Technology Gaps, Trade and Income*

Thomas Sampson†
London School of Economics

September 2022

Abstract

This paper quantifies the contribution of technology gaps to international income inequality. I develop an endogenous growth model where cross-country differences in R&D efficiency and cross-industry differences in innovation and adoption opportunities together determine equilibrium technology gaps, trade patterns and income inequality. Higher R&D efficiency countries are richer and have comparative advantage in more innovation-dependent industries. I calibrate R&D efficiency by country and innovation-dependence by industry using R&D, patent and bilateral trade data. Counterfactual analysis implies technology gaps account for one-quarter to one-third of nominal wage variation within the OECD.

Keywords: Technology gaps, Development accounting, Comparative advantage, Innovation, Technology diffusion, Endogenous growth.

*I am grateful to Pol Antràs, Arnaud Costinot, Ha Nguyen, Natalia Ramondo, Andrés Rodríguez-Clare, Pau Roldan-Blanco and Christopher Tonetti, as well as many seminar and conference participants, for helpful comments and suggestions. Fergal Hanks provided excellent research assistance. An early version of this paper circulated under the title "The Global Productivity Distribution and Ricardian Comparative Advantage".

†Centre for Economic Performance, Department of Economics, London School of Economics, Houghton Street, London, WC2A 2AE, United Kingdom. E-mail: t.a.sampson@lse.ac.uk.

Most innovation takes place in a small number of rich, industrialized economies.[1] And since technology diffusion is not instantaneous, more innovative firms and countries use better technologies. This paper studies the technology gaps that arise from innovation and diffusion (Parente and Prescott 1994; Buera and Oberfield 2020). What determines the size of international technology gaps? How do technology gaps differ across industries? And how important are technology gaps in explaining cross-country variation in wages and incomes?

The paper makes two main contributions. First, it develops a theory of equilibrium international technology gaps when productivity levels are determined by the innovation and adoption investments of heterogeneous firms. The theory assumes that firms behind the technology frontier benefit from an advantage of backwardness (Gerschenkron 1962) and that knowledge spillovers are stronger within than across countries (Keller 2002). I use the model to isolate the mechanisms through which country-level differences in the efficiency of innovation affect technology gaps, trade flows and incomes. The model is both analytically tractable and sufficiently rich to be used for quantitative analysis.

The paper's second contribution is to calibrate the model and quantify how international variation in innovativeness affects trade and incomes. The calibration strategy exploits the model's prediction that cross-country differences in innovation efficiency generate Ricardian comparative advantage due to sectoral heterogeneity in the innovation and diffusion technologies. Consequently, trade data can be used to infer sufficient statistics that capture how innovation efficiency affects technology gaps in each industry. The quantification provides a novel way to evaluate international variation in living standards due to technology gaps. By contrast, previous development accounting research identifies productivity differences with the Solow residual after accounting for factor endowments (Caselli 2005), or estimates the effect of misallocation on the efficiency with which given technologies and factors are used (Acemoglu and Zilibotti 2001; Hsieh and Klenow 2009; Hsieh et al. 2019).

The components of the model are introduced in Section 1. First, the *efficiency of R&D* varies across countries due to differences in national innovation systems (Nelson 1993). Countries with better national innovation systems have an absolute advantage in R&D. Second, firms choose whether to upgrade their productivity through innovative R&D or through technology adoption (Benhabib, Perla and Tonetti 2014; König, Lorenz and Zilibotti 2016). Firms are heterogeneous in their R&D capabilities and, in equilibrium, there exists a capability threshold above which firms select into R&D. In countries with higher R&D efficiency the threshold is lower, which implies the share of firms that innovate is greater. Allowing for firm-level selection between R&D and adoption is a key distinction between the theory and existing quantitative models of trade and

[1] For example, the US and Japan accounted for 48% of applications filed under the World Intellectual Property Organization's Patent Cooperation Treaty in 2014, while producing 28% of world GDP.

productivity.

Third, there are knowledge spillovers within and across countries. Knowledge is used as an input to both R&D and technology adoption and the knowledge level in each country is an average of the domestic productivity frontier and global knowledge capital. The weight given to the domestic frontier determines the *localization of knowledge spillovers*. There is also an *advantage of backwardness* that increases the efficiency of technology investment for less productive firms, regardless of whether they choose innovation or adoption (Gerschenkron 1962; Griffith, Redding and Van Reenen 2004). I allow both the localization of knowledge spillovers and the advantage of backwardness to be industry-specific.[2]

The presence of international knowledge spillovers and an advantage of backwardness ensures that on a balanced growth path technology gaps (i.e. relative productivity levels) are stable, both between domestic firms and across countries. Section 2 characterizes balanced growth in a global economy with many countries and industries and studies equilibrium technology gaps. Countries with higher R&D efficiency are more productive and richer. Likewise, within country-industry pairs, firms that perform R&D are more productive than those that adopt and productivity is increasing in R&D capability among innovative firms.

owever, the size of technology gaps is endogenous and differs by industry depending upon the dispersion and concentration forces. The dispersion force results from the localization of knowledge spillovers. The concentration force comes from global knowledge spillovers and the advantage of backwardness. Within countries, a greater advantage of backwardness strengthens the concentration force and reduces productivity variation. Across countries, not only is this effect present, but the localization of knowledge spillovers also plays a role. More localized spillovers magnify the advantage of firms in more productive countries and widen technology gaps.

