

Online Appendix for “Technology Gaps, Trade and Income”

Thomas Sampson

October 2022

A Proofs and Derivations

A.1 Equilibrium definition from Section I.D

The representative consumer in country s with initial assets a_s chooses a consumption path to maximize utility subject to the budget constraint:

$$(38) \quad \dot{a}_s = \iota_s a_s + w_s - z_s c_s.$$

Solving the intertemporal optimization problem gives the Euler equation:

$$(39) \quad \frac{\dot{c}_s}{c_s} = \iota_s - \rho - \frac{\dot{z}_s}{z_s}.$$

The transversality condition for intertemporal optimization in country s is:

$$(40) \quad \lim_{\tilde{t} \rightarrow \infty} \left\{ a_s(\tilde{t}) \exp \left[- \int_t^{\tilde{t}} \iota_s(\hat{t}) d\hat{t} \right] \right\} = 0.$$

Aggregate consumption in country s is given by:

$$c_s L_s = \prod_{j=1}^J \left(\frac{X_{js}}{\mu_j} \right)^{\mu_j}, \text{ with } X_{js} = \left(\sum_{\tilde{s}=1}^S x_{j\tilde{s}s}^{\frac{\sigma-1}{\sigma}} \right)^{\frac{\sigma}{\sigma-1}} \text{ and } \sum_{j=1}^J \mu_j = 1,$$

where X_{js} denotes consumption of industry j output in country s and $x_{j\tilde{s}s}$ is industry j output from country \tilde{s} that is consumed in country s . Solving consumers' intratemporal optimization problem yields:

$$(41) \quad P_{js} X_{js} = \mu_j z_s c_s L_s,$$

$$(42) \quad z_s = \prod_{j=1}^J P_{js}^{\mu_j},$$

$$(43) \quad x_{j\bar{s}s} = \left(\tau_{j\bar{s}s} \frac{p_{j\bar{s}}}{P_{js}} \right)^{-\sigma} X_{js},$$

$$(44) \quad P_{js} = \left(\sum_{\bar{s}=1}^S \tau_{j\bar{s}s}^{1-\sigma} p_{j\bar{s}}^{1-\sigma} \right)^{\frac{1}{1-\sigma}}.$$

Summing up across firms using (2) we have that aggregate production employment L_{js}^P in industry j and country s is:

$$(45) \quad L_{js}^P = M_{js} \left(\frac{\beta p_{js}}{w_s} \right)^{\frac{1}{1-\beta}} \int_{\theta} \theta^{\frac{1}{1-\beta}} dH_{js}(\theta),$$

where $H_{js}(\theta)$ denotes the cumulative distribution function of productivity. Similarly, aggregate output is:

$$(46) \quad Y_{js} = M_{js} \left(\frac{\beta p_{js}}{w_s} \right)^{\frac{\beta}{1-\beta}} \int_{\theta} \theta^{\frac{1}{1-\beta}} dH_{js}(\theta).$$

Let $l_{js}^R(\psi, \theta)$ and $l_{js}^A(\psi, \theta)$ denote the optimal R&D and adoption employment of a firm with capability ψ and productivity θ . Then $L_{js}^R = M_{js} \int_{(\psi, \theta)} l_{js}^R(\psi, \theta) d\tilde{H}_{js}(\psi, \theta)$ gives aggregate R&D employment, while $L_{js}^A = M_{js} \int_{(\psi, \theta)} l_{js}^A(\psi, \theta) d\tilde{H}_{js}(\psi, \theta)$ gives aggregate adoption employment. The labor market clearing condition in each country s is:

$$(47) \quad L_s = \sum_{j=1}^J (L_{js}^P + L_{js}^R + L_{js}^A + L_{js}^E),$$

where L_{js}^E is aggregate employment in entry.

Output market clearing requires that domestic output Y_{js} equals the sum of sales to all countries inclusive of the iceberg trade costs:

$$(48) \quad Y_{js} = \sum_{\tilde{s}=1}^S \tau_{j\tilde{s}} x_{j\tilde{s}}.$$

Asset market clearing requires that total asset holdings equal the aggregate value of all domestic firms:

$$(49) \quad a_s L_s = \sum_{j=1}^J M_{js} \int_{(\psi, \theta)} V_{js}(\psi, \theta) d\tilde{H}_{js}(\psi, \theta).$$

An equilibrium of the global economy is defined by time paths for consumption per capita c_s , assets per capita a_s , the wage w_s , the interest rate ι_s , the consumption price z_s , consumption levels X_{js} and $x_{j\tilde{s}}$, prices P_{js} and p_{js} , production employment L_{js}^P , industry output Y_{js} , the mass of firms M_{js} , knowledge levels χ_{js}^R and χ_{js}^A , global knowledge capital χ_j , R&D employment L_{js}^R , adoption employment L_{js}^A , entry employment L_{js}^E and the joint distribution of firms' capabilities and productivity levels $\tilde{H}_{js}(\psi, \theta)$ for all countries s , $\tilde{s} = 1, \dots, S$ and all industries $j = 1, \dots, J$ such that: (i) individuals choose consumption per capita to maximize utility subject to the budget constraint (38) giving the Euler equation (39) and the transversality condition (40); (ii) individuals' intratemporal consumption choices imply consumption levels and prices satisfy (41)-(44); (iii) firms choose production employment to maximize production profits implying industry level production employment and output are given by (45) and (46), respectively; (iv) firms' productivity levels evolve according to the R&D technology (3) and the adoption technology (7) and firms choose R&D and adoption employment to maximize their value (8); (v) the R&D and adoption knowledge levels are given by (4) and $\chi_{js}^A = \eta \chi_{js}^R$; (vi) global knowledge capital evolves according to (5); (vii) there is free entry and entrants draw capability and productivity levels from the joint distribution $\tilde{H}_{js}(\psi, \theta)$ implying the free entry condition (9) holds and the mass of firms evolves according to (10), and; (viii) labor, output and asset market clearing imply (47)-(49) hold.

A.2 Growth rates on balanced growth path from Section II

The first step in solving the model is to derive a set of restrictions on equilibrium growth rates that must hold on any balanced growth path. Let g_j be the growth rate of global knowledge capital χ_j . Differentiating (4) and $\chi_{js}^A = \eta \chi_{js}^R$ yields:

$$\frac{\dot{\chi}_{js}^A}{\chi_{js}^A} = \frac{\dot{\chi}_{js}^R}{\chi_{js}^R} = \frac{\kappa_j}{1 + \kappa_j} \frac{\dot{\theta}_{js}^{\max}}{\theta_{js}^{\max}} + \frac{g_j}{1 + \kappa_j}.$$

It follows that on a balanced growth path the productivity frontier θ_{js}^{\max} , together with the R&D and adoption knowledge levels, must grow at constant rate g_j in all countries.³⁶ Consequently, the productivity distribution $H_{js}(\theta)$ shifts outwards at rate g_j for all s . This means $H_{js}(\theta, t) = H_{js}\left(e^{g_j(\tilde{t}-t)}\theta, \tilde{t}\right)$ for all times t, \tilde{t} and productivity levels θ . The productivity growth rate of each industry is constant across countries because $\kappa_j < \infty$ ensures the existence of some global knowledge spillovers.

Now let q_s be the growth rate of consumption per capita c_s . On a balanced growth path the individual's budget constraint (38) implies:

$$(50) \quad \frac{\dot{w}_s}{w_s} = \frac{\dot{a}_s}{a_s} = q_s + \frac{\dot{z}_s}{z_s}.$$

while substituting the free entry condition (9) into the asset market clearing condition (49) gives:

$$a_s L_s = \sum_{j=1}^J M_{js} f^E w_s.$$

Since there is no population growth it follows that $\dot{M}_{js} = 0$.

Next, the growth rate of production employment can be obtained by differentiating (45). Since the productivity distribution $H_{js}(\theta)$ shifts outwards at rate g_j this yields:

$$\frac{\dot{L}_{js}^P}{L_{js}^P} = \frac{1}{1-\beta} \left(\frac{\dot{p}_{js}}{p_{js}} - \frac{\dot{w}_s}{w_s} + g_j \right).$$

On a balanced growth path $\dot{L}_{js}^P = 0$. Therefore, substituting (50) into the expression above we obtain:

$$(51) \quad q_s = \frac{\dot{p}_{js}}{p_{js}} + g_j - \frac{\dot{z}_s}{z_s}.$$

Now, differentiating the industry price index (44) yields:

$$\frac{\dot{P}_{js}}{P_{js}} = \frac{\sum_{\tilde{s}=1}^S \tau_{j\tilde{s}s}^{1-\sigma} p_{j\tilde{s}}^{1-\sigma} \frac{\dot{p}_{j\tilde{s}}}{p_{j\tilde{s}}}}{P_{js}^{1-\sigma}},$$

which is time invariant if and only if output prices p_{js} grow at the same rate in all countries implying:

³⁶To see this, note that the R&D technology (3) implies balanced growth is possible only if the productivity frontier and the R&D knowledge level grow at the same rate in each country.

$$(52) \quad \frac{\dot{P}_{js}}{P_{js}} = \frac{\dot{p}_{j\tilde{s}}}{p_{j\tilde{s}}},$$

for all $s, \tilde{s} = 1, \dots, S$. Differentiating the consumption price equation (42) gives:

$$\frac{\dot{z}_s}{z_s} = \sum_{j=1}^J \mu_j \frac{\dot{P}_{js}}{P_{js}}.$$

Multiplying both sides of (51) by μ_j , summing across industries and using the previous expression, (52) and $\sum_{j=1}^J \mu_j = 1$ we obtain:

$$\sum_{j=1}^J \mu_j q_s = q_s = \sum_{j=1}^J \mu_j g_j,$$

which shows that the growth rate of consumption per capita is the same in all countries. The numeraire condition $\sum_{s=1}^S z_s c_s L_s = 1$ then implies:

$$(53) \quad \frac{\dot{z}_s}{z_s} = -q,$$

and substituting this result into (51) shows that output prices p_{js} and, therefore, also industry prices P_{js} decline at rate g_j . Note also that using (41) to substitute for X_{js} in (43) and appealing to (53) together with the fact prices decline at rate g_j implies $x_{j\tilde{s}s}$ grows at rate g_j . It then follows from the industry output market clearing condition (48) that industry output Y_{js} also grows at rate g_j .

Finally, substituting (53) into the Euler equation (39) yields that the interest rate is time invariant, constant across countries and given by $\iota_s = \rho$. Since the discount rate $\rho > 0$ and nominal assets per capita remain constant over time, the transversality condition (40) is satisfied.

Collecting together the results above, we have on a balanced growth path the growth rate of consumption per capita $q = \sum_{j=1}^J \mu_j g_j$ is the same in all countries and equals a weighted average of productivity growth in the J industries where the weights are given by the industry expenditure shares. Consumption prices z_s decline at rate q , while nominal wages w_s and assets per capita a_s remain constant over time. This implies real wages and assets per capita grow at rate q . Employment in production, R&D, adoption and entry in each country-industry pair is time invariant, as is the mass of firms M_{js} . Industry output Y_{js} and the quantity sold in each market $x_{j\tilde{s}s}$ grow at rate g_j , while prices p_{js} and P_{js} decline at rate g_j .

A.3 Solution to firm's R&D problem in Section II.A

Firms take the time paths of w_s , p_{js} , χ_{js}^R , χ_{js}^A and ι_s as given. In particular, suppose the economy is on a balanced growth path, implying w_s is time invariant, p_{js} declines at rate g_j , χ_{js}^R and χ_{js}^A both grow at rate g_j , and $\iota_s = \rho$.

