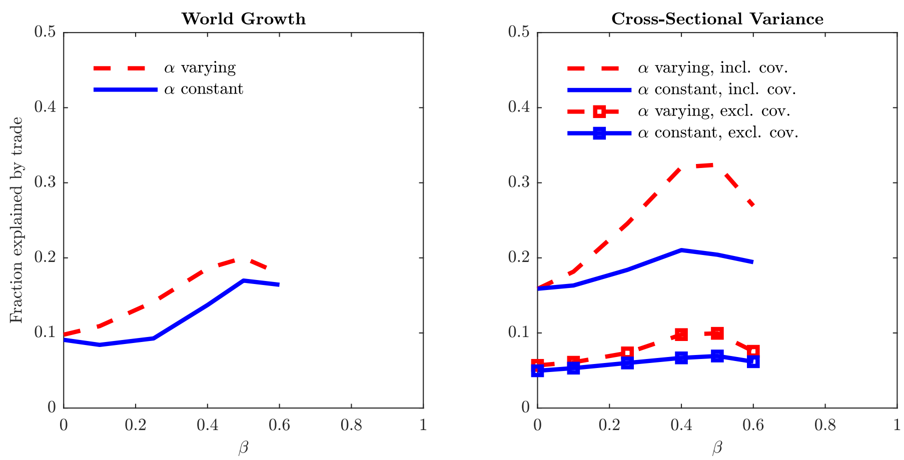

Figure OA.5: The contribution of changes in trade costs to changes in TFP, Targeted Learning.

Note: This figure reports the fraction of TFP growth accounted for by trade costs, for various values of β, according to two decompositions described in Section 4.4. 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. See the note to Figure 6 for details.

Figure OA.5 reports the fraction of TFP growth accounted for by trade costs for the extension of the benchmark model with targeted learning. In particular, we present results for the case with $\phi = 1.5$. Notice that now we must restrict the values of $\beta < 1/\phi$.[7] As the

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

figure clearly illustrates, for each value of β the results are qualitatively very similar to an environment without targeted learning and a higher value of β.

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

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

Conditioning on a value of α, learning is faster when learning is more targeted ($\Gamma(1-\beta\phi)$ is increasing in ϕ). However, conditioning on a calibration target, the faster learning plays no role because α would be recalibrated to absorb the change.

E.6 Variety

An implication of the learning from producers specification is that the rate of increase of a country's stock of knowledge grows without bound as the share of its expenditure spent on domestic goods shrinks to zero. In that case, only the most productive managers would be able to sell goods domestically, so the insights drawn from these firms would 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.

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[8]

$$\dot{\lambda}_{it} = \Gamma(1-\beta)\alpha_{it}\pi_{ii}\left(\frac{\lambda_i}{\pi_{ii}}\right)^{\beta}. \tag{10}$$

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

[8] 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 (10) the arrival rate is $\alpha_{it}\pi_{ii}$.

balanced growth path, the detrended stock of knowledge is

$$\hat{\lambda}_i \propto \hat{\alpha}_i^{1/(1-\beta)} \pi_{ii}.$$

E.7 The Role of Trade Exposure

A key driver of the dynamic gains in the model with learning from sellers is increasing exposure to more productive trading partners. To further highlight this effect, we construct counterfactual evolutions of stocks of knowledge and, therefore, counterfactual TFP series for each country to illustrate these effects. We distinguish between two ways a country's trade exposure could change. We first consider a counterfactual path in which each country's own trade share changes, but the fraction of imports coming from each trading partner is held fixed. Second, we consider the additional gains coming from shifts in the distribution of imports among trading partners. As we show, both of these component is important for the case of miracle economies.

To be specific, given initial vector of countries' stocks of knowledge, λ_0, a path of arrival rates for each country, α_t, we can construct a sequence of countries' stocks of knowledge corresponding to any sequence of trade matrices, Π_t. We call this path of stocks of knowledge that emerges $\lambda_t(\Pi, \alpha_t)$. In turn, we call the sequence of countries' TFP that emerges $TFP(\lambda_t, \Pi_t)$. We can now decompose the total gains from trade into a static component, a dynamic component with fixed trade exposures, and a dynamic component capturing changes in the composition of imports:

$$\log \frac{TFP(\lambda_t, \Pi_t)}{TFP(\lambda_0, \Pi_0)} = \log \frac{TFP(\lambda_0, \Pi_t)}{TFP(\lambda_0, \Pi_0)} + \log \frac{TFP(\lambda_t(\Pi_t^{FE}, \alpha_t), \Pi_t)}{TFP(\lambda_0, \Pi_t)} + \log \frac{TFP(\lambda_t(\Pi_t, \alpha_t), \Pi_t)}{TFP(\lambda_t(\Pi_t^{FE}, \alpha_t), \Pi_t)}$$

where the elements of the trade matrix for the fixed-exposure counterfactual are computed as $\pi_{ijt}^{FE} \equiv (1 - \pi_{iit}) \frac{\pi_{ij0}}{\sum_{k \neq i} \pi_{ik0}}$, all $j \neq i$.

Figure OA.6 shows actual changes in TFP against each of these components of the total gains from trade, under the assumption that learning is from sellers, $\beta = 0.6$, and the arrival rate of ideas is kept at its 1962 value, $\alpha_{it} = \alpha_{i0}$. The left panel holds fixed all countries' stocks of knowledge and, therefore, it illustrates the static gains from trade. The second panel considers the additional changes in TFP that occur when each country's stock of knowledge changes according to the counterfactual pattern of trade in which each country's own trade share evolves as in the data but the composition of imports is held fixed at its initial values. Finally, the third panel considers the additional changes in TFP due to changes in stocks of knowledge responding to observed trade patterns.