In steady state, the strength of the dispersion and concentration forces in each industry can be summarized by a single sufficient statistic: the elasticity of a country's relative average productivity to its R&D efficiency. I call this elasticity the industry's *innovation-dependence*. The theory implies that innovation-dependence is decreasing in the advantage of backwardness and increasing in the localization of knowledge spillovers. When innovation-dependence is low, the gap between leaders and followers is small, whereas high innovation-dependence increases the productivity advantage that accrues to innovators. Consequently, international technology gaps are greater in more innovation-dependent industries. It follows that countries with higher R&D efficiency have Ricardian comparative advantage in industries with lower advantage of backwardness and more localized knowledge spillovers.[3]

[2]Peri (2005) shows that the impact of international borders on knowledge flows varies by sector. Doraszelski and Jaumandreu (2013) find that the effect of current productivity on future productivity growth, conditional on R&D investment, differs across industries.

[3]This prediction provides an endogenous growth formalization of Krugman's (1985) argument that comparative

While cross-industry variation in innovation-dependence determines the pattern of comparative advantage, the level of innovation-dependence determines the cross-country wage and income inequality. In a single sector economy with free trade, the elasticity of a country's relative wage to its R&D efficiency is proportional to innovation-dependence. With trade costs and many sectors, the relationship is more complex, but the mechanism is the same: when industries are more innovation-dependent, countries with higher R&D efficiency have a greater technological advantage and this leads to larger differences in wages and income per capita. The model formalizes this intuition and unpacks the determinants of innovation-dependence.

To quantify the importance of technology gaps, Section 3 calibrates a first-order approximation to the model using data on 25 OECD economies. Two key sets of parameters are required: R&D efficiency by country and innovation-dependence by industry. Both are calibrated by matching model-implied moments to their empirical counterparts. Since the share of firms that choose R&D rather than adoption is increasing in R&D efficiency, the industry-level ratio of R&D expenditure to value-added is larger in countries with higher R&D efficiency. Using this moment, I calibrate R&D efficiency from cross-country, within-industry variation in innovation intensity. I obtain two independent measures of R&D efficiency using data on innovation inputs (R&D) and innovation outputs (patents), respectively.

Given R&D efficiency, I estimate innovation-dependence for 22 goods industries using the gravity equation for bilateral trade implied by the model. The innovation-dependence of each industry is estimated to match the observed correlation between R&D efficiency and trade flows. I estimate innovation-dependence separately using the R&D efficiency measure calibrated from R&D data and the measure calibrated from patent data, but the two sets of estimates are similar with a correlation of 0.88. In both cases, innovation-dependence is highest in the Computers, Machinery and equipment, and Chemicals industries and lowest in Mining. An out-of-sample validation test confirms that countries with higher R&D efficiency have a comparative advantage in industries with larger estimated innovation-dependence.

Using the calibrated model, I quantify the impact of R&D efficiency differences by comparing the calibrated equilibrium to a counterfactual economy where R&D efficiency is the same in all countries. The counterfactual analysis shows that technology gaps account for an important share of variation in both nominal wages and real income per capita across OECD countries.

Eliminating R&D efficiency differences increases nominal wages relative to the US by around 20% for the average sample country. Since richer countries tend to be more innovative, equalizing R&D efficiency also reduces cross-country inequality in wages and incomes. The results imply that R&D efficiency differences account for one-quarter to one-third of nominal wage dispersion within

advantage can be characterized in terms of technology gaps when countries are ranked by their technological level and industries by their technological intensity.

the OECD. Under the assumption that innovation-dependence is zero in the services sector, I also find that R&D efficiency accounts for around 15% of variation in real income per capita. Section 4 shows that these conclusions are robust to generalizing how the R&D and adoption technologies differ across countries and industries, and to incorporating inter-industry knowledge spillovers in the model.

As well as generating wage and income differences, technology gaps have large effects on comparative advantage. For example, in the calibration based on R&D data, eliminating R&D efficiency differences increases exports relative to the US for the average country by 94 log points more for the Chemicals industry at the 90th percentile of the innovation-dependence distribution than for the Agriculture industry at the 10th percentile.

By quantifying the impact of differences in innovativeness on living standards, this paper provides new evidence on the sources of international inequality and contributes to a small quantitative literature on technology gaps. Parente and Prescott (1994) calibrate a single sector model with exogenous growth and show that observed income disparities could be explained by plausible cross-country differences in the research technology, which they label barriers to technology adoption. Likewise, Klenow and Rodríguez-Clare (2005) argue that variation in R&D investment may be sufficient to generate observed international productivity gaps. Using a directed technical change model, Gancia, Müller and Zilibotti (2013) estimate the barriers to adoption needed to fit cross-country output differences and find that if all countries used frontier technologies then GDP per worker of the average OECD economy relative to the US would increase from 0.68 to 0.91. Relative to these studies, the paper's contribution is to quantify the impact of technology gaps using innovation and trade data without targeting observed income differences in the calibration.

In related development accounting research, Alviarez, Cravino and Ramondo (2021) use cross-country variation in the market shares of multinational firms to estimate that firm-embedded productivity differences account for one-third of cross-country income dispersion in their sample of mostly European countries. Although R&D efficiency is only one of the possible sources of firm-embedded productivity, the finding that it accounts for around 15% of income dispersion implies that innovativeness is an important determinant of firm-embedded productivity.

In common with this paper, Buera and Oberfield (2020) and Cai, Li and Santacreu (2022) develop dynamic quantitative trade models incorporating knowledge diffusion. I differ from these papers in modelling technology upgrading by incumbent firms, rather than building upon the framework for studying trade and innovation pioneered by Eaton and Kortum (2001, 2002). But a more important distinction is that, whereas their work asks how trade liberalization affects productivity, I study the extent to which technology gaps explain income differences. Eaton and Kortum (1999) also build a model of innovation and diffusion in five leading research countries, but their objective is to estimate the extent of international technology diffusion.