Taking the time derivative of ϕ and using the R&D technology (3) implies:

$$(54) \quad \frac{\dot{\phi}}{\phi} = \frac{1}{1-\beta} [\psi B_s \phi^{-\gamma_j(1-\beta)} (l^R)^\alpha - (\delta + g_j)].$$

Substituting the production profits function (2) into the value function (8), using $\iota_s = \rho$ and changing variables from θ to ϕ , the optimization problem of a firm with capability ψ can be written as:

$$\max_{\phi, l^R} \int_t^\infty e^{-(\rho+\zeta)(\tilde{t}-t)} w_s \left[\frac{1-\beta}{\beta} \left(\frac{\beta p_{js} \chi_{js}^R}{w_s} \right)^{\frac{1}{1-\beta}} \phi - l^R \right] d\tilde{t},$$

subject to the growth of ϕ being given by (54) and an initial value for ϕ at time t . Since w_s is constant, p_{js} declines at rate g_j and χ_{js}^R grows at rate g_j , the payoff function depends upon time only through exponential discounting meaning the firm faces a discounted infinite-horizon optimal control problem of the type studied in Section 7.5 of Acemoglu (2009) with state variable ϕ and control variable l^R .

The current-value Hamiltonian for the firm's problem is:

$$\mathcal{H}(\phi, l^R, \lambda) = \left[\frac{1-\beta}{\beta} \left(\frac{\beta p_{js} \chi_{js}^R}{w_s} \right)^{\frac{1}{1-\beta}} \phi - l^R \right] w_s + \lambda \frac{\phi}{1-\beta} [\psi B_s \phi^{-\gamma_j(1-\beta)} (l^R)^\alpha - (\delta + g_j)],$$

where λ is the current-value costate variable. From Theorem 7.13 in Acemoglu (2009), any solution must satisfy:

$$(55) \quad 0 = \frac{\partial \mathcal{H}}{\partial l^R} = -w_s + \lambda \frac{\alpha}{1-\beta} \psi B_s \phi^{1-\gamma_j(1-\beta)} (l^R)^{\alpha-1},$$

$$(56) \quad \begin{aligned} (\rho + \zeta) \lambda - \dot{\lambda} &= \frac{\partial \mathcal{H}}{\partial \phi} = \frac{1-\beta}{\beta} \left(\frac{\beta p_{js} \chi_{js}^R}{w_s} \right)^{\frac{1}{1-\beta}} w_s \\ &\quad + \frac{\lambda}{1-\beta} \{ [1 - \gamma_j(1-\beta)] \psi B_s \phi^{-\gamma_j(1-\beta)} (l^R)^\alpha - (\delta + g_j) \}, \\ 0 &= \lim_{\tilde{t} \rightarrow \infty} \left[e^{-(\rho+\zeta)(\tilde{t}-t)} \mathcal{H}(\phi, l^R, \lambda) \right], \end{aligned}$$

where equation (56) is the transversality condition. Differentiating the upper expression with respect to time gives:

$$(57) \quad (1-\alpha) \frac{\dot{l}^R}{l^R} = [1 - \gamma_j(1-\beta)] \frac{\dot{\phi}}{\phi} + \frac{\dot{\lambda}}{\lambda},$$

and using the first order conditions of the Hamiltonian to substitute for λ and $\dot{\lambda}$, and (54) to substitute for $\dot{\phi}$ yields:

$$(58) \quad \frac{\dot{l}^R}{l^R} = \frac{1}{1-\alpha} \left[\rho + \zeta + \gamma_j (\delta + g_j) - \alpha \beta^{\frac{\beta}{1-\beta}} \psi B_s \left(\frac{p_{js} \chi_{js}^R}{w_s} \right)^{\frac{1}{1-\beta}} \phi^{1-\gamma_j(1-\beta)} (l^R)^{\alpha-1} \right].$$

Equations (54) and (58) are an autonomous nonlinear system of differential equations in (ϕ, l^R) whose unique steady state $(\phi_{js}^*, l_{js}^{R*})$ is given by (12) and (13). Suppose we write the system as:

$$\begin{pmatrix} \dot{\phi} \\ \dot{l}^R \end{pmatrix} = F \begin{pmatrix} \phi \\ l^R \end{pmatrix}.$$

At the steady state, the Jacobian $\mathcal{D}F$ of the function F is:

$$\mathcal{D}F \begin{pmatrix} \phi_{js}^* \\ l_{js}^{R*} \end{pmatrix} = \begin{pmatrix} -\gamma_j (\delta + g_j) & \frac{\alpha}{1-\beta} \frac{\phi_{js}^*}{l_{js}^{R*}} (\delta + g_j) \\ -\frac{1-\gamma_j(1-\beta)}{1-\alpha} \frac{l_{js}^{R*}}{\phi_{js}^*} [\rho + \zeta + \gamma_j (\delta + g_j)] & \rho + \zeta + \gamma_j (\delta + g_j) \end{pmatrix}.$$

The trace of the Jacobian is $\rho + \zeta$ which is positive. The determinant of the Jacobian is:

$$\left| \mathcal{D}F \begin{pmatrix} \phi_{js}^* \\ l_{js}^{R*} \end{pmatrix} \right| = -(\delta + g_j) [\rho + \zeta + \gamma_j (\delta + g_j)] \frac{\gamma_j(1-\beta) - \alpha}{(1-\alpha)(1-\beta)},$$

which is negative by Assumption 1. This means the Jacobian has one strictly negative and one strictly positive eigenvalue. Therefore, by Theorem 7.19 in Acemoglu (2009), the steady state is locally saddle-path stable. There exists an open neighborhood of the steady state such that if the firm's initial ϕ lies within this neighborhood, the system of differential equations given by (54) and (58) has a unique solution. The solution converges to the steady state along the stable arm of the system as shown in Figure 1 in the paper. From equation (57) it follows that $\dot{\lambda} \rightarrow 0$ as the solution converges to the steady state. Since $\rho + \zeta > 0$ this implies the solution satisfies the transversality condition (56).

The solution to (54) and (58) is a candidate for a solution to the firm's problem. To show it is in fact the unique solution we can use Theorem 7.14 in Acemoglu (2009). Suppose λ is the current-value costate variable obtained from the solution to (54) and (58). Equation (55) implies λ is always strictly positive. Therefore, given any path for ϕ on which ϕ is always positive we have $\lim_{\bar{t} \rightarrow \infty} \left[e^{-(\rho+\zeta)(\bar{t}-t)} \lambda \phi \right] \geq 0$. Now define:

$$\begin{aligned} \bar{\mathcal{H}}(\phi, \lambda) &= \max_{l^R} \mathcal{H}(\phi, l^R, \lambda), \\ &= \left[\frac{1-\beta}{\beta} \left(\frac{\beta p_{js} \chi_{js}^R}{w_s} \right)^{\frac{1}{1-\beta}} w_s - \frac{\lambda(\delta + g_j)}{1-\beta} \right] \phi + \frac{1-\alpha}{\alpha} w_s^{\frac{-\alpha}{1-\alpha}} \left(\frac{\alpha \lambda \psi B_s}{1-\beta} \right)^{\frac{1}{1-\alpha}} \phi^{\frac{1-\gamma_j(1-\beta)}{1-\alpha}}, \end{aligned}$$

where the second line follows from solving the maximization problem in the first line. Assumption 1 implies $\bar{\mathcal{H}}(\phi, \lambda)$ is strictly concave in ϕ . Thus, the sufficiency conditions of Theorem 7.14 in Acemoglu (2009) hold, implying the solution to (54) and (58) is the unique solution to the firm's optimal control problem.

A.4 Proof of Proposition 1

On a balanced growth path the productivity distribution $H_{js}(\theta)$ must shift outwards at rate g_j . The evolution of $H_{js}(\theta)$ depends upon productivity growth at surviving firms and how the productivity distribution of entrants compares to that of exiting firms. Entrants draw their capability and productivity from the joint distribution of ψ and θ among incumbents and all incumbents face instantaneous exit probability ζ . Therefore, if all incumbent firms are in steady state, each new firm enters with its steady state productivity level and net entry does not affect $H_{js}(\theta)$. Since all surviving firms grow at rate g_j in steady state, it follows that firm-level productivity dynamics are consistent with balanced growth if and only if all incumbent firms are in steady state.

Parts (i) and (ii) of the proposition follow immediately from the solution of the firm's intertemporal optimization problem in Section II.A. For part (iii), consider two firms in the same country

and industry with capabilities ψ and ψ' , respectively. The ratio of these firms' steady state productivity levels is:

$$\frac{\theta_{js}^*(\psi')}{\theta_{js}^*(\psi)} = \begin{cases} \left(\frac{\psi'}{\psi}\right)^{\frac{1-\beta}{\gamma_j(1-\beta)-\alpha}}, & \psi' \geq \psi \geq \psi_{js}^*, \\ \left(\frac{\psi'}{\psi_{js}^*}\right)^{\frac{1-\beta}{\gamma_j(1-\beta)-\alpha}}, & \psi' \geq \psi_{js}^* \geq \psi, \\ 1, & \psi_{js}^* \geq \psi' \geq \psi. \end{cases}$$

When both firms perform R&D, technology gaps and productivity inequality are strictly increasing in α and β and strictly decreasing in γ_j . Conditional on ψ_{js}^* , productivity inequality between R&D and adoption firms is also strictly increasing in α and β and strictly decreasing in γ_j . There is no productivity inequality within adopters. However, a higher advantage of backwardness or a lower R&D efficiency reduces industry-level productivity inequality by increasing ψ_{js}^* and decreasing the fraction of firms that choose R&D.

Combining these results, it follows that aggregate productivity inequality within each country-industry pair is strictly increasing in α , β and B_s and strictly decreasing in γ_j . From (2) inequality in production employment, revenue and profits are also strictly increasing in α , β and B_s and strictly decreasing in γ_j .

A.5 Derivation of balanced growth path equilibrium equations (16)-(18)

Suppose the global economy is on a balanced growth path. Using (2), (8), (12) and (13) implies that on a balanced growth path the steady state value of a firm with capability $\psi \geq \psi_{js}^*$ is:

$$V_{js}(\psi, \theta_{js}^*) = \left(1 - \beta - \frac{\alpha(\delta + g_j)}{\rho + \zeta + \gamma_j(\delta + g_j)}\right) \frac{w_s}{\rho + \zeta} \times \left[\alpha^\alpha \beta \gamma_j^\beta B_s \psi \left(\frac{p_{js} \chi_{js}^R}{w_s}\right)^{\gamma_j} \frac{(\delta + g_j)^{\alpha-1}}{[\rho + \zeta + \gamma_j(\delta + g_j)]^\alpha} \right]^{\frac{1}{\gamma_j(1-\beta)-\alpha}},$$

where $\theta_{js}^* = \chi_{js}^R (\phi_{js}^*)^{1-\beta}$ is the firm's steady state productivity, which is growing over time. The steady state value of firms with capability $\psi \leq \psi_{js}^*$, which choose adoption, is given by the same expression, but with $\psi = \psi_{js}^*$. Assumption 1 implies $1 - \beta > \frac{\alpha(\delta + g_j)}{\rho + \zeta + \gamma_j(\delta + g_j)}$ which ensures $V_{js}(\psi, \theta_{js}^*)$ is positive.