In the figure we label the observations of the miracle economies. Each of these countries

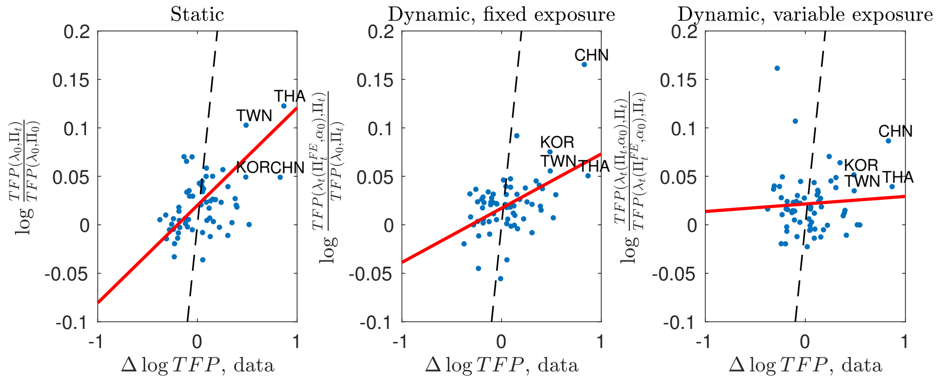

Figure OA.6: Decomposition of Changes in TFP, 1962-2000

Note: This plot shows actual changes in TFP against the various components of predicted changes in TFP, under the assumption that learning is from sellers, $\beta = 0.6$, and the arrival rate of ideas is kept at its 1962 value, $\alpha_{it} = \alpha_{i0}$. The left panel holds fixed all counties stocks of knowledge and, therefore, it illustrates the static gains from trade. The second panel considers the additional changes in TFP that occur when the stock of knowledge of each countries changes in response to the observe change in the own trade share, but the relative exposure across foreign countries is held fixed to their initial values, i.e., $\pi_{ijt}^{FE} = (1 - \pi_{iit})\pi_{ij0}/\sum_{k \neq i} \pi_{ik0}$, all $j \neq i$. Finally, the third panel considers the additional changes in TFP that occur when the stock of knowledge of each countries changes in response to the observe changes in the relative exposure across foreign countries.

opened up their economies and experienced large gains from trade. Interestingly, both components that comprise the dynamic gains from trade were important for these countries. These countries increased their expenditures on imports, a fact that explained the large increases in TFP featured in the first two panels, but also redirected their trade towards more productive countries, explaining the increase in TFP in the third panel.

E.8 Using PPP as a Calibration Target

In this appendix we discuss an alternative calibration strategy that considers the explicit mapping between the PPP data from the Penn World Tables (PWT) and the model. In particular, we use the definition of PPP and the structure of the model to obtain a system of equations that is used to solve for the theoretical price index p_i, given data from the PWT. We then use the value of the theoretical price index to perform the calibration strategy described in the main text.

We first describe the mapping between the theoretical price index and the PPP in the PWT. Let $p_i(s)$ and $c_i(s)$ be country i's price and consumption of good s. Following the Geary-Khamis approach to construct PPP (as described by Feenstra et al., 2015), PPP is defined as part of a system of equations. Note that in an open economy, there are two separate PPP indices for each country: PPP_i^q, which is used to deflate country i's final expenditure, and PPP_i^o, which is used to deflate a country's output. The international price of a good is defined as a quantity-weighted average of each country's price for the good relative to that country's PPP_i^o:

$$p^{int}(s) = \frac{\sum_i \frac{p_i(s)}{PPP_i^o} c_i(s)}{\sum_i c_i(s)} \tag{11}$$

where PPP_i^q is defined as the ratio of i's expenditure to i's implied expenditure if each good where priced at the international price, i.e.,

$$PPP_i^q = \frac{\int p_i(s) c_i(s) ds}{\int p^{int}(s) c_i(s) ds}. \tag{12}$$

The definition for PPP_i^o is similar, although we omit it here because it is not needed for our calibration.

We next construct PPP_i^q in our model and show the relationship to the theoretical price index in each country. Each good is characterized by the vector of productivities at which the various countries can produce the good, $\mathbf{q} = (q_1, ..., q_N)$. Alternatively, given each country's stock of knowledge, $\lambda = (\lambda_1, ..., \lambda_N)$, we can characterize each good by

$$\mathbf{u} = (u_1, ..., u_N) = \left(\frac{q_1}{\lambda_1^{1/\theta}}, ..., \frac{q_N}{\lambda_N^{1/\theta}} \right).$$

Note that $(u_1, ..., u_N)$ are independent and identically distributed, each drawn from distribution with CDF $e^{-u^{-\theta}}$. Let U denote the distribution of $\mathbf{u}$. The following two lemmas will help.

Claim 6 *The price index of good* $\mathbf{u}$ *in* i *can be expressed as* $p_i(\mathbf{u}) = p_i r_i(\mathbf{u})$ *where* $r_i(\mathbf{u}) \equiv \min\left\{ \frac{A_{i1}}{u_1}, ..., \frac{A_{iN}}{u_N} \right\}$ *and* $A_{ij} \equiv \frac{p_j^\eta w_j^{1-\eta} \kappa_{ij}}{p_i \lambda_j^{1/\theta}}$

Proof. If country j can produce a good at productivity q_j, its cost of delivering a unit of that good to i is $\frac{p_j^\eta w_j^{1-\eta} \kappa_{ij}}{q_j}$, or, in terms of $u_j \equiv \frac{q_j}{\lambda_j^{1/\theta}}$, $\frac{p_j^\eta w_j^{1-\eta} \kappa_{ij}}{u_j \lambda_j^{1/\theta}} = \frac{p_i}{u_j} A_{ij}$. Therefore

$$p_i(\mathbf{u}) = \min\left\{ \frac{p_i}{u_1} A_{i1}, ..., \frac{p_i}{u_N} A_{iN} \right\} = p_i r_i(\mathbf{u})$$

■

Claim 7 A_{ij} *can be expressed in terms of data and parameters.*

Proof. As described in Section 4.2 of the main text, we derive expressions for the iceberg trade costs, $\kappa_{ij} = \frac{p_i}{p_j} g_1 (\pi_{ii}, \pi_{ij})$ and a country's stock of knowledge, $\lambda_i = g_2 (\pi_{ii}) \left(\frac{w_i}{p_i}\right)^{(1-\eta)\theta}$ where g_1 and g_2 depend only on trade shares and parameters. We can thus express A_{ij} as

$$A_{ij} = \frac{p_j^{\eta} w_j^{1-\eta} \kappa_{ij}}{p_i \lambda_j^{1/\theta}} = \left[\frac{1}{\lambda_j}\left(\frac{w_j}{p_j}\right)^{(1-\eta)\theta}\right]^{\frac{1}{\theta}} \left(\frac{p_j}{p_i}\kappa_{ij}\right) = g_2 (\pi_{ii})^{-1/\theta} g_1 (\pi_{ii}, \pi_{ij}).$$

■

We are now in position to construct PPP in the model. Equations (11) and (12) can be expressed as

$$p^{int}(\mathbf{u}) = \frac{\sum_i \frac{p_i(\mathbf{u})}{PPP_i^o} c_i(\mathbf{u})}{\sum_i c_i(\mathbf{u})}$$

and

$$PPP_i^q = \frac{\int p_i(\mathbf{u}) c_i(\mathbf{u})\, dU(\mathbf{u})}{\int p^{int}(\mathbf{u})\, c_i(\mathbf{u})\, dU(\mathbf{u})}.$$