The theoretical framework in this paper builds upon research modelling the effect of international knowledge diffusion on productivity in single sector economies (Parente and Prescott 1994; Howitt 2000; Buera and Oberfield 2020; Lind and Ramondo 2022) and studying how endogenous innovation affects comparative advantage (Grossman and Helpman 1990; Somale 2021; Cai, Li and Santacreu 2022). It is also related to product cycle theories of imitation and trade (Krugman 1979; Grossman and Helpman 1991), learning-by-doing models of how initial conditions shape long-run comparative advantage (Redding 1999), recent papers on innovation and/or imitation by incumbent firms (Atkeson and Burstein 2010; Perla and Tonetti 2014; Akcigit and Kerr 2018), and to work by Akcigit, Ates and Impullitti (2018) who study how technology gaps arise from endogenous firm-level innovation in a Schumpeterian economy with two asymmetric countries. Relative to these literatures, the theoretical contribution of this paper lies in developing a model with endogenous innovation and adoption that facilitates the quantitative analysis of technology gaps in the global economy. In particular, allowing for asymmetric countries and industries, trade costs, and firm-level selection between R&D and adoption enhances the mapping between model and data.

The methodology used to estimate innovation-dependence from bilateral trade data is related to empirical studies that test for comparative advantage using the interaction of country and industry characteristics (Romalis 2004; Nunn 2007; Manova 2013) and, particularly, to the approach developed by Costinot (2009) to reveal cross-country variation in institutional quality. Costinot, Donaldson and Komunjer (2012), Hanson, Lind and Muendler (2013) and Levchenko and Zhang (2016) show how structural gravity models can be used to infer productivity differences from trade flows. In their work Hanson, Lind and Muendler (2013) and Levchenko and Zhang (2016) analyze how the pattern of comparative advantage changes over time, while remaining agnostic about mechanisms, whereas this paper provides a theory and quantification of cross-sectional variation in steady state technology gaps. An alternative approach to measuring international technology differences is to use data on the adoption of specific technologies (Caselli and Coleman 2001; Comin, Hobijn and Rovito 2009; Comin and Mestieri 2018). Consistent with my model, such studies find that the rate at which new technologies are adopted differs greatly across countries and is strongly positively correlated with GDP per capita. The framework presented in this paper shows how technology gaps can be quantified using trade data even when direct measures of technology use are unavailable, as is the case for many technologies and sectors.

1 Technology Gap Model

This section develops a model of technology gaps and trade. There are S countries indexed by s and J industries indexed by j. I assume that output and factor markets are competitive, time

t is continuous and all model parameters are time invariant. To simplify notation, I suppress the dependence of endogenous variables on time except when necessary to avoid confusion.

1.1 Production

Within each country, all firms in a given industry produce the same homogeneous output good. However, output is differentiated by country of origin following Armington (1969), implying that the output price p_{js} in industry j is country-specific.

Firms differ in their productivity θ, which is a time-varying, firm-level state variable. Labor is the only factor of production and a firm with productivity θ that employs l^P production workers produces output:

$$y = \theta \left(l^P\right)^{\beta}, \qquad \text{with } 0 < \beta < 1. \tag{1}$$

The assumption $\beta < 1$ implies that there are decreasing returns to scale in production.[4]

At each moment in time, firms in industry j and country s choose production employment to maximize the flow of production profits $\pi^P = p_{js}y - w_s l^P$ taking the output price p_{js}, the wage w_s and productivity θ as given. Solving the profit maximization problem yields:

$$l^P_{js}(\theta) = \left(\frac{\beta p_{js}\theta}{w_s}\right)^{\frac{1}{1-\beta}}, \qquad \pi^P_{js}(\theta) = (1-\beta)\left(\frac{\beta}{w_s}\right)^{\frac{\beta}{1-\beta}} \left(p_{js}\theta\right)^{\frac{1}{1-\beta}}. \tag{2}$$

Employment, output and profits are all increasing in the firm's productivity and the output price, but decreasing in the wage level.

1.2 Technology Investment

Each firm's productivity grows over time at a rate that depends upon its investment in technology upgrading. Firms can choose between two types of technology investment: R&D and adoption. R&D investment seeks to create new ideas and technologies through innovation, while adoption is aimed at learning about and implementing existing production techniques.

R&D technology. Firms are heterogeneous in their R&D capability ψ. R&D capability is a time invariant firm characteristic that increases the efficiency of R&D investment. A firm with capability ψ and productivity θ that employs l^R workers to undertake R&D (and does not invest in adoption) has productivity growth given by:

[4] Decreasing returns to scale ensure the firm's static profit maximization problem is concave. Concavity could also result from firms facing downward sloping demand curves. In an environment where each firm produces a differentiated variety with a constant elasticity of substitution between varieties and there is monopolistic competition between firms, the firm would face an equivalent optimization problem. However, this alternative would make the model analytically intractable in general equilibrium given the existence of many asymmetric countries and industries.

$$\frac{\dot{\theta}}{\theta} = \psi B_s \left(\frac{\theta}{\chi_{js}^R}\right)^{-\gamma_j} \left(l^R\right)^{\alpha} - \delta, \tag{3}$$

where $B_s > 0$, $\gamma_j > 0$, $\delta > 0$ and $\alpha \in (0, 1)$ are parameters, and χ_{js}^R denotes the R&D knowledge level in industry j and country s. The knowledge level χ_{js}^R is non-rival and does not vary across firms.