Section II.A showed that on a balanced growth path each new firm enters with the steady state productivity level corresponding to its capability. Since entrants' capabilities have distribution $G(\psi)$, substituting the above expression for $V_{js}(\psi, \theta_{js}^*)$ into the free entry condition (9) yields:

$$(59) \quad f^E = \left(1 - \beta - \frac{\alpha(\delta + g_j)}{\rho + \zeta + \gamma_j(\delta + g_j)}\right) \frac{\Psi_{js}}{\rho + \zeta} \\ \times \left[\alpha^\alpha \beta^{\gamma_j \beta} B_s \left(\frac{p_{js} \chi_{js}^R}{w_s} \right)^{\gamma_j} \frac{(\delta + g_j)^{\alpha-1}}{[\rho + \zeta + \gamma_j(\delta + g_j)]^\alpha} \right]^{\frac{1}{\gamma_j(1-\beta)-\alpha}}.$$

Next, observe that on a balanced growth path:

$$\int_{\theta} \theta^{\frac{1}{1-\beta}} dH_{js}(\theta) = \int_{\psi^{\min}}^{\psi^{\max}} (\chi_{js}^R)^{\frac{1}{1-\beta}} \phi_{js}^* dG(\psi),$$

where ϕ_{js}^* is given by (12) for R&D firms and by (12) with $\psi = \psi_{js}^*$ for adopters. Thus, by substituting (13) and (45) into the labor market clearing condition (47) and using (10) with $\dot{M}_{js} = 0$ to solve for L_{js}^E we obtain:

$$(60) \quad L_s = \sum_{j=1}^J M_{js} \left\{ \left(1 + \frac{\alpha}{\beta} \frac{\delta + g_j}{\rho + \zeta + \gamma_j(\delta + g_j)}\right) \Psi_{js} \right. \\ \left. \times \left[\alpha^\alpha \beta^{\gamma_j - \alpha} B_s \left(\frac{p_{js} \chi_{js}^R}{w_s} \right)^{\gamma_j} \frac{(\delta + g_j)^{\alpha-1}}{[\rho + \zeta + \gamma_j(\delta + g_j)]^\alpha} \right]^{\frac{1}{\gamma_j(1-\beta)-\alpha}} + f^E \zeta \right\}.$$

Similarly, substituting (12), (41), (43) and (46) into the goods market clearing condition (48) and using (59) we obtain:

$$(61) \quad \sum_{\bar{s}=1}^S \left(\frac{\tau_{j\bar{s}} p_{j\bar{s}}}{P_{j\bar{s}}} \right)^{1-\sigma} \mu_{j\bar{s}} z_{\bar{s}} c_{\bar{s}} L_{\bar{s}} = f^E (\rho + \zeta) \left(1 - \beta - \frac{\alpha(\delta + g_j)}{\rho + \zeta + \gamma_j(\delta + g_j)}\right)^{-1} M_{js} w_s.$$

On a balanced growth path $\dot{a}_s = 0$ and $\iota_s = \rho$. Therefore, the individual's budget constraint implies:

$$(62) \quad z_s c_s = \rho a_s + w_s,$$

while substituting the free entry condition (9) into the asset market clearing condition (49) gives:

$$(63) \quad a_s L_s = \sum_{j=1}^J M_{js} w_s f^E.$$

Equations (59)-(63) together with R&D knowledge levels (4), knowledge capital growth rates (5), consumption prices (42) and industry price indices (44) form a system of $4JS + 4S + J$ equations. Together with the numeraire condition $\sum_{s=1}^S z_s c_s L_s = 1$, the steady state relative productivity levels in (12) and the initial global knowledge capital in each industry χ_j these equations determine the $4JS + 4S + J$ unknowns $w_s, a_s, c_s, z_s, g_j, p_{js}, P_{js}, M_{js}$ and χ_{js}^R for all industries $j = 1, \dots, J$ and all countries $s = 1, \dots, S$.

To simplify this system, start by substituting (59) and (61) into (60) giving:

$$(64) \quad L_s = \sum_{j=1}^J \frac{\mu_j}{\rho + \zeta} \left(\zeta + \beta\rho + \frac{\alpha\rho(\delta + g_j)}{\rho + \zeta + \gamma_j(\delta + g_j)} \right) \sum_{\bar{s}=1}^S \left(\frac{\tau_{j\bar{s}} p_{j\bar{s}}}{P_{j\bar{s}}} \right)^{1-\sigma} \frac{z_{\bar{s}} c_{\bar{s}} L_{\bar{s}}}{w_s}.$$

Using (44) to obtain the industry price index, (59) to substitute for p_{js} , (4) to give χ_{js}^R and (12) to solve for relative steady state productivity levels then implies:

$$(65) \quad \left(\frac{p_{js}}{P_{j\bar{s}}} \right)^{1-\sigma} = \frac{w_s^{1-\sigma} \left(B_s \Psi_{js}^{\frac{\gamma_j(1-\beta)}{1+\kappa_j} - \alpha} \right)^{\frac{(\sigma-1)(1+\kappa_j)}{\gamma_j}}}{\sum_{\bar{s}=1}^S \tau_{j\bar{s}}^{1-\sigma} w_{\bar{s}}^{1-\sigma} \left(B_{\bar{s}} \Psi_{j\bar{s}}^{\frac{\gamma_j(1-\beta)}{1+\kappa_j} - \alpha} \right)^{\frac{(\sigma-1)(1+\kappa_j)}{\gamma_j}}}.$$

Substituting this expression into (64) and using (62) yields equation (16). Equation (17) can be derived in a similar manner by substituting (61) and (65) into the asset market clearing condition (63). Finally, substituting steady state R&D employment (13) together with (59), (61) and (65) into (5) yields equation (18).

A.6 Proof of existence and uniqueness of balanced growth path in single sector economy with free trade

Equations (16)-(18) are a non-linear system of $2S + J$ equations in the unknown wages w_s , asset holdings a_s and growth rates g_j . Existing methods are insufficient to prove this system has a unique solution. Therefore, to establish sufficient conditions for a unique balanced growth, I impose

assumptions that make this system separable in w_s , a_s and g_j . Free trade implies that asset holdings a_s can be eliminated from the labor market clearing equation (16). And setting $J = 1$ implies that w_s and a_s can be eliminated from the growth equation (18).

I will start by proving that, under free trade, equations (16) and (17) yield a unique solution for w_s and a_s given growth rates g_j . I will then show that, when the economy has a single sector, there exists a unique equilibrium growth rate. Together these results imply that a single sector economy with free trade has a unique balanced growth path.

Using the numeraire condition $\sum_{s=1}^S z_s c_s L_s = 1$ and equation (62) gives $\sum_{s=1}^S (\rho a_s + w_s) L_s = 1$. Substituting this expression into (19) with $\tau_{j\tilde{s}} = 1$ for all j, s, \tilde{s} gives:

$$(66) \quad Z_{js} = \frac{w_s^{-\sigma} \left(B_s \Psi_{js}^{\frac{\gamma_j(1-\beta)}{1+\kappa_j} - \alpha} \right)^{\frac{(\sigma-1)(1+\kappa_j)}{\gamma_j}}}{\sum_{\tilde{s}=1}^S w_{\tilde{s}}^{1-\sigma} \left(B_{\tilde{s}} \Psi_{j\tilde{s}}^{\frac{\gamma_j(1-\beta)}{1+\kappa_j} - \alpha} \right)^{\frac{(\sigma-1)(1+\kappa_j)}{\gamma_j}}},$$

and using this expression in (16) implies that the S -dimensional wage vector $\mathbf{w} = (w_1, \dots, w_S)$ satisfies $\mathbf{f}(\mathbf{w}) = 0$ where $\mathbf{f} : \mathbb{R}_{++}^S \rightarrow \mathbb{R}^S$ and element s of the vector \mathbf{f} is given by:

$$f_s(\mathbf{w}) = \sum_{j=1}^J \frac{\mu_j}{\rho + \zeta} \left(\zeta + \beta\rho + \frac{\alpha\rho(\delta + g_j)}{\rho + \zeta + \gamma_j(\delta + g_j)} \right) \frac{w_s^{-\sigma} \left(B_s \Psi_{js}^{\frac{\gamma_j(1-\beta)}{1+\kappa_j} - \alpha} \right)^{\frac{(\sigma-1)(1+\kappa_j)}{\gamma_j}}}{\sum_{\tilde{s}=1}^S w_{\tilde{s}}^{1-\sigma} \left(B_{\tilde{s}} \Psi_{j\tilde{s}}^{\frac{\gamma_j(1-\beta)}{1+\kappa_j} - \alpha} \right)^{\frac{(\sigma-1)(1+\kappa_j)}{\gamma_j}}} - L_s.$$

Suppose the growth rates g_j for $j = 1, \dots, J$ are known. To prove that $\mathbf{f}(\mathbf{w}) = 0$ implies a unique solution for wages I use results from Allen, Arkolakis and Li (2015). For all $s = 1, \dots, S$ define the scaffold function $\mathbf{F} : \mathbb{R}_{++}^{S+1} \rightarrow \mathbb{R}^S$ by:

$$F_s(\tilde{\mathbf{w}}, w_s) = \sum_{j=1}^J \frac{\mu_j}{\rho + \zeta} \left(\zeta + \beta\rho + \frac{\alpha\rho(\delta + g_j)}{\rho + \zeta + \gamma_j(\delta + g_j)} \right) \frac{w_s^{-\sigma} \left(B_s \Psi_{js}^{\frac{\gamma_j(1-\beta)}{1+\kappa_j} - \alpha} \right)^{\frac{(\sigma-1)(1+\kappa_j)}{\gamma_j}}}{\sum_{\tilde{s}=1}^S \tilde{w}_{\tilde{s}}^{1-\sigma} \left(B_{\tilde{s}} \Psi_{j\tilde{s}}^{\frac{\gamma_j(1-\beta)}{1+\kappa_j} - \alpha} \right)^{\frac{(\sigma-1)(1+\kappa_j)}{\gamma_j}}} - L_s.$$

Note that $f_s(\mathbf{w}) = F_s(\mathbf{w}, w_s)$ for all s and the function \mathbf{F} is continuously differentiable.

To prove existence it is now sufficient to show that conditions (i)-(iii) of Lemma 1 in Allen, Arkolakis and Li (2015) are satisfied. Condition (i) follows from observing that, for any $\tilde{\mathbf{w}}$, $F_s(\tilde{\mathbf{w}}, w_s)$ is strictly decreasing in w_s , positive for w_s sufficiently close to zero and negative for w_s sufficiently large. To see that condition (ii) holds, note that $1 - \sigma < 0$ implying $F_s(\tilde{\mathbf{w}}, w_s)$ is strictly increasing in $\tilde{w}_{\hat{s}}$ for all \hat{s} .

Now, given $\lambda > 0$ and $\tilde{\mathbf{w}} \in \mathbb{R}_{++}^S$ define $w_s(\lambda)$ by $F_s[\lambda\tilde{\mathbf{w}}, w_s(\lambda)] = 0$. Let $u \in (0, 1)$ be such that $-1 + \sigma u < 0$. Then $F_s[\lambda\tilde{\mathbf{w}}, \lambda^{1-u}w_s(1)]$ is strictly negative if $\lambda > 1$ and strictly positive if $\lambda < 1$. Since $F_s(\tilde{\mathbf{w}}, w_s)$ is strictly decreasing in w_s it follows that $w_s(\lambda) < \lambda^{1-u}w_s(1)$ if $\lambda > 1$ and $w_s(\lambda) > \lambda^{1-u}w_s(1)$ if $\lambda < 1$. Therefore, when $\lambda \rightarrow \infty$, $\frac{\lambda}{w_s(\lambda)} \rightarrow \infty$ and when $\lambda \rightarrow 0$, $\frac{\lambda}{w_s(\lambda)} \rightarrow 0$ implying condition (iii) holds. Thus, a solution exists.

To prove uniqueness I use Theorem 2 in Allen, Arkolakis and Li (2015). Since $f_s(\mathbf{w})$ is strictly increasing in $w_{\hat{s}}$ whenever $\hat{s} \neq s$, $\mathbf{f}(\mathbf{w})$ satisfies gross substitution. Also, $f_s(\mathbf{w})$ can be written as $f_s(\mathbf{w}) = \tilde{f}_s(\mathbf{w}) - L_s$ where $\tilde{f}_s(\mathbf{w})$ is positive and homogeneous of degree minus one, while L_s is positive and homogeneous of degree zero in \mathbf{w} . Consequently, Theorem 2 in Allen, Arkolakis and Li (2015) implies the solution is unique.

Using the solution for wages and equation (66) for Z_{js} , assets a_s are given immediately by (17). This completes the proof that under free trade there exists a unique solution for w_s and a_s given growth rates g_j .

Now suppose the economy has a single sector. Setting $J = 1$ and substituting (16) into (18) yields:

$$(67) \quad \frac{g[\rho + \zeta + \gamma(\delta + g)]}{\alpha(\rho + \zeta)(\delta + g)} \left(\zeta + \beta\rho + \frac{\alpha\rho(\delta + g)}{\rho + \zeta + \gamma(\delta + g)} \right) = \sum_{s=1}^S \frac{L_s}{\Psi_s} \int_{\psi_s^*}^{\psi^{\max}} \lambda_s(\psi) \psi^{\frac{1}{\gamma(1-\beta)-\alpha}} dG(\psi).$$

This expression holds regardless of whether there are trade costs. The left hand side is a strictly increasing function of g with range $[0, \infty)$, while the right hand side is a positive constant. Thus, there exists a unique equilibrium productivity growth rate g and, with a single sector, the consumption growth rate also equals g . It follows immediately that, if $J = 1$ and there are no trade costs, the global economy has a unique balanced growth path.