Let X_i be the nominal expenditure in i. Note that $c_i(\mathbf{u}) = C_i \left(\frac{p_i(\mathbf{u})}{p_i}\right)^{-\varepsilon} = X_i p_i^{\varepsilon-1} p_i(\mathbf{u})^{-\varepsilon}$, so these can be written as

$$p^{int}(\mathbf{u}) = \frac{\sum_i \frac{p_i(\mathbf{u})}{PPP_i^o} p_i^{\varepsilon-1} p_i(\mathbf{u})^{-\varepsilon} X_i}{\sum_i p_i^{\varepsilon-1} p_i(\mathbf{u})^{-\varepsilon} X_i} = \frac{\sum_i r_i(\mathbf{u})^{1-\varepsilon} \frac{X_i}{PPP_i^o}}{\sum_i r_i(\mathbf{u})^{-\varepsilon} \frac{X_i}{p_i}}$$

and

$$PPP_i^q = \frac{X_i}{\int p^{int}(\mathbf{u})\, p_i^{\varepsilon-1} p_i(\mathbf{u})^{-\varepsilon} X_i dU(\mathbf{u})} = \frac{p_i}{\int p^{int}(\mathbf{u})\, r_i(\mathbf{u})^{-\varepsilon} dU(\mathbf{u})}.$$

Combining these by eliminating $p^{int}(\mathbf{u})$ and rearranging gives

$$p_i = PPP_i^q \int r_i(\mathbf{u})^{-\varepsilon} \frac{\sum_{\tilde{i}} r_{\tilde{i}}(\mathbf{u})^{1-\varepsilon} \frac{X_{\tilde{i}}}{PPP_{\tilde{i}}^o}}{\sum_{\tilde{i}} r_{\tilde{i}}(\mathbf{u})^{-\varepsilon} \frac{X_{\tilde{i}}}{p_{\tilde{i}}}} dU(\mathbf{u}). \tag{13}$$

Given parameters and data on each country's PPP^q, PPP^o, nominal expenditure, and trade shares, (13) for countries $i = 1, ..., N$ gives a system of N equations for the N unknown price indices. In practice, we compute the integrals by simulating a world with a large number of goods and draw a vector $\mathbf{u}$ for each good.

Figure OA.7 illustrates the results from two calibration strategies: (i) the benchmark calibration where we assume that the theoretical price index equals the PPP of the final expenditure in the PWT; (ii) an the alternative, more theoretically rigorous strategy, where we calculate the theoretical price index using the mapping implied by the theory, as described in this appendix. The left panel plots the theoretical price index against the PPP of final expenditure in the data. The middle panel plots the stocks of knowledge implied by our benchmark strategy against those implied by our alternative, more rigorous strategy. The right panel plots the trade costs implied by the two calibration strategies. Figure OA.7 shows that the difference between the two calibration strategies is quantitatively negligible.

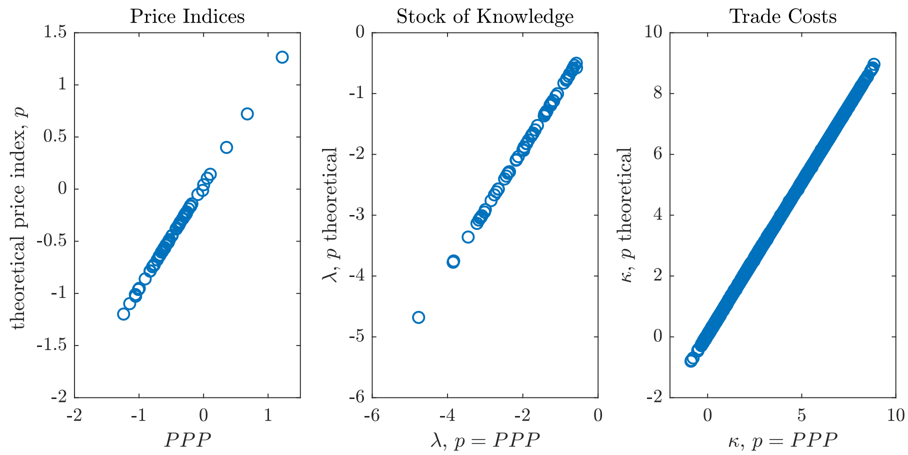

Figure OA.7: Comparing Alternative Calibration Strategies

Note: These subplot compare the results from two calibration strategies: (i) the benchmark calibration where we assume that the theoretical price index equals the PPP of the final expenditure in the PWT; (ii) the alternative, more theoretically rigorous strategy, where we calculate the theoretical price index using the mapping implied by the theory, as described in this appendix. The left panel plots the theoretical price index against the PPP of final expenditure in the data. The middle panel plots the stocks of knowledge implied by our benchmark strategy against those implied by our alternative, more rigorous strategy. The right panel plots the trade costs implied by the two calibration strategies.

F Explaining the Initial Distribution of TFP

We next assess the role of trade barriers in accounting for initial cross-country TFP differences. To this end, we use as a baseline 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.

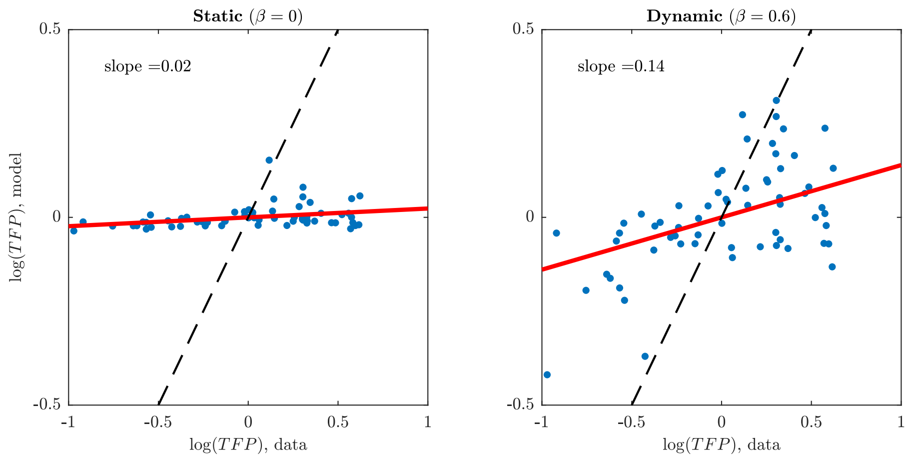

Figure OA.8: Openness and the Distribution of TFP in 1962

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 $\beta = 0.6$ 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 OA.8 compares the implied distribution of TFP on the balanced growth path of the model to the actual distribution of TFP in 1962, the first year in our sample. Each dot represents a country. The first panel shows the case of $\beta = 0$, 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, these static gains generate only a small amount of cross-country differences in productivity.