Conditional on current productivity and R&D employment, equation (3) shows that productivity growth is increasing in the firm's R&D capability. At the same time, conditional on capability and R&D employment, productivity growth is decreasing in the firm's current productivity with elasticity γ_j. This implies that there exists an advantage of backwardness, which benefits firms further from the technology frontier.[5] I allow the strength of the advantage of backwardness γ_j to vary by industry to capture differences in the extent to which generating new ideas and techniques is harder for more productive firms.

The returns to R&D also depend upon country-level R&D efficiency B_s, which captures variation in the quality of a country's national innovation system. Countries with a higher R&D efficiency B_s have an absolute advantage in R&D. The parameter α determines the returns to scale in R&D, while δ is the rate at which a firm's technical knowledge depreciates causing its productivity to decline. The assumption $\delta > 0$ captures obsolescence of previously acquired techniques as well as loss of knowledge resulting from labor force turnover.[6]

Because knowledge is partially non-excludable, R&D generates knowledge spillovers that allow firms to build upon the knowledge created by past innovations. The knowledge level χ_{js}^R captures these spillovers and the specification of the R&D technology implies that a higher knowledge level increases productivity growth all else equal.

Analysis of cross-border knowledge flows finds that domestic spillovers are stronger than international spillovers (Branstetter 2001; Keller 2002) and that the geographic localization of spillovers may vary by industry due to differences in the importance of tacit knowledge, cross-border communication, and whether production techniques must be adapted to local requirements (Evenson and Westphal 1995; Peri 2005). To model this geography of knowledge spillovers, I assume the knowledge level χ_{js}^R depends upon both the domestic productivity frontier and global knowledge capital accumulated through past R&D investments.

Formally, let ω index firms and let Ω_{js} denote the set of firms operating in industry j in country s. Define $\theta_{js}^{\max}(\omega) = \sup_{\tilde{\omega} \in \Omega_{js}, \tilde{\omega} \neq \omega} \{\theta(\tilde{\omega})\}$ as the supremum of the productivity of all firms in

[5] Using industry level data for OECD countries, Griffith, Redding and Van Reenen (2004) find that the effect of R&D on productivity growth is increasing in distance to the frontier. At the firm level, Bartelsman, Haskel and Martin (2008) and Griffith, Redding and Simpson (2009) both estimate that lower productivity relative to the domestic frontier raises productivity growth in the UK.

[6] Doraszelski and Jaumandreu (2013) estimate the persistence of productivity and conclude that "old knowledge is hard to keep" (p.1341).

industry j in country s excluding firm ω. The definition implies $\theta_{js}^{\max}(\omega)$ is exogenous to firm ω. In equilibrium, there will always be a continuum of firms at the productivity frontier.[7] Therefore, $\theta_{js}^{\max}(\omega) = \theta_{js}^{\max}$ and does not vary with ω. The R&D knowledge level χ_{js}^{R} of industry j in country s is then given by:

$$\chi_{js}^{R} = \left(\theta_{js}^{\max}\right)^{\frac{\kappa_j}{1+\kappa_j}} \chi_j^{\frac{1}{1+\kappa_j}}, \tag{4}$$

where χ_j denotes global knowledge capital in industry j. This specification assumes all knowledge spillovers occur within industries.[8]

The knowledge level depends upon domestic spillovers through $\theta_{js}^{\max}$ and global spillovers through χ_j. The parameter $\kappa_j > 0$ determines the localization of knowledge spillovers, which varies by industry. A higher κ_j implies spillovers are more localized because the elasticity of the knowledge level to the domestic productivity frontier is increasing in κ_j, while the elasticity to global knowledge capital is decreasing. Since knowledge spillovers are localized, firms in countries with a greater frontier productivity benefit from access to a higher knowledge level.

Global knowledge capital χ_j is a state variable of the world economy that increases over time as R&D investment leads to the creation of new ideas and technologies. I assume growth in χ_j depends upon a weighted sum of R&D investment by all firms in all countries:

$$\frac{\dot{\chi}_j}{\chi_j} = \sum_{s=1}^{S} M_{js} \int_{\psi^{\min}}^{\psi^{\max}} \lambda_{js}(\psi) l_{js}^{R}(\psi) dG(\psi). \tag{5}$$

where M_{js} denotes the mass of firms that produce good j in country s, $\lambda_{js}(\psi) \geq 0$ determines the strength of R&D spillovers and $G(\psi)$ is the cumulative distribution function of R&D capabilities, which is assumed to be continuous with support $\left[\psi^{\min}, \psi^{\max}\right]$ and does not vary across countries. Note that this specification allows the strength of R&D spillovers to vary by country, industry and the firm's R&D capability. However, adoption investment does not affect global knowledge capital because it does not generate new ideas.

Adoption technology. Although innovation and imitation are closely related activities (Rosenberg 1990), adoption differs from R&D in two important ways. First, it does not require the rare combination of firm capability and institutional support that enables knowledge creation. Therefore, I assume neither firm-level R&D capability ψ nor country-level R&D efficiency B_s affect the efficiency of adoption investment B^A. It follows that firms with higher R&D capability and countries with higher R&D efficiency have a relative advantage at innovation compared to adoption.

[7] In steady state this observation follows from assuming that there exist a continuum of firms with each capability ψ. Outside steady state it also requires assuming an initial condition in which there are either zero or a continuum of firms with each (ψ, θ) pair.

[8] Section 4.2 generalizes the model to include inter-industry knowledge spillovers.