Equation (67) can be used to characterize the determinants of the equilibrium growth rate in a single sector economy. Growth is higher when R&D spillovers $\lambda_s(\cdot)$ are stronger and when there is more employment in R&D. This generates a scale effect whereby growth is increasing in the size L_s of each country. It also implies growth is increasing in the R&D efficiency B_s of each country because higher R&D efficiency reduces the R&D threshold ψ_s^* . Similarly, growth declines

when adoption becomes more attractive relative to R&D due to an increase in either the adoption knowledge premium η or adoption efficiency B^A .

Growth is higher in the open economy than in autarky because the R&D spillovers specified in (5) are global in scope. However, growth does not depend upon the localization of knowledge spillovers κ , which affects countries' relative knowledge levels, but not the rate of increase of global knowledge capital. The growth rate is also independent of the level of trade costs.

A.7 Proof of Proposition 2

To derive (20) start by substituting the free entry condition (59) into (12) and using $\theta_{js}^* = \chi_{js}^R (\phi_{js}^*)^{1-\beta}$ to obtain:

$$(68) \quad (\theta_{js}^*)^{\frac{1}{1-\beta}} = \left[f^E \frac{\rho + \zeta}{(\delta + g_j)^{\frac{1}{\alpha}}} \left(\frac{1 - \beta \rho + \zeta + \gamma_j(\delta + g_j)}{\alpha \delta + g_j} - 1 \right)^{-1} \frac{B_s^{\frac{1}{\alpha}}}{\Psi_{js}} \right]^{\frac{\alpha}{\gamma_j(1-\beta)}} \psi^{\frac{1}{\gamma_j(1-\beta)-\alpha}} (\chi_{js}^R)^{\frac{1}{1-\beta}},$$

where $\psi = \psi_{js}^*$ for firms that choose adoption. Setting $\psi = \psi^{\max}$ in this expression and using (4) to substitute for χ_{js}^R then implies:

$$\theta_{js}^{*\max} = \left[f^E \frac{\rho + \zeta}{(\delta + g_j)^{\frac{1}{\alpha}}} \left(\frac{1 - \beta \rho + \zeta + \gamma_j(\delta + g_j)}{\alpha \delta + g_j} - 1 \right)^{-1} \frac{B_s^{\frac{1}{\alpha}}}{\Psi_{js}} \right]^{\frac{\alpha(1+\kappa_j)}{\gamma_j}} (\psi^{\max})^{\frac{(1-\beta)(1+\kappa_j)}{\gamma_j(1-\beta)-\alpha}} \chi_j.$$

Substituting this expression and (4) back into (68) and integrating over the capability distribution yields:

$$(69) \quad \bar{\theta}_{js}^* = \left[f^E \frac{\rho + \zeta}{(\delta + g_j)^{\frac{1}{\alpha}}} \left(\frac{1 - \beta \rho + \zeta + \gamma_j(\delta + g_j)}{\alpha \delta + g_j} - 1 \right)^{-1} \right]^{\frac{\alpha(1+\kappa_j)}{\gamma_j}} \times (\psi^{\max})^{\frac{(1-\beta)\kappa_j}{\gamma_j(1-\beta)-\alpha}} \chi_j \left(B_s \Psi_{js}^{\frac{\gamma_j(1-\beta)}{1+\kappa_j}-\alpha} \right)^{\frac{1+\kappa_j}{\gamma_j}},$$

and dividing this equation by the equivalent expression for country \tilde{s} gives (20).

Using (41) and (43) the exports of country s to country \tilde{s} in industry j are given by:

$$EX_{js\tilde{s}} = \tau_{js\tilde{s}}^{1-\sigma} \left(\frac{p_{js}}{P_{j\tilde{s}}} \right)^{1-\sigma} \mu_j z_{\tilde{s}} c_{\tilde{s}} L_{\tilde{s}}.$$

Substituting (65) into this expression and taking logs we obtain equation (23) where:

$$v_{j\bar{s}}^2 = \log(\mu_j z_{\bar{s}} c_{\bar{s}} L_{\bar{s}}) - \log \left[\sum_{\hat{s}=1}^S \tau_{j\hat{s}\bar{s}}^{1-\sigma} w_{\hat{s}}^{1-\sigma} \left(B_{\hat{s}} \Psi_{j\hat{s}}^{\frac{\gamma_j(1-\beta)}{1+\kappa_j} - \alpha} \right)^{\frac{(\sigma-1)(1+\kappa_j)}{\gamma_j}} \right],$$

and substituting (69) into this expression gives equation (22) where:

$$v_{j\bar{s}}^1 = v_{j\bar{s}}^2 - (\sigma-1) \log \left\{ \left[f^E \frac{\rho + \zeta}{(\delta + g_j)^{\frac{1}{\alpha}}} \left(\frac{1-\beta}{\alpha} \frac{\rho + \zeta + \gamma_j(\delta + g_j)}{\delta + g_j} - 1 \right)^{-1} \right]^{\frac{\alpha(1+\kappa_j)}{\gamma_j}} (\psi^{\max})^{\frac{(1-\beta)\kappa_j}{\gamma_j(1-\beta)-\alpha}} \chi_j \right\}.$$

Next, differentiating the definition of Ψ_{js} and using that the R&D threshold ψ_{js}^* is given by (11) yields:

$$\frac{\partial \log \Psi_{js}}{\partial \log B_s} = \frac{-1}{\gamma_j(1-\beta) - \alpha} \frac{(\psi_{js}^*)^{\frac{1}{\gamma_j(1-\beta)-\alpha}} G(\psi_{js}^*)}{\Psi_{js}},$$

and differentiating (69) then implies:

$$\frac{\partial \log \bar{\theta}_{js}^*}{\partial \log B_s} = \frac{1 + \kappa_j}{\gamma_j} \left[1 - \frac{\gamma_j(1-\beta) - \alpha(1 + \kappa_j)}{(1 + \kappa_j)[\gamma_j(1-\beta) - \alpha]} \frac{(\psi_{js}^*)^{\frac{1}{\gamma_j(1-\beta)-\alpha}} G(\psi_{js}^*)}{\Psi_{js}} \right],$$

which is strictly positive. Inspection of this expression shows immediately that $\frac{\partial^2 \log \bar{\theta}_{js}^*}{\partial \kappa_j \partial \log B_s} > 0$ and differentiating with respect to γ_j gives:

$$\begin{aligned} \frac{\partial^2 \log \bar{\theta}_{js}^*}{\partial \gamma_j \partial \log B_s} &= \frac{-1}{\gamma_j} \frac{\partial \log \bar{\theta}_{js}^*}{\partial \log B_s} - \frac{\alpha(1-\beta)\kappa_j}{\gamma_j [\gamma_j(1-\beta) - \alpha]^2} \frac{(\psi_{js}^*)^{\frac{1}{\gamma_j(1-\beta)-\alpha}} G(\psi_{js}^*)}{\Psi_{js}} \\ &\quad - \frac{\gamma_j(1-\beta) - \alpha(1 + \kappa_j)}{\gamma_j [\gamma_j(1-\beta) - \alpha]} \frac{\partial}{\partial \gamma_j} \left[\frac{(\psi_{js}^*)^{\frac{1}{\gamma_j(1-\beta)-\alpha}} G(\psi_{js}^*)}{\Psi_{js}} \right]. \end{aligned}$$

The first two terms on the right hand side of this expression are negative. Computing the derivative in the third term and using the definition of Ψ_{js} to collect terms gives:

$$\begin{aligned} \frac{\partial}{\partial \gamma_j} \left[\frac{(\psi_{js}^*)^{\frac{1}{\gamma_j(1-\beta)-\alpha}} G(\psi_{js}^*)}{\Psi_{js}} \right] &= \frac{(\psi_{js}^*)^{\frac{1}{\gamma_j(1-\beta)-\alpha}} G(\psi_{js}^*)}{\Psi_{js}^2} \left[\frac{\log \eta}{\gamma_j(1-\beta)-\alpha} \int_{\psi_{js}^*}^{\psi^{\max}} \psi^{\frac{1}{\gamma_j(1-\beta)-\alpha}} dG(\psi) \right. \\ &\quad \left. + \log \eta \frac{\psi_{js}^* G'(\psi_{js}^*)}{G(\psi_{js}^*)} \Psi_{js} + \frac{1-\beta}{[\gamma_j(1-\beta)-\alpha]^2} \int_{\psi_{js}^*}^{\psi^{\max}} (\log \psi - \log \psi_{js}^*) \psi^{\frac{1}{\gamma_j(1-\beta)-\alpha}} dG(\psi) \right], \end{aligned}$$

which is positive since $\eta > 1$. It follows that $\frac{\partial^2 \log \bar{\theta}_{js}^*}{\partial \gamma_j \partial \log B_s} < 0$ as claimed in Proposition 2.

A.8 Derivation of balanced growth path consumption prices from Section II.D

From (42) and (44) we have:

$$z_s = \prod_{j=1}^J \left(\sum_{\tilde{s}=1}^S \tau_{j\tilde{s}s}^{1-\sigma} p_{j\tilde{s}}^{1-\sigma} \right)^{\frac{\mu_j}{1-\sigma}},$$

and combining (4), (59) and (68) with $\psi = \psi^{\max}$ gives:

$$\begin{aligned} p_{js} &= \beta^{-\beta} (\psi^{\max})^{\frac{-(1-\beta)\kappa_j}{\gamma_j(1-\beta)-\alpha}} \left[f^E(\rho + \zeta) \left(1 - \beta - \frac{\alpha(\delta + g_j)}{\rho + \zeta + \gamma_j(\delta + g_j)} \right)^{-1} \right]^{\frac{\gamma_j(1-\beta)-\alpha(1+\kappa_j)}{\gamma_j}} \\ &\quad \times \left[\frac{\alpha(\delta + g_j)^{\frac{\alpha-1}{\alpha}}}{\rho + \zeta + \gamma_j(\delta + g_j)} \right]^{\frac{-\alpha(1+\kappa_j)}{\gamma_j}} \frac{w_s}{\chi_j} \left(B_s \Psi_{js}^{\frac{\gamma_j(1-\beta)}{1+\kappa_j} - \alpha} \right)^{-\frac{1+\kappa_j}{\gamma_j}}. \end{aligned}$$

Using these two expressions to obtain the ratio of consumption prices in countries s and \tilde{s} then yields:

$$\frac{z_s}{z_{\tilde{s}}} = \prod_{j=1}^J \left[\frac{\sum_{\hat{s}=1}^S \tau_{j\hat{s}s}^{1-\sigma} w_{\hat{s}}^{1-\sigma} \left(B_{\hat{s}} \Psi_{j\hat{s}}^{\frac{\gamma_j(1-\beta)}{1+\kappa_j} - \alpha} \right)^{\frac{(\sigma-1)(1+\kappa_j)}{\gamma_j}}}{\sum_{\hat{s}=1}^S \tau_{j\hat{s}\tilde{s}}^{1-\sigma} w_{\hat{s}}^{1-\sigma} \left(B_{\hat{s}} \Psi_{j\hat{s}}^{\frac{\gamma_j(1-\beta)}{1+\kappa_j} - \alpha} \right)^{\frac{(\sigma-1)(1+\kappa_j)}{\gamma_j}}} \right]^{\frac{\mu_j}{1-\sigma}}.$$

A.9 Derivation of model approximation in Sections III.A and III.D

The assumption that the capability distribution is truncated Pareto with lower bound $\psi^{\min} = 1$ and shape parameter k means $G(\psi) = \frac{1-\psi^{-k}}{1-(\psi^{\max})^{-k}}$. Using this functional form in (15) to calculate Ψ_{js} yields:

$$\Psi_{js} = \frac{k}{1 - (\psi^{\max})^{-k}} \left[\frac{(\psi_{js}^*)^{\frac{1}{\gamma_j(1-\beta)-\alpha}-k} - (\psi^{\max})^{\frac{1}{\gamma_j(1-\beta)-\alpha}-k}}{k - \frac{1}{\gamma_j(1-\beta)-\alpha}} + \frac{(\psi_{js}^*)^{\frac{1}{\gamma_j(1-\beta)-\alpha}} - (\psi_{js}^*)^{\frac{1}{\gamma_j(1-\beta)-\alpha}-k}}{k} \right].$$

Letting $\psi^{\max} \rightarrow \infty$ and collecting terms gives equation (24).