The second panel assumes that $\beta = 0.6$ 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 positive slope of the red regression line 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 OA.9. 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 4.4, we can do this in two ways. In (14) the contribution of trade is evaluated at common arrival rates $\bar{\alpha}$, while in (15) it is evaluated at the country-specific arrival rates.

$$\ln \frac{TFP_i(\alpha_i, \kappa_{ij})}{TFP_i(\bar{\alpha}, \bar{\kappa})} = \underbrace{\ln \frac{TFP_i(\bar{\alpha}, \kappa_{ij})}{TFP_i(\bar{\alpha}, \bar{\kappa})}}_{\text{cont. from trade}} + \underbrace{\ln \frac{TFP_i(\alpha_i, \kappa_{ij})}{TFP_i(\bar{\alpha}, \kappa_{ij})}}_{\text{cont. from arrival rates}} \tag{14}$$

$$\ln \frac{TFP_i(\alpha_i, \kappa_{ij})}{TFP_i(\bar{\alpha}, \bar{\kappa})} = \underbrace{\ln \frac{TFP_i(\alpha_i, \bar{\kappa})}{TFP_i(\bar{\alpha}, \bar{\kappa})}}_{\text{cont. from arrival rates}} + \underbrace{\ln \frac{TFP_i(\alpha_i, \kappa_{ij})}{TFP_i(\alpha_i, \bar{\kappa})}}_{\text{cont. from trade}} \tag{15}$$

Each panel of Figure OA.9 has four lines. The two solid lines correspond to the decomposition in (14) 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 (15) 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-sectional variance in 1962 and 2000, respectively.

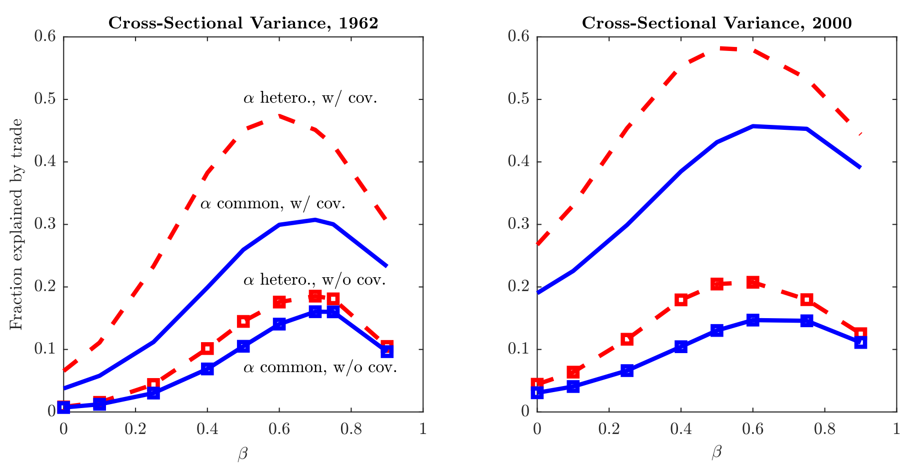

Figure OA.9: Mean squared error of predicted TFP

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 (14) in which the contribution of trade is evaluated at a common arrival rate; the dashed lines to correspond to (15) 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.

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 OA.8. As we consider specifications for which the strength of diffusion is larger, the model accounts for more of the variation in TFP. Again, while the contribution of trade to initial differences in TFP initially rises with the strength of diffusion, it is greatest for cases with intermediate values of the diffusion parameter, β. For a large range of values of β, however, variation in trade barriers accounts on average for up to 45% of the cross-sectional dispersion in TFP in 1962.[9]

G Reduced-Form Evidence

In this appendix we present suggestive reduced-form evidence of the mechanisms emphasized by the theory: whether openness and trade with more advanced countries is associated with higher productivity. In addition, this evidence provides interesting non-targeted moments related to the mechanisms in the model and, therefore, we compare the same reduced-form regressions in the actual data and the model generated data.

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 its trading partners. The first panel of Figure OA.10 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 OA.10 provides simple reduced-form evidence suggesting a correlation between the level of TFP and openness and in the exposure to productive trading partners. It is a natural step to evaluate the ability of the model to account for this reduced-form evidence.

Figure OA.11 presents the reduced-form relationship between openness and the level of TFP, using model generated data when $\beta = 0.6$, the arrival rate is the same in each country, $\hat{\alpha}_i = \hat{\alpha}$, and the model is on a balanced growth path consistent with the calibrated trade costs in 1962. Given these assumptions, trade is the only force driving TFP differences across countries. The model reproduces remarkably well the quantitative patterns reported in Figure OA.10.

We now turn to the time series implications of the theory. Over time, among the many factors that would alter a country's productivity, the model emphasizes changes in open-

[9] These results suggest that there can very large gains in productivity if the poorest countries in our sample, e.g., Niger, were to face the trade cost of open developed economies, e.g., Belgium. Appendix H.4 illustrates these possibilities.

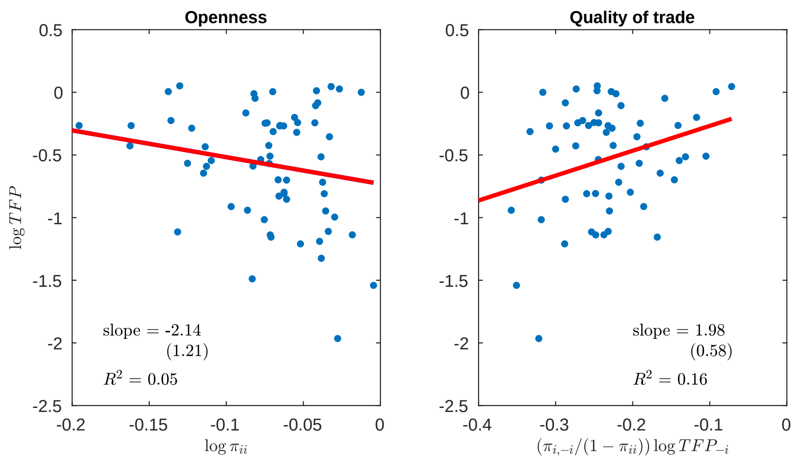

Figure OA.10: Cross-sectional TFP differences in 1962

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.