Second, while both R&D and adoption draw upon existing knowledge, prior inventions are more useful to adopters than innovators. Consequently, I assume that the adoption knowledge level is greater than the R&D knowledge level $\chi_{js}^{A}=\eta\chi_{js}^{R}$ with $\eta>1$.[9]

Suppose the productivity growth of a firm that employs l^R R&D workers and l^A adoption workers is given by:

$$\frac{\dot{\theta}}{\theta}=\theta^{-\gamma_j}\left\{\left[\psi B_s\left(\chi_{js}^{R}\right)^{\gamma_j}\right]^{\frac{1}{\alpha}}l^R+\left[B^A\left(\chi_{js}^{A}\right)^{\gamma_j}\right]^{\frac{1}{\alpha}}l^A\right\}^{\alpha}-\delta. \tag{6}$$

This specification implies that the decreasing returns to scale in technology investment generated by $\alpha<1$ apply to the firm's combined employment of R&D and adoption workers, meaning that no firm will invest in both R&D and adoption simultaneously. Firms invest in R&D if $\psi B_s\left(\chi_{js}^{R}\right)^{\gamma_j}>B^A\left(\chi_{js}^{A}\right)^{\gamma_j}$ and adoption otherwise. Moreover, for firms that choose adoption, productivity growth is given by:

$$\frac{\dot{\theta}}{\theta}=B^A\left(\frac{\theta}{\chi_{js}^{A}}\right)^{-\gamma_j}\left(l^A\right)^{\alpha}-\delta, \tag{7}$$

showing that the adoption technology has the same functional form as the R&D technology in equation (3).

Optimal technology investment. Each firm chooses paths for employment in R&D and adoption to maximize its value subject to productivity growth satisfying equation (6). Firms take the current and future values of χ_{js}^{R} and χ_{js}^{A} as given when making technology investments. Because technology investments affect productivity growth, but not the current value of θ, the technology investment problem is separable from the firm's static production decision.

Let $V_{js}(\psi,\theta)$ be the value of a firm with capability ψ and productivity θ. $V_{js}(\psi,\theta)$ equals the expected present discounted value of the firm's production profits minus its technology investment costs on the optimal investment path:

$$V_{js}(\psi,\theta)=\sup_{\{l^R,l^A\}}\left\{\int_t^{\infty}\exp\left[-\int_t^{\tilde{t}}\left(\iota_s+\zeta\right)d\hat{t}\right]\left[\pi_{js}^{P}(\theta)-w_s\left(l^R+l^A\right)\right]d\tilde{t}\right\}, \tag{8}$$

where ι_s denotes the interest rate and $\pi_{js}^{P}(\theta)$ is given by (2). All endogenous variables in this expression, including the firm's value function, are time dependent.

[9] This specification is equivalent to assuming that R&D and adoption draw upon the same knowledge stock χ_{js}^{R}, but that knowledge is more useful in adoption than R&D (since $\eta>1$).

1.3 Entry

Entrants must pay a fixed cost to establish a firm. To set-up a unit flow of new firms, a potential entrant must hire f^E workers where $f^E > 0$ is an entry cost parameter. Following the idea flows literature I assume that the capability ψ and initial productivity θ of each entrant are determined by a random draw from the joint distribution of ψ and θ among incumbent firms in the entrants' country and industry at the time the firm is created. Thus, the distribution of productivity θ at each capability level ψ is the same for entrants and incumbents. This specification implies the existence of spillovers from incumbents to entrants within a country-industry pair.[10]

There is free entry and the free entry condition requires that the cost of entry equals the expected value of entry meaning:

$$f^E w_s = \int_{(\psi,\theta)} V_{js}(\psi,\theta) d\tilde{H}_{js}(\psi,\theta), \tag{9}$$

where $\tilde{H}_{js}(\psi,\theta)$ denotes the cumulative distribution function of (ψ,θ) across firms.

Let L^E_{js} be aggregate employment in entry in industry j and country s. Then the total flow of entrants in industry j and country s is L^E_{js}/f^E. Since firms die at rate ζ this means the mass of firms M_{js} evolves according to:

$$\dot{M}_{js} = -\zeta M_{js} + \frac{L^E_{js}}{f^E}. \tag{10}$$

1.4 Closing the Model

To complete the description of the model, we need to define consumer preferences, specify trade costs and impose market clearing conditions.

Each country has a representative consumer with identical preferences who consumes a single consumption good that is produced as a Cobb-Douglas aggregate of industry outputs. The representative consumer has intertemporal preferences with discount rate $\rho > 0$ and unit elasticity of intertemporal substitution, and allocates a fraction μ_j of expenditure to industry j. Let c_s denote consumption per capita, z_s be the price of the consumption good and L_s be the population of country s. There is no population growth.

Within industries let $\sigma > 1$ be the Armington demand elasticity, which determines the substitutability of output produced in different countries. Suppose trade costs take the iceberg form,

[10] In Sampson (2016a) spillovers from incumbents to entrants lead to endogenous growth through a dynamic selection mechanism. In this paper the dynamic selection mechanism is absent because there is no fixed cost of production, meaning that firm exit is not endogenous. Instead, R&D investment by incumbent firms is the source of long-run growth. Garcia-Macia, Hsieh and Klenow (2019) estimate that most growth in US manufacturing comes from incumbent firms, rather than creative destruction or the introduction of new varieties.

such that $\tau_{js\tilde{s}}$ units of industry j output must be shipped from country s to country $\tilde{s}$ in order for one unit to arrive at the destination.