Next, differentiate the above expression for Ψ_{js} with respect to B_s to obtain:

$$\frac{\partial \log \Psi_{js}}{\partial \log B_s} = \frac{1}{\Psi_{js}} \frac{1}{1 - (\psi^{\max})^{-k}} \frac{(\psi_{js}^*)^{\frac{1}{\gamma_j(1-\beta)-\alpha}}}{\gamma_j(1-\beta)-\alpha} \left[1 - (\psi_{js}^*)^{-k} \right] \frac{\partial \log \psi_{js}^*}{\partial \log B_s}.$$

From equation (11) we have $\frac{\partial \log \psi_{js}^*}{\partial \log B_s} = -1$. Consequently, letting $\psi^{\max} \rightarrow \infty$ and taking a first order approximation for large ψ_{js}^* implies:

$$\frac{\partial \log \Psi_{js}}{\partial \log B_s} \approx \frac{-1}{\gamma_j(1-\beta)-\alpha}.$$

Substituting this equation into (21) gives $ID_j = \frac{(1-\beta)\kappa_j}{\gamma_j(1-\beta)-\alpha}$ as claimed in the paper.

To obtain the expression for industry-level R&D intensity in equation (29), start by noting that RD_{js} is defined as:

$$RD_{js} = \frac{\int_{\theta} w_s l_{js}^R(\theta) dH_{js}(\theta)}{\int_{\theta} p_{js} y_{js}(\theta) dH_{js}(\theta)}.$$

Using equations (1), (2), (12) and (13) and the functional form for $G(\psi)$ to compute this ratio implies:

$$RD_{js} = \frac{\alpha(\delta + g_j)}{\rho + \zeta + \gamma_j(\delta + g_j)} \frac{\gamma_j(1-\beta) - \alpha}{k[\gamma_j(1-\beta) - \alpha] - 1} \frac{k}{1 - (\psi^{\max})^{-k}} \frac{(\psi_{js}^*)^{\frac{1}{\gamma_j(1-\beta)-\alpha}-k} - (\psi^{\max})^{\frac{1}{\gamma_j(1-\beta)-\alpha}-k}}{\Psi_{js}}.$$

Letting $\psi^{\max} \rightarrow \infty$, using the approximation to Ψ_{js} in (25) and substituting for ψ_{js}^* from (11) then gives equation (29).

B Model Extensions

B.1 Generalization of model

This appendix generalizes the baseline model in three ways. First, it allows for exogenous productivity differences at the country-industry level that are not caused by variation in R&D efficiency. Instead of equation (1), assume the production technology is:

$$y = A_{js}\theta (l^P)^\beta,$$

where A_{js} is a time invariant allocative efficiency term that varies by country and industry.

Second, it assumes that the extent to which a higher quality national innovation system increases R&D efficiency differs across industries. In particular, suppose R&D efficiency varies across industries as well as countries and is given by $B_{js} = B_s^{\nu_{0j}}$ where $\nu_{0j} > 0$ determines the elasticity of B_{js} to country-level R&D efficiency B_s . National innovation systems are more important in industries with higher ν_{0j} .

Third, it relaxes the assumption that the efficiency of technology adoption is constant across firms and countries. Suppose technology adoption is more efficient in countries with higher R&D efficiency and that firms with higher R&D capability also have higher adoption capability. This assumption is consistent with evidence that adoption and innovation draw upon similar capabilities (Rosenberg 1990). Instead of (7), I assume that the adoption technology is given by:

$$\frac{\dot{\theta}}{\theta} = \psi^{\nu_{1j}} B^A B_s^{\nu_{0j}\nu_{2j}} \left(\frac{\theta}{\chi_{js}^A} \right)^{-\gamma_j} (l^A)^\alpha - \delta,$$

where $\nu_{1j}, \nu_{2j} \in [0, 1)$. The parameter ν_{1j} sets the elasticity of a firm's adoption capability to its R&D capability, while ν_{2j} determines the elasticity of adoption efficiency to R&D efficiency. Both elasticities may vary by industry. Imposing $\nu_{1j}, \nu_{2j} < 1$ ensures that, as in the baseline model, the efficiency of R&D relative to adoption is increasing in ψ and B_s .

With these generalizations, the model can be solved using the same series of steps described in Section II. The main differences from the baseline model are as follows. The R&D threshold (11) is now given by:

$$\psi_{js}^* = \eta^{\frac{\gamma_j}{1-\nu_{1j}}} (B^A)^{\frac{1}{1-\nu_{1j}}} B_s^{\frac{-\nu_{0j}(1-\nu_{2j})}{1-\nu_{1j}}}.$$

Steady state relative productivity and R&D employment are still given by (12) and (13), respectively, except that in both equations p_{js} is multiplied by A_{js} and B_s is replaced by $B_{js} = B_s^{\nu_{0j}}$.

The adoption investment problem of a firm with R&D capability ψ is equivalent to the R&D investment problem of a firm with capability $\psi^{\nu_{1j}} (\psi_{js}^*)^{1-\nu_{1j}}$. Therefore, the steady state relative

productivity and adoption employment of a firm with capability $\psi < \psi_{js}^*$ equal the corresponding values for a hypothetical firm with capability $\psi^{\nu_{1j}} (\psi_{js}^*)^{1-\nu_{1j}}$ that chooses to invest in R&D. In addition, the average effective capability in industry j and country s is:

$$\Psi_{js} \equiv \int_{\psi_{js}^*}^{\psi^{\max}} \psi^{\frac{1}{\gamma_j(1-\beta)-\alpha}} dG(\psi) + (\psi_{js}^*)^{\frac{1-\nu_{1j}}{\gamma_j(1-\beta)-\alpha}} \int_{\psi^{\min}}^{\psi_{js}^*} \psi^{\frac{\nu_{1j}}{\gamma_j(1-\beta)-\alpha}} dG(\psi).$$

Given the above modifications to the definitions of ψ_{js}^* and Ψ_{js} , the general equilibrium equations (16)-(18) are unchanged, other than that the definition of Z_{js} becomes:

$$Z_{js} \equiv \frac{\sum_{\hat{s}=1}^S \tau_{j\hat{s}\bar{s}}^{1-\sigma} (\rho a_{\bar{s}} + w_{\bar{s}}) L_{\bar{s}} w_{\bar{s}}^{-\sigma} A_{j\hat{s}}^{\sigma-1} \left(B_s^{\nu_{0j}} \Psi_{j\hat{s}}^{\frac{\gamma_j(1-\beta)}{1+\kappa_j}-\alpha} \right)^{\frac{(\sigma-1)(1+\kappa_j)}{\gamma_j}}}{\sum_{\hat{s}=1}^S \tau_{j\hat{s}\bar{s}}^{1-\sigma} w_{\hat{s}}^{1-\sigma} A_{j\hat{s}}^{\sigma-1} \left(B_{\hat{s}}^{\nu_{0j}} \Psi_{j\hat{s}}^{\frac{\gamma_j(1-\beta)}{1+\kappa_j}-\alpha} \right)^{\frac{(\sigma-1)(1+\kappa_j)}{\gamma_j}}}.$$

Crucially, relative average steady state firm productivity levels are still given by (20) with B_s replaced by B_{js} , implying international technology gaps due to R&D efficiency are independent of A_{js} . However, allocative efficiency does affect income levels (through Z_{js}) and comparative advantage. In particular, the bilateral exports equation (23) is replaced by:

$$\log EX_{j\bar{s}} = v_{j\bar{s}}^4 + (\sigma - 1) \left(\frac{1 + \kappa_j}{\gamma_j} \nu_{0j} \log B_s + \frac{\gamma_j(1 - \beta) - \alpha(1 + \kappa_j)}{\gamma_j} \log \Psi_{js} + \log A_{js} - \log w_s - \log \tau_{j\bar{s}} \right),$$

where $v_{j\bar{s}}^4$ is an importer-industry specific term.

It follows from these observations that all the main theoretical results in the baseline model continue to hold, including Propositions 1, 2 and 3. However, in contrast to the baseline model, trade and income levels are affected by allocative efficiency differences, while industry-level variation in R&D efficiency depends upon ν_{0j} and the parameters ν_{1j} and ν_{2j} affect the equilibrium through ψ_{js}^* and Ψ_{js} .

Taking a first order approximation to Ψ_{js} for large ψ_{js}^* implies that in the generalized model:

$$(70) \quad \Psi_{js} \approx \frac{k [\gamma_j(1 - \beta) - \alpha]}{k [\gamma_j(1 - \beta) - \alpha] - \nu_{1j}} \left[\eta^{\gamma_j} B^A B_s^{-\nu_{0j}(1-\nu_{2j})} \right]^{\frac{1}{\gamma_j(1-\beta)-\alpha}},$$

and using this approximation to calculate innovation-dependence yields:

$$(71) \quad ID_j = \nu_{0j} \left[\frac{\kappa_j(1-\beta)}{\gamma_j(1-\beta) - \alpha} + \nu_{2j} \frac{\gamma_j(1-\beta) - \alpha(1+\kappa_j)}{\gamma_j[\gamma_j(1-\beta) - \alpha]} \right].$$

As in the baseline model, innovation-dependence is increasing in the localization of knowledge spillovers κ_j and decreasing in the advantage of backwardness γ_j . In addition, it is now increasing in both ν_{0j} and ν_{2j} . A higher ν_{0j} raises innovation-dependence by making the returns to R&D more sensitive to B_s , while an increase in ν_{2j} implies B_s has a stronger effect on adoption efficiency. However, innovation-dependence is independent of ν_{1j} . Variation in ν_{1j} affects both selection into R&D and firms' adoption capabilities. In the approximated model, these extensive and intensive margin effects exactly cancel, meaning that the elasticity of average effective capability Ψ_{js} to B_s does not depend upon ν_{1j} .

Next, equations (70) and (71) can be used to obtain generalized versions of the key equations needed to calibrate the model and undertake counterfactual analysis. First, Z_{js} can be written as:

$$Z_{js} = \sum_{\bar{s}=1}^S \frac{\tau_{j\bar{s}\bar{s}}^{1-\sigma} (\rho a_{\bar{s}} + w_{\bar{s}}) L_{\bar{s}} w_{\bar{s}}^{-\sigma} A_{j\bar{s}}^{\sigma-1} B_{\bar{s}}^{(\sigma-1)ID_j}}{\sum_{\hat{s}=1}^S \tau_{j\hat{s}\hat{s}}^{1-\sigma} w_{\hat{s}}^{1-\sigma} A_{j\hat{s}}^{\sigma-1} B_{\hat{s}}^{(\sigma-1)ID_j}}.$$

Except for the inclusion of the allocative efficiency terms, this equation is identical to the corresponding expression in the baseline model (equation 27). It follows that, conditional on knowing B_s and ID_j , wage and income differences due to variation in R&D efficiency can be calculated using (28) exactly as in the baseline model. In particular, it is not necessary to calibrate ν_{0j} , ν_{1j} or ν_{2j} .

Second, substituting (70) into the trade equation (23) and using (71) gives the bilateral exports equation (31) that is used to estimate innovation-dependence in Section III.D. Consequently, given values for B_s (up to a multiplicative constant), innovation-dependence can be estimated exactly as in the baseline model.