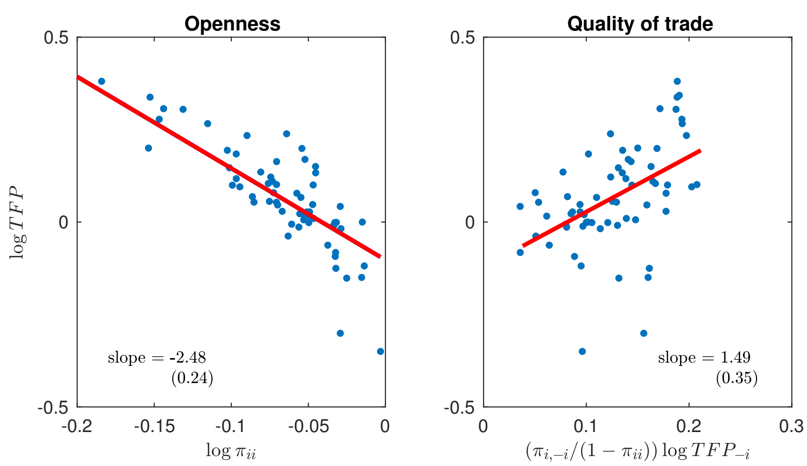

Figure OA.11: reduced-form Relationship between Openness and TFP in Model Generated Data, 1962

Note: These panels reproduce the reduced-form evidence in Figure OA.10, using model generated data when $\beta = 0.6$, and the arrival rate is the same in each country, $\hat{\alpha}_i = \hat{\alpha}$.

ness, changing exposure to trading partners, and changes in trading partners' productivity. Figure OA.12 shows some suggestive reduced-form evidence of these mechanisms.

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.

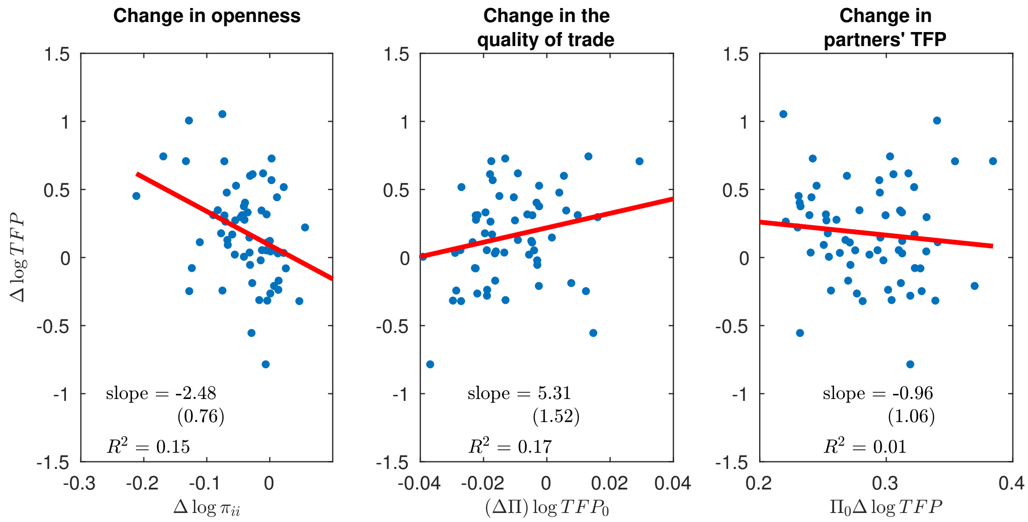

Figure OA.12: Openness and Changes in TFP, 1962-2000

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 second panel shows the association between the change in each country's composition of expenditures and its TFP growth. For each country, we compute the changes in exposure to trading partners with high initial TFP. Specifically, for country i we compute $\sum_j(\pi_{ij}^{2000} - \pi_{ij}^{1962}) \ln TFP_j^{1962}$. 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.

Figure OA.12 provides simple reduced-form evidence suggesting a correlation between TFP growth and changes in openness and in the exposure to productive trading partners. Again, it is a natural step to evaluate the ability of the model to account for this reduced-form evidence.

Figure OA.13 presents the reduced-form relationship between openness and changes in TFP, using model generated data when $\beta = 0.6$, arrival rates of ideas are kept at 1962 values, $\alpha_{it} = \alpha_{i0}$. The model qualitatively reproduces the patterns reported in Figure OA.12. As was the case in the data, changes in own trade share (left panel) and changes in exposure

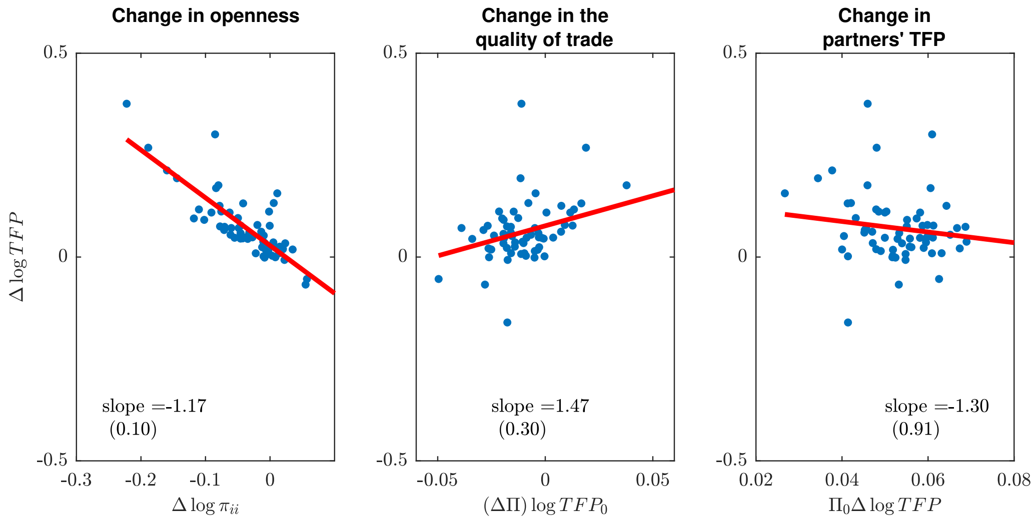

Figure OA.13: reduced-form Relationship between Openness and TFP Growth in Model Generated Data, 1962-2000

Note: These panels reproduce the reduced-form evidence in Figure OA.12, using model generated data when $\beta = 0.6$, and the arrival rate of ideas is kept at its 1962 value, $\alpha_{it} = \alpha_{i0}$.

to high-TFP trading partners (middle panel) correlate significantly with changes in TFP, while changes in trading partners' TFP is not significantly associated with TFP growth. Interestingly, the magnitude of the coefficients in the model generated data are smaller than in the actual data. This is true even though in our model the gains from trade are substantially higher than those of a static trade model. Through the lens of our model, this suggests that part of relationship between openness and growth is accounted for by the correlation between changes trade cost and changes in the arrival rate of ideas; countries whose TFP rose due to changes in trade barriers also tended to increase research intensity. This was already suggested by the importance of the covariance term in Figure 6.