The system of demand and price index equations implied by these preferences can be found in Appendix A.1. The Armington assumption is sufficient to generate constant elasticity demand, which implies equilibrium bilateral trade flows follow a gravity equation. In particular, demand in $\tilde{s}$ for industry j output produced in country s is given by:

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

where $P_{j\tilde{s}}$ denotes the price index for industry j in country $\tilde{s}$.

I assume there is no international lending, meaning asset markets clear at the national level. Labor markets clearing also occurs country-by-country, while output markets clear at the country-industry level. I let global consumption expenditure be the numeraire, implying $\sum_{s=1}^{S} z_s c_s L_s = 1$.

Finally, to ensure concavity in firms' intertemporal optimization problems, I assume that the returns to scale in production and R&D, the advantage of backwardness and the localization of knowledge spillovers satisfy the following restriction.

Assumption 1. *For all industries j, the parameters of the global economy satisfy:* $\frac{1}{1-\beta} > \gamma_j > \frac{\alpha}{1-\beta} + \frac{\kappa_j \gamma_j}{1+\kappa_j}$.

This completes the specification of the model. Appendix A.1 provides the full set of equilibrium equations and defines an equilibrium of the global economy. The economy's state variables are the joint distributions $\tilde{H}_{js}(\psi, \theta)$ of firms' capabilities and productivity levels for all country-industry pairs, global knowledge capital χ_j in each industry and the mass of firms M_{js} in all countries and industries. An initial condition is required to pin down the initial values of these state variables. Note that, apart from any differences in initial conditions, the only exogenous sources of cross-country variation are differences in R&D efficiency B_s, population L_s and trade costs $\tau_{js\tilde{s}}$.

2 Balanced Growth Path

This section characterizes a balanced growth path equilibrium of the global economy, focusing primarily on how R&D efficiency affects comparative advantage and international income inequality. Full details of the solution together with proofs of the propositions can be found in Appendix A.

Let $H_{js}(\theta)$ be the cumulative distribution function of productivity in industry j and country s. I define a balanced growth path as an equilibrium in which all aggregate country and industry variables have constant growth rates and the productivity distributions $H_{js}(\theta)$ shift outwards at constant rates. Appendix A.2 shows that, on any balanced growth path, the existence of cross-border knowledge spillovers implies $H_{js}(\theta)$ must shift outwards at the same rate g_j in all countries.[11] Moreover, rising productivity is the only source of growth and the growth rate of consumption per capita $q = \sum_j \mu_j g_j$ is the same everywhere. It follows that, on a balanced growth path, cross-country heterogeneity leads to differences in the levels, not growth rates, of endogenous variables.

2.1 Firm Productivity Dynamics

Productivity dynamics depend upon firms' technology investment choices. How do firms behave on a balanced growth path? To solve for optimal firm behavior, we start by determining whether firms invest in R&D or adoption. A higher capability ψ increases the returns to R&D investment, but not to adoption investment. Consequently, there exists a capability threshold ψ^*_{js} such that firms invest in R&D if and only if their capability exceeds ψ^*_{js}. From equation (6) and $\chi^A_{js} = \eta\chi^R_{js}$ we have:

$$\psi^*_{js} = \eta^{\gamma_j} \frac{B^A}{B_s}, \tag{11}$$

which implies that the R&D threshold ψ^*_{js} is increasing in the advantage of backwardness γ_j, decreasing in R&D efficiency B_s and independent of the firm's current productivity. This means that, on the extensive margin, there is more R&D in industries where the advantage of backwardness is smaller and in countries that are better at R&D.[12]

Now consider the R&D investment problem faced by a firm with capability $\psi \geq \psi^*_{js}$. Let $\phi \equiv \left(\theta/\chi^R_{js}\right)^{\frac{1}{1-\beta}}$ be the firm's productivity relative to the R&D knowledge level. I show in Appendix A.3 that changing variables from θ to ϕ allows the firm's problem to be written as an optimal control problem in which the payoff function depends upon time only through exponential discounting. Consequently, the firm's value is a stationary function of its relative productivity and the value function $V_{js}(\psi, \phi)$ satisfies the Hamilton-Jacobi-Bellman equation:

$$(\rho + \zeta)\, V_{js}(\psi, \phi) = \pi^P_{js}(\phi) + \frac{dV_{js}(\psi, \phi)}{dt},$$

where the profit flow $\pi^P_{js}(\phi)$ is given by substituting $\phi = \left(\theta/\chi^R_{js}\right)^{\frac{1}{1-\beta}}$ into equation (2). Appendix A.3 solves the firm's dynamic problem and shows that, on a balanced growth path, it has a unique,

[11] The assumption, embodied in equation (4), that the knowledge level χ^R_{js} is homogeneous of degree one in the pair $\left(\theta^{\max}_{js}, \chi_j\right)$ is a necessary condition for the existence of a balanced growth path. Since productivity growth depends upon current productivity relative to the R&D and adoption knowledge levels, this assumption ensures knowledge spillovers are sufficiently strong to sustain ongoing productivity growth and is analogous to Romer's (1990) assumption that knowledge production is linear in the existing knowledge stock.