Finally, note that industry-level R&D intensity satisfies:

$$(72) \quad RD_{js} = \frac{\alpha(\delta + g_j)}{\rho + \zeta + \gamma_j(\delta + g_j)} \frac{k[\gamma_j(1-\beta) - \alpha] - \nu_{1j}}{k[\gamma_j(1-\beta) - \alpha] - 1} \left[\eta^{\frac{\gamma_j}{1-\nu_{1j}}} (B^A)^{\frac{1}{1-\nu_{1j}}} B_s^{\frac{-\nu_{0j}(1-\nu_{2j})}{1-\nu_{1j}}} \right]^{\frac{\nu_{1j}}{\gamma_j(1-\beta) - \alpha} - k}.$$

Unlike in the baseline model, the elasticity of R&D intensity to R&D efficiency B_s differs across industries. Appendix D.4 explains how this expression can be used to calibrate R&D efficiency in the generalized model.

B.2 Inter-industry spillovers

Suppose the economy is unchanged from the baseline model except that the R&D knowledge level satisfies equation (36). It is straightforward to check that the balanced growth path solution to the baseline model is unaffected, except that equation (65) is replaced by:

$$\left(\frac{p_{js}}{P_{j\tilde{s}}}\right)^{1-\sigma} = \frac{w_s^{1-\sigma} \left(B_s \Psi_{js}^{\gamma_j(1-\beta)-\alpha}\right)^{\frac{\sigma-1}{\gamma_j}} \prod_{i=1}^J \left(B_s \Psi_{is}^{-\alpha}\right)^{\frac{(\sigma-1)\kappa_j d_{ij}}{\gamma_i}}}{\sum_{\tilde{s}=1}^S \tau_{j\tilde{s}\tilde{s}}^{1-\sigma} w_{\tilde{s}}^{1-\sigma} \left(B_{\tilde{s}} \Psi_{j\tilde{s}}^{\gamma_j(1-\beta)-\alpha}\right)^{\frac{\sigma-1}{\gamma_j}} \prod_{i=1}^J \left(B_{\tilde{s}} \Psi_{i\tilde{s}}^{-\alpha}\right)^{\frac{(\sigma-1)\kappa_j d_{ij}}{\gamma_i}}},$$

which implies that equation (19) becomes:

$$Z_{js} \equiv \frac{\sum_{\tilde{s}=1}^S \tau_{j\tilde{s}\tilde{s}}^{1-\sigma} (\rho a_{\tilde{s}} + w_{\tilde{s}}) L_{\tilde{s}} w_{\tilde{s}}^{-\sigma} \left(B_s \Psi_{js}^{\gamma_j(1-\beta)-\alpha}\right)^{\frac{\sigma-1}{\gamma_j}} \prod_{i=1}^J \left(B_s \Psi_{is}^{-\alpha}\right)^{\frac{(\sigma-1)\kappa_j d_{ij}}{\gamma_i}}}{\sum_{\tilde{s}=1}^S \tau_{j\tilde{s}\tilde{s}}^{1-\sigma} w_{\tilde{s}}^{1-\sigma} \left(B_{\tilde{s}} \Psi_{j\tilde{s}}^{\gamma_j(1-\beta)-\alpha}\right)^{\frac{\sigma-1}{\gamma_j}} \prod_{i=1}^J \left(B_{\tilde{s}} \Psi_{i\tilde{s}}^{-\alpha}\right)^{\frac{(\sigma-1)\kappa_j d_{ij}}{\gamma_i}}}.$$

Using equations (36) and (68) also yields that the technology gap between countries s and \tilde{s} in industry j satisfies:

$$\frac{\bar{\theta}_{js}^*}{\bar{\theta}_{j\tilde{s}}^*} = \left[\frac{B_s}{B_{\tilde{s}}} \left(\frac{\Psi_{js}}{\Psi_{j\tilde{s}}} \right)^{\gamma_j(1-\beta)-\alpha} \right]^{\frac{1}{\gamma_j}} \prod_{i=1}^J \left[\frac{B_s}{B_{\tilde{s}}} \left(\frac{\Psi_{is}}{\Psi_{i\tilde{s}}} \right)^{-\alpha} \right]^{\frac{\kappa_j d_{ij}}{\gamma_i}}.$$

Consequently, innovation-dependence can be defined as:

$$ID_{js} \equiv \frac{\partial \log}{\partial \log B_s} \left[\left(B_s \Psi_{js}^{\gamma_j(1-\beta)-\alpha} \right)^{\frac{1}{\gamma_j}} \prod_{i=1}^J \left(B_s \Psi_{is}^{-\alpha} \right)^{\frac{\kappa_j d_{ij}}{\gamma_i}} \right].$$

Now, taking a first order approximation to Ψ_{js} for large ψ_{js}^* gives equation (25). It follows that in the approximated model innovation-dependence is given by equation (37), Z_{js} can be written as in equation (27), and bilateral exports satisfy (31). This means that the calibration and counterfactual analysis presented in Section III are unaffected by the inclusion of inter-industry domestic spillovers.

An alternative approach to incorporating inter-industry spillovers in the model is to assume that inter-industry spillovers affect global knowledge capital χ_j . Suppose, for example, that growth in χ_j is given by:

$$\frac{\dot{\chi}_j}{\chi_j} = \sum_{i=1}^J \tilde{d}_{ij} \sum_{s=1}^S M_{is} \int_{\psi_{\min}}^{\psi_{\max}} \lambda_{is}(\psi) l_{is}^R(\psi) dG(\psi), \quad \text{with } \sum_{i=1}^J \tilde{d}_{ij} = 1.$$

This expression generalizes equation (5) by allowing R&D investment in any industry to contribute

to the growth of global knowledge capital in all other industries. The parameter \tilde{d}_{ij} determines the strength of spillovers from industry i to industry j .

With this modification to the model, the balanced growth path equilibrium conditions are unchanged except that, instead of equation (18), productivity growth satisfies:

$$g_j = \sum_{i=1}^J \tilde{d}_{ij} \sum_{s=1}^S \mu_i \frac{\alpha (\delta + g_i)}{\rho + \zeta + \gamma_i (\delta + g_i)} \frac{Z_{is}}{\Psi_{is}} \int_{\psi_{is}^*}^{\psi^{\max}} \lambda_{is}(\psi) \psi^{\frac{1}{\gamma_i(1-\beta)} - \alpha} dG(\psi).$$

Since the calibration and counterfactual analysis do not use this equation, it immediately follows that allowing inter-industry spillovers to affect global knowledge capital does not affect any of the quantitative results in this paper.

C Data

R&D: R&D intensity is calculated as the industry-level ratio of business R&D expenditure in the OECD's ANBERD database to current price value-added in the OECD's STAN database for 2 digit ISIC Revision 4 manufacturing industries (OECD 2018a,b). To reduce the number of missing observations, I merge industries 10 (Food), 11 (Beverages) and 12 (Tobacco) into a combined industry labelled 1012 and industries 31 (Furniture), 32 (Other manufacturing) and 33 (Repair and installation of machinery and equipment) into a combined industry labelled 3133. This leaves 20 industries in the sample.

I use R&D data from 2010-14 for country-year pairs where R&D intensity is observed for at least two-thirds of industries. The sample includes 25 OECD countries: Australia, Austria, Belgium, Canada, Chile, Czech Republic, Denmark, Finland, France, Germany, Hungary, Ireland, Italy, Japan, Korea, Mexico, Netherlands, Norway, Poland, Portugal, Slovenia, Spain, Turkey, UK and USA. R&D data for Belgium, France and the UK is allocated across industries based on product field, whereas firms' main activity is used for all other countries. Median log R&D intensity b_s^R is computed over all sample industries and years with available data. Because US R&D intensity is missing for a small number of industry-year pairs, I first compute each country's median log R&D intensity relative to Germany, which has no missing data, and then normalize $B_{US} = 1$.

Patents: Counts of triadic patent families by inventor's country and priority date for 2010-14 are from the OECD's Patents by technology database (OECD 2020). The data is for International Patent Classification 4 digit classes and is converted to the 20 ISIC 2 digit manufacturing industries in the R&D intensity sample using the probability based mapping from Lybbert and Zolas (2014). Since industry-level count data for triadic patent families can be volatile from year-to-year, I use average patents and average value-added per year during the sample period to compute patenting

intensity. Industry value-added at current national prices is taken from the OECD's STAN database (OECD 2018b) and converted to US dollars using exchange rates from the IMF's International Financial Statistics (IMF 2018). Median log patenting intensity b_s^P is computed over all sample industries with available data.

International patent applications filed under the Patent Cooperation Treaty in 2014 are from the World Intellectual Property Organization (WIPO 2016). Share of US and Japan in world GDP at market exchange rates in 2014 calculated from the World Bank's World Development Indicators (World Bank 2022).

Trade, output and value-added: Bilateral trade for 2 digit ISIC Revision 4 goods industries is from the OECD's STAN Bilateral Trade by Industry and End-use database (OECD 2018c). Sales of domestic production to the domestic market EX_{jss} is calculated as the difference between output and the sum of exports to all destinations. Output at current national prices is taken from the STAN Database for Structural Analysis (OECD 2018b) and converted to US dollars using exchange rates from the IMF's International Financial Statistics (IMF 2018).

The trade sample comprises imports of the 25 countries where R&D efficiency is observed from all 117 partner countries that have a population greater than 1 million in 2010 and for which nominal wages per efficiency unit of labor employed can be calculated using the Penn World Tables 9.0 (Feenstra, Inklaar and Timmer 2015). The data covers 22 industries: the 20 manufacturing industries included in the R&D intensity sample, Agriculture, forestry and fishing (labelled 0103), and; Mining and quarrying (labelled 0508).

Gravity variables are from the CEPII gravity dataset (Head and Mayer 2014). Distance is population weighted. The Common language dummy denotes country-pairs that share a common official or primary language. The Free trade agreement dummy denotes country-pairs that have notified a regional trade agreement to the World Trade Organization.

Industry growth rates are estimated using OECD STAN data on value-added volumes per person engaged from 1995-2014 (OECD 2018b). The sample comprises the 27 OECD countries that report data for at least half the sample years in at least half the sample industries. Each industry's growth rate is estimated as the time trend from a regression of log value-added volume per person engaged on a trend and country fixed effects.

Country-level variables: GDP, population, nominal wages, physical capital per employee and human capital are from the Penn World Table 9.0 (Feenstra, Inklaar and Timmer 2015). Nominal wages are calculated as labor's share of GDP times output-side GDP at current purchasing power parties (PPPs) times the price level of current GDP divided by persons engaged. The wage variable used to estimate innovation-dependence is the nominal wage per efficiency unit of labor employed, which is calculated as the nominal wage divided by human capital. Physical capital per employee is given by the capital stock at current PPPs divided by persons engaged.

Working age population is measured as the population aged 15-64 from the World Bank's World Development Indicators (World Bank 2021). GDP per capita is defined as GDP per member of the working age population, where GDP is output-side real GDP at chained PPPs from the Penn World Table.

The Worldwide Governance Indicators are from the World Bank (World Bank 2018a). Financial development, measured as private credit by deposit money banks and other financial institutions as a share of GDP is from the World Bank's Financial Structure Database (Čihák et al. 2012). Data for Canada is unavailable after 2008, so I extrapolate by holding Canadian financial development constant at its 2008 value. Business environment is measured by a country's global distance to the frontier for Ease of doing business from the World Bank's Doing Business data set (World Bank 2018b). All these variables are time-varying.

UK firm-level R&D: The share of firms that perform R&D $ShRD_{js}$ and the share of value-added produced by firms that perform R&D $ShVA_{js}$ are computed from the UK's Annual Business Survey, which is a representative sample of production, construction, distribution and service industries (ONS 2021). The Annual Business Survey data is reported for UK SIC 2007 industries, which corresponds to ISIC Revision 4. The data does not cover Northern Ireland.

Firms are asked whether they have "plans to carry out in-house Research and Development during the next two years". I identify firms that answer yes to this question as R&D firms and drop non-respondents from the calculations. Value-added is measured as approximate gross value-added at basic prices. The R&D and value-added shares are computed for each 2 digit goods industry using sampling weights and I measure the average shares for 2008-09. For the Coke and refined petroleum products industry (19), the data implies that R&D firms are, on average, smaller than other firms, so I set $ShVA_{js} = ShRD_{js}$.