This reduced-form evidence is in principle consistent with a world in which variables other than trade costs affect both TFP and trade flows. The structural analysis in the main text provides a framework to quantify the contribution of measured change in trade costs to account for changes in TFP and trade. Therefore, it provides a structural assessment of the role of these mechanisms.

H Additional Illustrative Examples

H.1 Gains from trade in a Symmetric Economy

Figure OA.14 replicates Figure 1, showing each country's stock of knowledge and real income as a function of (the inverse of) trade costs for a symmetric economy.

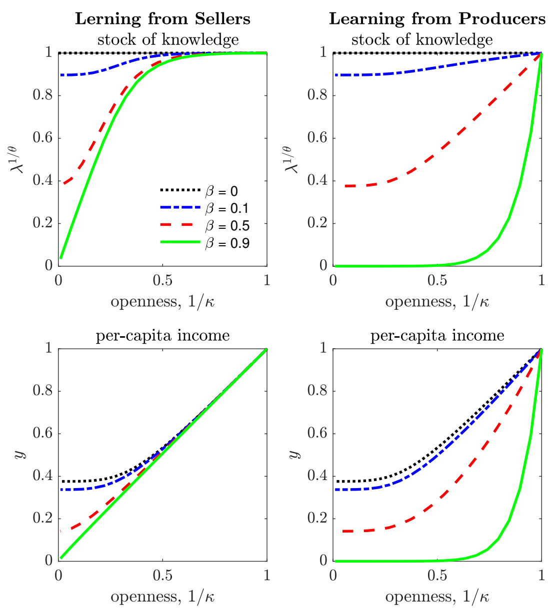

Figure OA.14: Gains from trade in a Symmetric Economy

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

H.2 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? More specifically, for which specifications of learning and which parameter values do the gains from trade accrue quickly?

Consider a world economy that starts with $n-1$ open economies and a single deviant economy that are on a balanced growth path. Trade is costless among the open economies but shipping a good to/from the deviant economy is, initially, infinitely costly. Figure OA.15 shows the evolution of the real income in the initially deviant economy following a trade liberalization in which trade with the deviant country becomes costless. 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 real income in the deviant country.[10] The paths of the (detrended) stocks of knowledge solve the differential equations in (5) and (8), depending on whether insights are drawn from sellers or producers.[11] 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.

The right panel of Figure OA.15 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 eventually 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, around 50% of the dynamic gains from trade have been realized.

We can obtain 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

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

where $\Omega_{ii} \equiv \frac{\pi_{ii}^{1-\beta}\lambda_i^{\beta}}{\sum_j \pi_{ij}^{1-\beta}\lambda_j^{\beta}}$ 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_{ii} is significantly larger than than π_{ii}, a case in which i's stock of

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

[11] We set $\beta = 0.5$. The rest of the parameters follow the calibration in footnote 27.

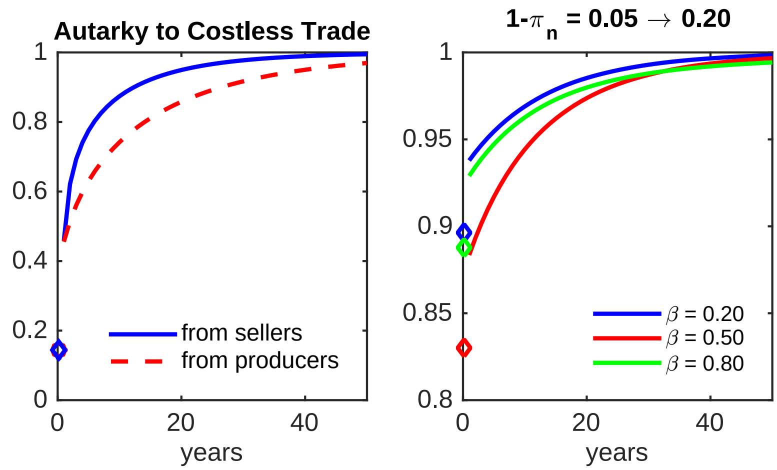

Figure OA.15: Dynamics Following a Trade Liberalization.

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. See footnote 27 for additional details of the calibration used in this figure.

knowledge is much larger than that of its trading partners.

H.3 Core-Periphery Structure

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 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 OA.16 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 earlier discussion, 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.

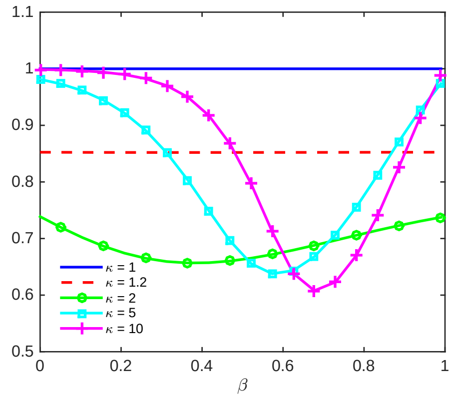

Figure OA.16: A Core-Periphery Economy

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.

H.4 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.[12] This exercise is presented in Figure OA.17.

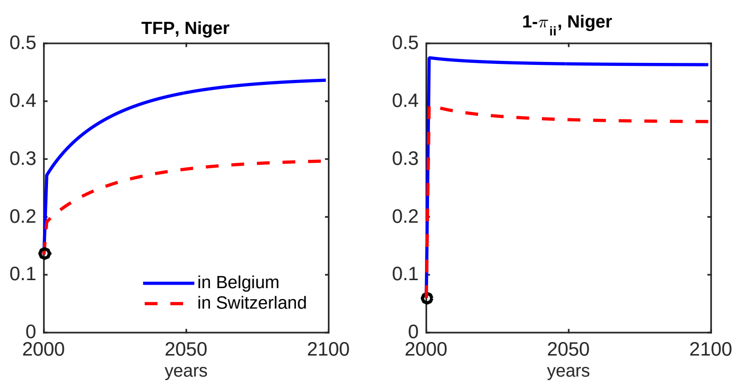

Figure OA.17: Transitional Dynamics after Assigning Belgium and Switzerland's Trade Costs to Niger.

The left panel plots the evolution of TFP, normalized by the TFP of the US. On impact the TFP jumps due to the static gains from trade, reflecting increased specialization and

[12] 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.

comparative advantage. Given that Niger is a very isolated economy, compared to the more integrated developed countries, the static gains increase Niger's income by 5 to 10 percentage points. Over time, as firms in Niger interact and get 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 are 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.