[12] I assume the parameter values are such that $\psi^*_{js} \in \left(\psi^{\min}, \psi^{\max}\right)\ \forall s$ implying both adoption and R&D take place in every country.

locally saddle-path stable steady state and that the firm's steady state relative productivity and R&D employment are given by:

$$\phi_{js}^{*}=\left[\alpha\beta^{\frac{\beta}{1-\beta}}\left(\psi B_{s}\right)^{\frac{1}{\alpha}}\left(\frac{p_{js}\chi_{js}^{R}}{w_{s}}\right)^{\frac{1}{1-\beta}}\frac{\left(\delta+g_{j}\right)^{\frac{\alpha-1}{\alpha}}}{\rho+\zeta+\gamma_{j}\left(\delta+g_{j}\right)}\right]^{\frac{\alpha}{\gamma_{j}(1-\beta)-\alpha}}, \tag{12}$$

$$l_{js}^{R*}=\left[\alpha\beta^{\frac{\beta}{1-\beta}}\left(\psi B_{s}\right)^{\frac{1}{\gamma_{j}(1-\beta)}}\left(\frac{p_{js}\chi_{js}^{R}}{w_{s}}\right)^{\frac{1}{1-\beta}}\frac{\left(\delta+g_{j}\right)^{\frac{\gamma_{j}(1-\beta)-1}{\gamma_{j}(1-\beta)}}}{\rho+\zeta+\gamma_{j}\left(\delta+g_{j}\right)}\right]^{\frac{\gamma_{j}(1-\beta)}{\gamma_{j}(1-\beta)-\alpha}}. \tag{13}$$

The steady state and transition dynamics are shown in Figure 1. Along the stable arm, relative productivity and R&D employment increase over time for firms that start with ϕ below ϕ_{js}^{*}, while the opposite is true for firms with initial ϕ above ϕ_{js}^{*}. The existence of an advantage of backwardness is necessary for the stability of the steady state because it introduces a negative relationship between productivity levels and productivity growth, all else constant.

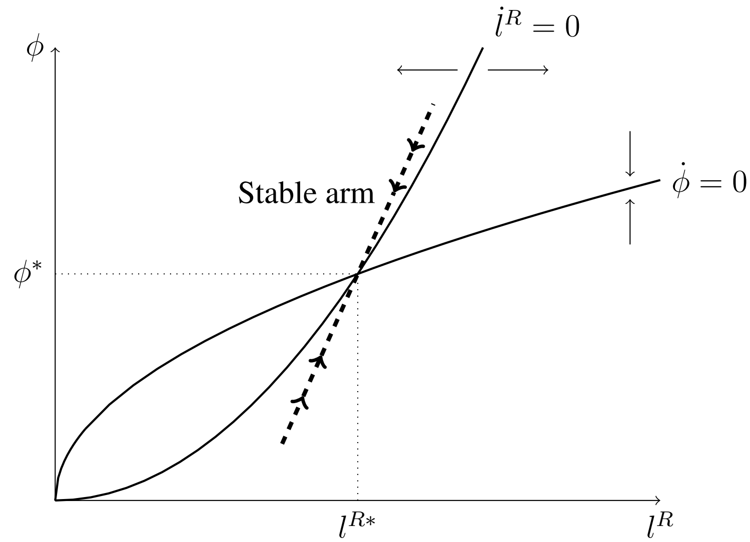

Figure 1: Firm steady state and transition dynamics

The steady state has several important properties. First, in steady state all surviving R&D firms in an industry have the same productivity growth rate g_j, meaning that the steady state satisfies Gibrat's law for surviving firms. Second, ϕ_{js}^{*} is increasing in ψ implying that, within each country-industry pair, more capable firms have higher steady state relative productivity levels. This explains why, even though R&D capability differs across firms, steady state growth rates do not. The advantage of backwardness raises the R&D efficiency of less productive firms and, in steady state, this exactly offsets the disadvantage from low ψ implying all firms grow at the same rate.

Third, the steady state is consistent with two key stylized facts about R&D highlighted by Klette and Kortum (2004): (i) productivity and R&D investment are positively correlated across firms since ϕ^*_{js} and l^{R*}_{js} are both increasing in ψ, and; (ii) among firms with positive R&D investment, R&D intensity is independent of firm size. To see this observe that using (1), (2) and (13) implies the steady state ratio of R&D investment to sales satisfies:

$$\frac{w_s l^{R*}_{js}}{p_{js} y_{js}\left(\phi^*_{js}\right)} = \frac{\alpha\left(\delta + g_j\right)}{\rho + \zeta + \gamma_j\left(\delta + g_j\right)}, \tag{14}$$

which is constant within each industry. R&D intensity is increasing in the returns to scale in R&D α, the knowledge depreciation rate δ and the industry growth rate g_j, and decreasing in the advantage of backwardness γ_j, the interest rate ρ and the firm exit rate ζ.

Fourth, inequality in productivity levels and size between R&D firms is endogenous and steady state inequality is strictly increasing in α and β and strictly decreasing in γ_j.[13] An increase in α raises the returns to scale in R&D which disproportionately benefits higher capability firms that employ more R&D workers. Similarly, an increase in β raises the returns to scale in production giving higher capability, larger firms a greater incentive to raise productivity by increasing R&D investment. By contrast, a higher advantage of backwardness γ_j reduces steady state technology gaps between firms.[14]

The adoption investment problem faced by firms with capability below the R&D threshold ψ^*_{js} is formally equivalent to the R&D investment problem of a firm with threshold capability ψ^*_{js}.[15] It follows that the steady state relative productivity and adoption employment of firms with capability below ψ^*_{js} are given by (12) and (13), respectively, but with $\psi = \psi^*_{js}$. By allowing firms to draw upon existing knowledge, adoption permits firms with capability below the R&D threshold to attain the same steady state productivity level as a firm with R&D capability ψ^*_{js}. Consequently, adopters constitute a fringe of firms with mass $M_{js}G\left(\psi^*_{js}\right)$ that compete with innovators and all have the same steady state productivity.

[13]See the proof of Proposition 1. All inequality results hold for any measure of inequality that respects scale independence and second order stochastic dominance. See Lemma 2 in Sampson (2016b) for a proof of how elasticity changes affect inequality.