To measure R&D intensity, I match the Annual Business Survey with the Business Enterprise Research and Development data set (ONS 2017) and compute the R&D intensity of each firm that performs R&D as the ratio of total R&D expenditure to approximate gross value-added at basic prices. R&D intensity $FiRD_j$ is then calculated as the median of all firm-level observations pooled for 2008-09 for each 2 digit goods industry and for the services sector. Due to sample size restrictions on data disclosure, R&D intensity for the Agriculture, forestry and fishing industry (0103) and the Coke and refined petroleum products industry (19) are calculated using 2008-13 data.

Additional calibration parameters: Expenditure shares are calculated as the industry's share of domestic absorption, where domestic absorption is defined as output plus imports minus exports. Output at current national prices is taken from the OECD's STAN Database for Structural Analysis (OECD 2018b) and converted to US dollars using exchange rates from the IMF's International Financial Statistics (IMF 2018). Imports and exports by industry are from the OECD's STAN

Bilateral Trade by Industry and End-use database (OECD 2018c). The calibrated expenditure shares are averages over all OECD countries for which data is available for all industries in 2012.

The exit rate is the average across OECD countries in 2012 of the death rate of employer enterprises in the business economy excluding holding companies. Data on death rates is from the OECD Structural and Demographic Business Statistics Business Demography Indicators using the ISIC Revision 4 classification (OECD 2018d).

Caliendo and Parro (2015) estimate trade elasticities for ISIC Revision 3 goods sectors at approximately the 2 digit level of aggregation. I take the benchmark estimates from the 99% sample in their Table 1. Caliendo and Parro do not use the estimated elasticities for the Basic metals, Machinery and Auto sectors because these elasticities are not robust across specifications. For these sectors, I set the trade elasticity equal to the estimated aggregate elasticity. Caliendo and Parro's sectors map one-to-one into 2 digit ISIC Revision 4 industries with the following exceptions: I map Textile to the Textiles (13), Wearing apparel (14) and Leather (15) industries; Paper to the Paper (17) and Printing (18) industries; Chemicals to the Chemicals (20) and Pharmaceutical (21) industries, and; for the Computers (26) industry I take the average of the trade elasticities in the Office, Communication and Medical sectors.

Out-of-sample comparative advantage test: R&D intensity is calculated from Eurostat data as the ratio of business expenditure on R&D to value-added at factor costs for 2 digit NACE Revision 2 manufacturing industries, which correspond directly to ISIC Revision 4 industries (Eurostat 2018a,b). As for the baseline sample, I merge industries 10, 11 and 12 and industries 31, 32 and 33, which leaves 20 industries. R&D efficiency is computed as the median log R&D intensity over all sample industries and years from 2008-15, where the sample includes those country-year pairs where R&D intensity is observed for at least half of all industries. These sample selection criteria are weaker than for the baseline OECD sample, which allows for a larger sample. Nine countries meet the criteria: Bulgaria, Croatia, Cyprus, Estonia, Greece, Lithuania, Romania, Slovakia and Sweden. To compute R&D efficiency from patent data for these nine countries, I use the same procedure as for the baseline sample, except that the data covers 2008-15 and industry value-added data is from Eurostat.

All other variables for the out-of-sample test are taken from the same sources used for the baseline estimation, except for industry output, which is from Eurostat. The sample covers bilateral trade in 20 manufacturing industries with 117 partner countries that have a population greater than 1 million in 2010 and for which the nominal wage per efficiency unit of labor employed can be calculated using the Penn World Tables 9.0.

D Calibration

D.1 Patent data calibration

Let $Patents_{js}$ be the number of patents generated by industry j in country s . Suppose $Patents_{js} = \Lambda_j^0 RDX_{js}^\Lambda$ where RDX_{js} denotes R&D expenditure in industry j and country s , Λ_j^0 is an industry-specific constant that captures cross-industry differences in the extent to which innovations can be patented and the benefits of patenting, and Λ is the elasticity of industry patenting to R&D expenditure. Let VA_{js} denote industry value-added. Then patenting intensity $PAT_{js} \equiv \frac{Patents_{js}^{\frac{1}{\Lambda}}}{VA_{js}}$ satisfies:

$$\begin{aligned} PAT_{js} &= (\Lambda_j^0)^{\frac{1}{\Lambda}} \frac{RDX_{js}}{VA_{js}}, \\ &= (\Lambda_j^0)^{\frac{1}{\Lambda}} \frac{\alpha(\delta + g_j)}{\rho + \zeta + \gamma_j(\delta + g_j)} \frac{k [\gamma_j(1 - \beta) - \alpha]}{k [\gamma_j(1 - \beta) - \alpha] - 1} \eta^{-k\gamma_j} \left(\frac{B_s}{B^A} \right)^k, \end{aligned}$$

where the second equality uses equation (29). Comparing the expression above to equation (29) implies:

$$\frac{PAT_{js}}{PAT_{j\bar{s}}} = \frac{RD_{js}}{RD_{j\bar{s}}} = \left(\frac{B_s}{B_{\bar{s}}} \right)^k.$$

It follows that, as an alternative to using R&D intensity data, R&D efficiency can also be calibrated from within industry, cross-country variation in patenting intensity.

The patent data used to calibrate R&D efficiency covers the same period, countries and industries as the R&D data. Since there is home bias in patent applications, I only count triadic patent families that have been filed jointly at the US, Japanese and European patent offices. Patenting intensity is calculated assuming the elasticity of patenting to R&D Λ equals one. A unit elasticity is consistent with the firm-level estimates of Lewbel (1997) and the conclusions of Griliches (1990). In the robustness checks detailed in Appendix D.2, I allow for an elasticity below one.

D.2 Robustness checks in Section III.F

Table A1 reports a series of robustness checks on the baseline counterfactual results. For each robustness check, I first recalibrate the model and then calculate the counterfactual changes in wages and income per capita relative to the US when differences in R&D efficiency are eliminated. In each calibration, innovation-dependence and trade costs are estimated including the productivity and comparative advantage controls from columns (3) and (4) of Table 1, except in the importer

fixed effects calibrations in columns (3) and (4) of Table A1 where the productivity controls are omitted. In all cases, I set innovation-dependence equal to zero whenever its point estimate is negative.

The first robustness check adds another control when estimating innovation-dependence – the interaction of industry dummy variables with the importer’s log GDP per capita. GDP per capita proxies for omitted variables that affect productivity and comparative advantage and may be correlated with R&D efficiency. However, because it is partly determined by R&D efficiency, it is not included in the baseline specification. Column (1) reports the results when R&D data is used to measure R&D efficiency, while patent data is used in column (2). The difference from the baseline results is negligible.

Second, I estimate innovation-dependence including importer fixed effects in equation (33) and dropping the productivity controls, which only vary by importer. This specification estimates innovation-dependence up to an additive constant. Consequently, I normalize the innovation-dependence estimates by setting the innovation-dependence of the Coke and refined petroleum products industry equal to zero. This normalization is conservative compared to the positive innovation-dependence estimates for the Coke industry obtained in Table 1. When importer fixed effects are included, the Coke industry has the second lowest innovation-dependence estimate for both the R&D and patent data calibrations (ahead of only Mining and quarrying). For the R&D data calibration, including importer fixed effects slightly reduces counterfactual wage and income changes. Column (3) reports that the wage dispersion ratio equals 0.27 and the income dispersion ratio is 0.12. The results for the patent data calibration in column (4) are also lower than in the baseline, though the differences are small.

Next, I repeat the baseline R&D and patent data calibrations, except that I set innovation-dependence to zero in all industries where estimated innovation-dependence is insignificant at the 10 percent level. This change reduces wage and income variation caused by differences in R&D efficiency, but as columns (5) and (6) show the counterfactual results differ little from the baseline results in Table 3.

The baseline patent data calibration in column (2) of Table 3 calculates patenting intensity PAT_{js} under the assumption that the elasticity of patenting to R&D expenditure $\Lambda = 1$. Griliches (1990) concludes that the firm-level patenting elasticity is probably close to unity, but also acknowledges that estimates below one are common in the literature. Therefore, in column (7) I calibrate the model assuming $\Lambda = 0.5$. Reducing Λ increases the variation in implied R&D efficiency given observed differences in patenting and value-added, which in turn compresses the innovation-dependence estimates obtained from equation (33). Together these effects lead to small declines in the wage and income dispersion ratios.

Column (8) reports an upper bound on the effect of eliminating R&D efficiency differences for

the R&D data calibration when the model is solved without taking a first order approximation. See Appendix D.3 below for details.

Columns (9)-(16) study the impact of calibrating the model using alternative values of the trade elasticity $\sigma - 1$, which in the baseline R&D and patenting calibrations equals 6.53. Columns (9) and (10) reduce the trade elasticity to 2.5 for the R&D intensity and patenting intensity calibrations, respectively. Columns (11) and (12) use an elasticity of 4.5, which is close to the aggregate elasticity estimated by Caliendo and Parro (2015). Columns (13) and (14) increases the elasticity to 8.5. The results show that increasing the trade elasticity reduces the magnitude of counterfactual wage and income changes because it leads to lower innovation-dependence estimates. The patent data calibration is more sensitive to changes in the trade elasticity than the R&D data calibration for which differences from the baseline results are not large. Finally, columns (15) and (16) use the industry-specific trade elasticities estimated by Caliendo and Parro (2015). Again, the results are similar to the baseline.

D.3 Model approximation

This appendix describes how to calculate an upper bound on the approximation error that results from using a first order approximation to Ψ_{js} in the counterfactual analysis. Comparing equations (24) and (25) shows that the approximation drops the term E_{js} given by:

$$(73) \quad E_{js} = 1 + \frac{(\psi_{js}^*)^{-k}}{k [\gamma_j(1 - \beta) - \alpha] - 1}.$$

Since ψ_{js}^* is decreasing in R&D efficiency B_s , this expression implies E_{js} is increasing in B_s . Not taking the approximation to Ψ_{js} leaves the equations used to solve the calibrated model unchanged (see equation 28), except that Z_{js} in equation (27) is replaced by:

$$(74) \quad Z_{js} = \frac{\sum_{\hat{s}=1}^S \tau_{j\hat{s}\hat{s}}^{1-\sigma} (\rho a_{\hat{s}} + w_{\hat{s}}) L_{\hat{s}} w_{\hat{s}}^{-\sigma} B_{\hat{s}}^{(\sigma-1)ID_j} E_{j\hat{s}}^{(\sigma-1) \left[1 - \beta - \frac{\alpha(1+\kappa_j)}{\gamma_j} \right]}}{\sum_{\hat{s}=1}^S \tau_{j\hat{s}\hat{s}}^{1-\sigma} w_{\hat{s}}^{1-\sigma} B_{\hat{s}}^{(\sigma-1)ID_j} E_{j\hat{s}}^{(\sigma-1) \left[1 - \beta - \frac{\alpha(1+\kappa_j)}{\gamma_j} \right]}},$$

where ID_j is still given by equation (26). Because Assumption 1 ensures $\gamma_j(1 - \beta) > \alpha(1 + \kappa_j)$, Z_{js} is increasing in E_{js} . It follows that using the approximation to Ψ_{js} reduces wage inequality caused by differences in R&D efficiency B_s .