The right panel shows the evolution of Niger's import share in the counterfactual experiments. The sharp increases in import share illustrates the large differences in trade costs characterizing poor economies.

I Additional Information about the Quantitative Exercises

In this Appendix we provide additional information about the quantitative exercises performed in the main body of the paper.

I.1 Trade Costs and Arrival Rates of Ideas, Miracle Economies

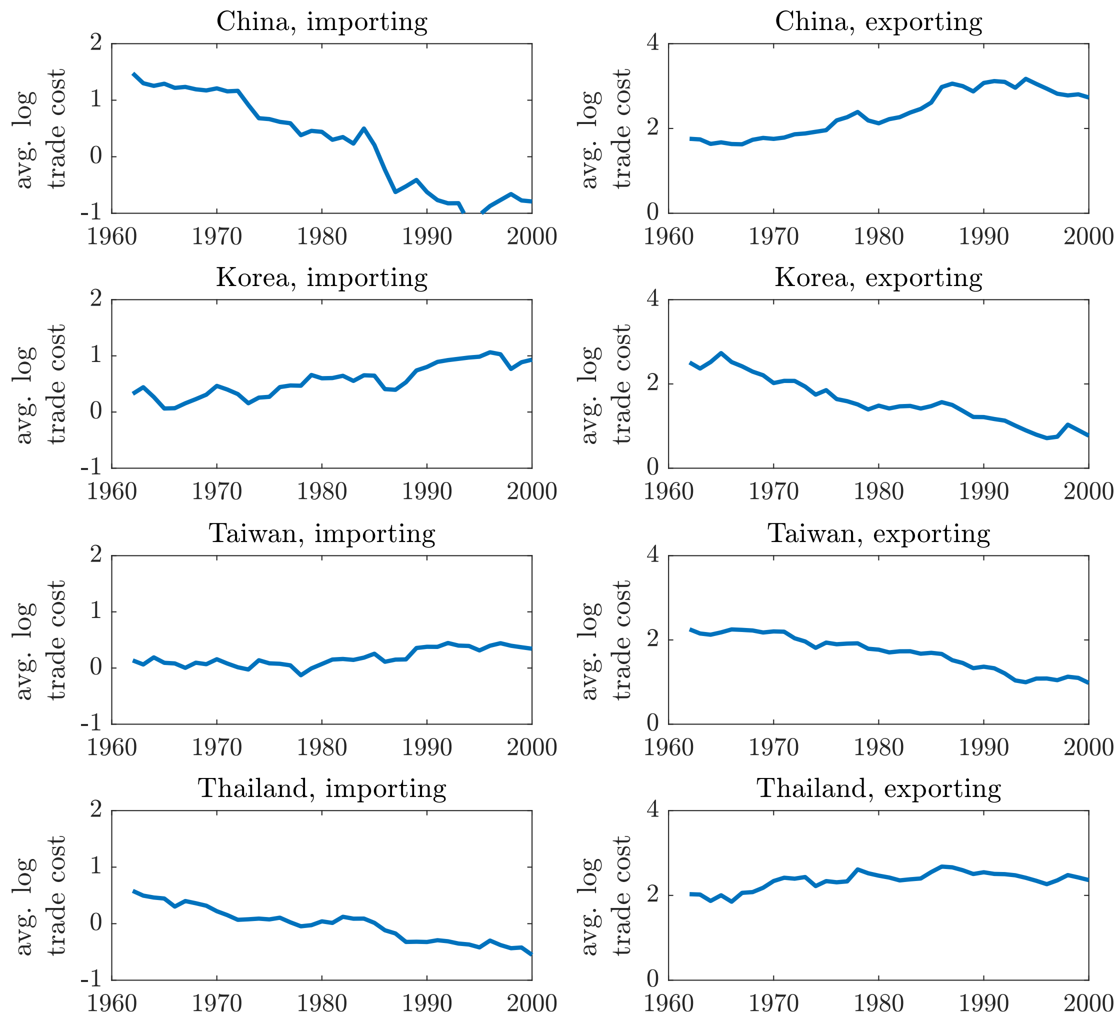

Figure OA.18: Evolution of Importing and Exporting Trade Costs: Miracle Economies.

Note: This figure plots the evolution of import-weighted (export-weighted) average of importing (exporting) log trade costs for the miracle economies.

Figure OA.18 shows the evolution of trade-weighted average log trade cost for the four miracle economies that we feature in the quantitative section. The left panels show the import-weighted average of the (log) cost of transporting a good to each individual miracle economy from every other country in our sample. The right panels show the export-weighted average of (log) cost of transporting a good from each individual miracle economy to every other country in our sample.

There are three important things to notice in these figure. First, importing costs tend to be lower than exporting costs, a well-known feature of trade costs that are consistent with trade patterns and relative prices in the data (e.g., Waugh, 2010). Second, importing and exporting costs exhibit diverging trends, reflecting trends in the price indexes of these countries relative to the rest of the world. For China and Thailand, average importing costs exhibit a substantial decline over the sample period, while exporting costs increase. For Korea and Taiwan, exporting costs exhibit a marked decline, while importing costs increase. The divergent trends are explained by the movements in the price indices of these countries relative to their trading partners over the sample period, declining for China and Thailand and increasing for Korea and Taiwan. Finally, the decline of importing/exporting costs is more important than the increase in exporting/importing costs.

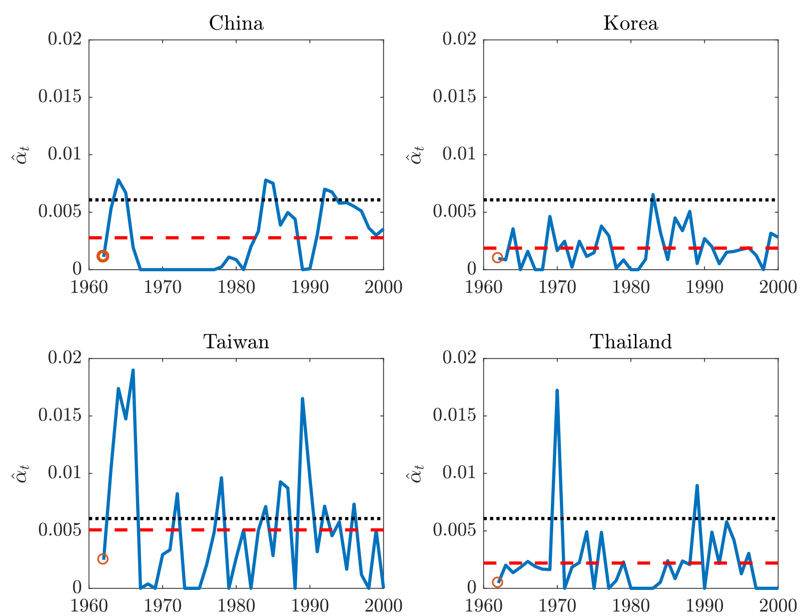

Figure OA.19: Evolution of the (detrended) Arrival Rates of Ideas for Miracle Economies.

Note: This figure plots the evolution the detrended arrival rate of ideas (solid line) together with its average for the same period (dashed line) and the corresponding average for the US (dotted line).

Figure OA.19 shows the evolution of each country's detrended arrival rates of ideas (solid line) together with its average for the sample period (dashed line) and the corresponding average of the detrended arrival rates of ideas of the US (dotted line).

Again, there are three important observations to make. First, for all the miracle economies the average of the detrended arrival rates of ideas is higher than the initial (circled) value, which is the value needed to rationalize the initial value of the stock of knowledge as a balanced growth path. Second, by construction, the evolution of the detrended arrival rates of ideas is volatile as it is chosen to account for the fluctuations in TFP that are not accounted for by the evolution of trade costs, itself a relatively smooth variable. Finally, as explained in footnote 43 of the main text, the detrended arrival rates of ideas is restricted to be non-negative.

I.2 Results by Country

Table 1: TFP Growth by Country, 1962-2000. Data vs. Models with $\beta = 0.6$.

	Data	Static Effect	Static+Dynamic Effect	
	$\ln \frac{TFP_i(\lambda_t,\kappa_t)}{TFP_i(\lambda_0,\kappa_0)}$	$\ln \frac{TFP_i(\lambda_0,\kappa_t)}{TFP_i(\lambda_0,\kappa_0)}$	$\ln \frac{TFP_i(\lambda(\alpha_0,\kappa_t),\kappa_t)}{TFP_i(\lambda(\alpha_0,\kappa_0),\kappa_0)}$	$\ln \frac{TFP_i(\lambda(\alpha_t,\kappa_t),\kappa_t)}{TFP_i(\lambda(\alpha_t,\kappa_0),\kappa_0)}$
Thailand	0.854	0.123	0.213	0.262
China	0.818	0.049	0.301	0.492
Ireland	0.531	0.237	0.268	0.336
Japan	0.503	0.000	0.007	0.043
Taiwan	0.476	0.103	0.193	0.255
Korea, Republic of	0.473	0.049	0.176	0.284
Belgium+Lux	0.395	0.350	0.376	0.329
Israel	0.384	0.018	0.048	0.075
Sri Lanka	0.377	0.020	0.051	0.058
Greece	0.366	0.027	0.077	0.027
Egypt	0.334	0.002	0.112	0.134
Finland	0.293	0.039	0.109	0.129
Pakistan	0.291	-0.011	0.019	0.046
Tunisia	0.243	0.057	0.126	0.098
Indonesia	0.225	0.023	0.132	0.197
Norway	0.208	-0.003	0.003	0.052
Italy	0.169	0.026	0.047	0.042
United Kingdom	0.142	0.026	0.045	0.020
Brazil	0.141	0.003	0.018	0.022
Mozambique	0.140	0.050	0.169	0.118
Turkey	0.118	0.025	0.059	0.030
India	0.111	0.007	0.062	0.088
Cote d'Ivoire	0.109	-0.012	0.021	0.020
Australia	0.104	0.017	0.060	0.063
Denmark	0.083	0.005	-0.001	-0.003
Portugal	0.078	0.059	0.109	0.068
Uruguay	0.077	0.023	0.057	0.045
Austria	0.075	0.050	0.112	0.102
Netherlands	0.049	0.037	0.046	0.072
Mali	0.041	-0.036	-0.054	-0.020
Sweden	0.039	0.034	0.047	0.054
Cameroon	0.034	-0.008	0.009	0.024

	Data	Static Effect	Static+Dynamic Effect	
	$\ln \frac{TFP_i(\lambda_t,\kappa_t)}{TFP_i(\lambda_0,\kappa_0)}$	$\ln \frac{TFP_i(\lambda_0,\kappa_t)}{TFP_i(\lambda_0,\kappa_0)}$	$\ln \frac{TFP_i(\lambda(\alpha_0,\kappa_t),\kappa_t)}{TFP_i(\lambda(\alpha_0,\kappa_0),\kappa_0)}$	$\ln \frac{TFP_i(\lambda(\alpha_t,\kappa_t),\kappa_t)}{TFP_i(\lambda(\alpha_t,\kappa_0),\kappa_0)}$
Spain	0.032	0.045	0.075	0.052
France	0.031	0.040	0.074	0.057
Ecuador	0.016	0.021	0.048	0.040
Paraguay	0.000	0.043	0.066	0.050
Zambia	-0.033	-0.133	-0.161	-0.194
New Zealand	-0.060	0.030	0.096	0.070
Syria	-0.070	-0.005	0.022	0.038
Mexico	-0.073	0.070	0.095	0.080
Morocco	-0.095	0.030	0.071	0.032
Germany	-0.099	0.035	0.055	0.044
Guatemala	-0.109	0.005	0.053	0.041
Tanzania	-0.112	-0.002	0.133	0.050
Chile	-0.120	0.034	0.071	0.083
Canada	-0.122	0.065	0.091	0.096
Argentina	-0.133	0.001	0.015	0.012
Colombia	-0.141	0.006	0.024	0.017
Philippines	-0.154	0.070	0.117	0.106
Dominican Republic	-0.189	-0.003	0.036	0.017
Switzerland	-0.198	0.014	0.009	0.021
United States	-0.199	0.018	0.045	0.041
South Africa	-0.201	0.005	0.034	0.024
Costa Rica	-0.204	-0.008	0.025	0.039
Jamaica	-0.220	-0.014	-0.007	-0.043
Kenya	-0.245	-0.019	0.019	-0.010
Peru	-0.253	-0.033	-0.068	-0.051
Senegal	-0.257	0.006	0.038	-0.030
Bolivia	-0.264	-0.005	-0.001	0.005
Honduras	-0.278	0.001	0.077	0.064
Uganda	-0.304	-0.005	0.156	0.031
Jordan	-0.334	-0.014	0.034	0.034
Niger	-0.342	0.010	0.079	0.017
Ghana	-0.398	0.005	0.002	0.001

References

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.

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.

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.

Feenstra, R. C., R. Inklaar, and M. P. Timmer (2015): "The next generation of the Penn World Table," The American Economic Review, 105, 3150–3182.

Lederman, D. and L. Saenz (2005): "Innovation and development around the world, 1960-2000," Policy Research Working Paper Series 3774, The World Bank.

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

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