The exponent of E_{js} in equation (74) is bounded above by $(\sigma - 1)(1 - \beta)$. Therefore, to obtain an upper bound on the approximation error, I start by setting the exponent equal to this upper bound

and assume that:

$$(75) \quad Z_{js} = \frac{\sum_{\bar{s}=1}^S \tau_{j\bar{s}\bar{s}}^{1-\sigma} (\rho a_{\bar{s}} + w_{\bar{s}}) L_{\bar{s}} w_{\bar{s}}^{-\sigma} B_{\bar{s}}^{(\sigma-1)ID_j} E_{j\bar{s}}^{(\sigma-1)(1-\beta)}}{\sum_{\hat{s}=1}^S \tau_{j\hat{s}\hat{s}}^{1-\sigma} w_{\hat{s}}^{1-\sigma} B_{\hat{s}}^{(\sigma-1)ID_j} E_{j\hat{s}}^{(\sigma-1)(1-\beta)}}.$$

The next step is to calibrate E_{js} . Let $ShVA_{js}$ denote the share of industry value-added produced by firms that perform R&D and note that:

$$E_{js} - 1 = \frac{(\eta^{\gamma_j} B^A)^{-k}}{k [\gamma_j(1-\beta) - \alpha] - 1} B_s^k = \frac{ShVA_{js} - ShRD_{js}}{1 - ShVA_{js}},$$

where the first equality is obtained by substituting equation (11) into equation (73), and the second equality uses $ShRD_{js} = \eta^{-k\gamma_j} (B_s/B^A)^k$, equation (29) and $RD_{js} = \frac{\alpha(\delta+g_j)}{\rho+\zeta+\gamma_j(\delta+g_j)} ShVA_{js}$. Using UK data to measure $ShVA_{js}$ and $ShRD_{js}$ allows me to calibrate E_{js} by industry in the UK. The calibrated values of B_s^k relative to the US can then be used to infer E_{js} in all other sample countries.

Finally, I calculate the counterfactual effect of eliminating R&D efficiency differences (i.e. setting both B_s and E_{js} equal across countries) when Z_{js} satisfies equation (75). To quantify the approximation error for a given calibration of R&D efficiency and innovation-dependence levels, the counterfactual analysis uses the calibrated parameters from the baseline R&D data calibration. When solving for real income per capita, I continue to assume that non-tradable prices are not directly affected by R&D efficiency.

The counterfactual results are shown in column (8) of Table A1. As noted above, R&D efficiency accounts for a larger share of international wage and income inequality when including variation in E_{js} , but the difference is small. On average, wages relative to the US increase by 20 log points compared to 18 log points in the baseline calibration, and the wage dispersion ratio is 0.36 compared to 0.32 in the baseline. Real income per capita relative to the US increases by 6.6 log points on average compared to 5.9 log points in the baseline, and the income dispersion ratio is 0.19 compared to 0.17 in the baseline. These comparisons show that the E_{js} term, which is dropped when taking a first order approximation to Ψ_{js} , is not quantitatively important for the counterfactual outcomes studied in the paper.

D.4 Calibration of R&D efficiency in generalized model from Section IV.A

The objective is to calibrate R&D efficiency in the generalized model. Using equation (72) and taking the ratio of RD_{js} across countries gives:

$$\frac{RD_{js}}{RD_{j\tilde{s}}} = \left(\frac{B_s}{B_{\tilde{s}}} \right)^{\frac{k[\gamma_j(1-\beta)-\alpha]^{-\nu_{1j}} \nu_{0j}(1-\nu_{2j})}{\gamma_j(1-\beta)-\alpha} \frac{1-\nu_{1j}}{1-\nu_{1j}}},$$

which shows that, unlike in the baseline model, the relative R&D intensity of different countries varies by industry. However, for any set of countries s , \hat{s} and \tilde{s} , the equation above implies:

$$\frac{\log\left(\frac{RD_{js}}{RD_{j\tilde{s}}}\right)}{\log\left(\frac{RD_{j\hat{s}}}{RD_{j\tilde{s}}}\right)} = \frac{\log\left(\frac{B_s}{B_{\tilde{s}}}\right)}{\log\left(\frac{B_{\hat{s}}}{B_{\tilde{s}}}\right)}.$$

After normalizing $B_{US} = 1$, this expression can be used to calibrate the ratio of log R&D efficiencies for any pair of sample countries. Arbitrarily fixing the value of $\log B_s$ in any one country then pins down log R&D efficiency for each country up to an unknown multiplicative constant, which is sufficient information to implement the quantitative analysis.

Let b_s^G denote the calibrated log R&D efficiency of country s in the generalized model. Formally, I compute b_s^G as:

$$b_s^G = \text{Median}_{\hat{s} \neq \tilde{s}} \left\{ K_{\hat{s}} \text{Median}_{j,t} \left[\frac{\log\left(\frac{RD_{jst}}{RD_{j\tilde{s}t}}\right)}{\log\left(\frac{RD_{j\hat{s}t}}{RD_{j\tilde{s}t}}\right)} \right] \right\}.$$

To understand this expression, start by noting that taking the median across industries and years of the term inside square brackets gives an estimate of log R&D efficiency of country s relative to country \hat{s} under the assumption that $b_{\hat{s}}^G = 0$. Multiplying this estimate by $K_{\hat{s}}$ then fixes R&D efficiency in one country. In particular, I choose $K_{\hat{s}}$ such that the difference between the log R&D efficiencies of Germany and the Czech Republic is the same as in the baseline R&D calibration in Section III.D. Finally, to obtain b_s^G , I take the median across all possible comparison countries \hat{s} .

Since US R&D intensity data is missing for a small number of industry-year pairs, I compute b_s^G with Germany as country \tilde{s} and then normalize $b_{US}^G = 0$. The medians are calculated over all sample industries and years from 2010-14 with available data and over all countries in the baseline sample.

References

- Acemoglu, Daron. 2009. "Introduction to Modern Economic Growth." Princeton University Press.
- Allen, Treb, Costas Arkolakis, and Xiangliang Li. 2015. "On the Existence and Uniqueness of Trade Equilibria." mimeo, Yale University.
- Caliendo, Lorenzo, and Fernando Parro. 2015. "Estimates of the Trade and Welfare Effects of

- NAFTA.” *Review of Economic Studies* 82 (1): 1-44.
- Čihák, Martin, Asli Demirgüç-Kunt, Erik Feyen, and Ross Levine. 2012. “Benchmarking Financial Development around the World.” World Bank Policy Research Working Paper 6175.
- Eurostat. 2018a. “BERD by NACE Rev. 2 activity.” (accessed September 10, 2018).
- Eurostat. 2018b. “Annual enterprise statistics for special aggregates of activities (NACE Rev. 2).” (accessed September 10, 2018).
- Feenstra, Robert C., Robert Inklaar, and Marcel P. Timmer. 2015. “The Next Generation of the Penn World Table.” *American Economic Review* 105 (10): 3150-3182, available for download at www.ggd.net/pwt.
- Griliches, Zvi. 1990. “Patent Statistics as Economic Indicators: A Survey Part I.” NBER Working Paper 3301.
- Head Keith, and Thierry Mayer. 2014. “Gravity Equations: Toolkit, Cookbook, Workhorse.” *Handbook of International Economics*, Vol. 4, eds. Gopinath, Helpman and Rogoff, Elsevier.
- International Monetary Fund (IMF). 2018. “International Financial Statistics.” (accessed June 13, 2018).
- Lewbel, Arthur. 1997. “Constructing Instruments for Regressions with Measurement Error When No Additional Data are Available, with an Application to Patents and R&D.” *Econometrica* 65 (5): 1201-1213.
- Lybbert, Travis, and Nikolas Zolas. 2014. “Getting Patents and Economic Data to Speak to Each Other: An ‘Algorithmic Links with Probabilities’ Approach for Joint Analyses of Patenting and Economic Activity.” *Research Policy* 43 (3): 530-42.
- Office for National Statistics (ONS). 2017. “Business Expenditure on Research and Development, 1994-2015.” 6th Edition, UK Data Service SN: 6690, <http://doi.org/10.5255/UKDA-SN-6690-6>.
- Office for National Statistics (ONS). 2021. “Annual Business Survey, 2005-19: Secure Access. [data collection].” 15th Edition, UK Data Service SN: 7451, <http://doi.org/10.5255/UKDA-SN-7451-15>.
- Organisation for Economic Co-operation and Development (OECD). 2018a. “ANBERD (Analytical Business Enterprise R&D) database.” (accessed February 8, 2018).
- Organisation for Economic Co-operation and Development (OECD). 2018b. “STAN Industrial Analysis 2016 ed.” (accessed May 11, 2018 and June 13, 2018).
- Organisation for Economic Co-operation and Development (OECD). 2018c. “Bilateral Trade in Goods by Industry and End-use.” (accessed June 12, 2018, August 24, 2018 and September 11, 2018).
- Organisation for Economic Co-operation and Development (OECD). 2018d. “SDBS Business Demography Indicators (ISIC Rev. 4).” (accessed August 24, 2018).

- Organisation for Economic Co-operation and Development (OECD). 2020. “Patents by technology.” (accessed April 2, 2020).
- Rosenberg, Nathan. 1990. “Science and Technology Policy for the Asian NICs: Lessons from Economic History.” in *Science and Technology. Lessons for Development Policy*, Robert E. Evenson and Gustav Ranis, eds. Westview Press.
- World Bank. 2018a. “Worldwide Governance Indicators.” (accessed July 2, 2018).
- World Bank. 2018b. “Doing Business.” (accessed July 2, 2018).
- World Bank. 2021. “World Development Indicators.” Series: SP.POP.1564.TOT, (accessed November 1, 2021).
- World Bank. 2022. “World Development Indicators.” Series: NY.GDP.MKTP.CD, (accessed August 24, 2022).
- World Intellectual Property Organization (WIPO). 2016. “U.S. Extends Lead in International Patent and Trademark Filings.” PR/2016/788, https://www.wipo.int/pressroom/en/articles/2016/article_0002.html, (accessed September 27, 2016).

Table A1: Counterfactual Robustness Checks

Robustness check		GDP per capita		Importer fixed effects		Significant innovation-dependence estimates		Patenting elasticity = 0.5	Approximation error
		R&D intensity	Patenting intensity	R&D intensity	Patenting intensity	R&D intensity	Patenting intensity	Patenting intensity	R&D intensity
		(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
(i) Nominal wage	Average change relative to US	0.18	0.13	0.15	0.13	0.17	0.13	0.23	0.20
	Dispersion ratio	0.32	0.25	0.27	0.26	0.31	0.25	0.25	0.36
(ii) Real income per capita	Average change relative to US	0.057	0.035	0.043	0.038	0.058	0.036	0.069	0.066
	Dispersion ratio	0.17	0.11	0.12	0.11	0.17	0.11	0.12	0.19

Robustness check		Trade elasticity = 2.5		Trade elasticity = 4.5		Trade elasticity = 8.5		Industry-specific trade elasticities	
		R&D intensity	Patenting intensity	R&D intensity	Patenting intensity	R&D intensity	Patenting intensity	R&D intensity	Patenting intensity
		(9)	(10)	(11)	(12)	(13)	(14)	(15)	(16)
(i) Nominal wage	Average change relative to US	0.23	0.22	0.19	0.16	0.17	0.13	0.19	0.14
	Dispersion ratio	0.38	0.43	0.34	0.31	0.32	0.24	0.36	0.26
(ii) Real income per capita	Average change relative to US	0.062	0.059	0.058	0.046	0.060	0.041	0.083	0.045
	Dispersion ratio	0.18	0.18	0.17	0.14	0.17	0.12	0.24	0.13

Row (i) reports the average counterfactual log wage change relative to the US, and the ratio of the standard deviation of the counterfactual log wage change to the standard deviation of observed log wages. Row (ii) gives the same statistics for real GDP per capita, defined as GDP per member of the working age population. Counterfactual sets R&D efficiency equal across countries. Observed wages and GDP per capita calculated from the Penn World Tables 9.0 and World Development Indicators in 2012. For columns (1) and (2) innovation-dependence is estimated including the interaction of industry dummy variables with the importer's log GDP per member of the working age population as an additional control. For columns (3) and (4) innovation-dependence is estimated including importer fixed effects as controls and the innovation-dependence estimate for Coke and refined petroleum products is normalized to zero. For columns (5) and (6) all innovation-dependence estimates that are insignificant at the 10 percent level are set equal to zero. For column (7) R&D efficiency is calculated from patenting data assuming that the elasticity of patenting to R&D expenditure equals 0.5. Column (8) uses the baseline R&D intensity calibration and reports an upper bound on the effect of eliminating R&D efficiency differences when the model is solved without taking a first order approximation. Industry-specific trade elasticities used in columns (15) and (16) from Caliendo and Parro (2015).