[14]In most heterogeneous firm models, such as Melitz (2003), the lower bound is the only endogenously determined parameter of the productivity distribution. This holds not only in static economies, but also in the growth models of Sampson (2016a) and Perla, Tonetti and Waugh (2021). An exception is Bonfiglioli, Crinò and Gancia (2018) who allow firms to choose between receiving productivity draws from distributions with different shapes.

[15]To see this, substitute $\chi^A_{js} = \eta\chi^R_{js}$ and (11) into (7) to obtain:

$$\frac{\dot{\theta}}{\theta} = \psi^*_{js} B_s \left(\frac{\theta}{\chi^R_{js}}\right)^{-\gamma_j} \left(l^A\right)^{\alpha} - \delta.$$

This expression is equivalent to the R&D technology (3) except that the firm's R&D capability ψ has been replaced by the capability threshold ψ^*_{js}.

The discussion above characterizes the productivity dynamics of incumbent firms. However, the evolution of the industry productivity distribution $H_{js}(\theta)$ also depends upon entry and exit. Recall that all firms exit exogenously at rate ζ and that entering firms draw their capability and productivity from the joint distribution of ψ and θ among incumbents. Consequently, net entry does not affect $H_{js}(\theta)$ because the productivity distributions of entering, exiting and incumbent firms are identical. Moreover, if all incumbent firms with capability ψ are in steady state, then each new firm that draws capability ψ enters at its steady state productivity level. Since all surviving firms grow at rate g_j in steady state, it follows that the industry productivity distribution shifts outwards at rate g_j provided all incumbent firms are in steady state.

By contrast, if any incumbent firms are not in steady state, then the shape of the productivity distribution $H_{js}(\theta)$ varies over time as firms transition towards steady state.[16] This is not consistent with balanced growth. Therefore, entry, exit and firms' optimal R&D and adoption investment decisions generate balanced growth if and only if all incumbent firms are in steady state. Proposition 1 summarizes the model's predictions regarding firm-level productivity outcomes on a balanced growth path.

Proposition 1. *Suppose Assumption 1 holds. On a balanced growth path equilibrium all firms in the same industry grow at the same rate and within any country-industry pair:*
(i) Firms that invest in R&D have higher productivity than firms that invest in adoption;
(ii) Among firms that invest in R&D, productivity and R&D employment are strictly increasing in firm capability;
(iii) Productivity inequality between firms is strictly decreasing in the industry's advantage of backwardness, but strictly increasing in the returns to scale in production and R&D and the country's R&D efficiency.

2.2 General Equilibrium

Having characterized firm-level behavior, we can now solve for a balanced growth path equilibrium. For this purpose, let Ψ_{js} be defined by:

$$\Psi_{js} \equiv \int_{\psi^*_{js}}^{\psi^{\max}} \psi^{\frac{1}{\gamma_j(1-\beta)-\alpha}} dG(\psi) + \left(\psi^*_{js}\right)^{\frac{1}{\gamma_j(1-\beta)-\alpha}} G\left(\psi^*_{js}\right). \tag{15}$$

Ψ_{js} is the average effective capability of firms in industry j and country s accounting for the fact that adoption is equivalent to R&D with capability ψ^*_{js}. Ψ_{js} is strictly increasing in the R&D threshold ψ^*_{js} and, therefore, strictly decreasing in country-level R&D efficiency B_s. It captures

[16] Formally, a balanced growth path only requires a mass M_{js} of firms to be in steady state, which allows for individual firms with zero mass to deviate from steady state. I overlook this distinction since it does not matter for industry or aggregate outcomes.

the benefits resulting from selection into adoption, which are larger in countries with lower R&D efficiency.

Using the individual's budget constraint, the definitions of the R&D and adoption knowledge levels, the free entry condition, the goods, labor and asset market clearing conditions and firms' steady state productivity and employment levels, Appendix A.5 shows that on a balanced growth path labor market clearing requires:

$$L_s = \sum_{j=1}^{J} \frac{\mu_j}{\rho+\zeta} \left(\zeta + \beta\rho + \frac{\alpha\rho\left(\delta+g_j\right)}{\rho+\zeta+\gamma_j\left(\delta+g_j\right)} \right) Z_{js}, \tag{16}$$

while asset market clearing implies that asset holdings per capita a_s are given by:

$$a_s L_s = \sum_{j=1}^{J} \frac{\mu_j}{\rho+\zeta} \left(1 - \beta - \frac{\alpha\left(\delta+g_j\right)}{\rho+\zeta+\gamma_j\left(\delta+g_j\right)} \right) w_s Z_{js}, \tag{17}$$

and the productivity growth rate in industry j satisfies:

$$g_j = \sum_{s=1}^{S} \mu_j \frac{\alpha\left(\delta+g_j\right)}{\rho+\zeta+\gamma_j\left(\delta+g_j\right)} \frac{Z_{js}}{\Psi_{js}} \int_{\psi_{js}^*}^{\psi^{\max}} \lambda_{js}(\psi)\psi^{\frac{1}{\gamma_j(1-\beta)-\alpha}} dG(\psi), \tag{18}$$

where:

$$Z_{js} \equiv \sum_{\tilde{s}=1}^{S} \frac{\tau_{js\tilde{s}}^{1-\sigma}\left(\rho a_{\tilde{s}} + w_{\tilde{s}}\right) L_{\tilde{s}} w_s^{-\sigma} \left(B_s \Psi_{js}^{\frac{\gamma_j(1-\beta)}{1+\kappa_j}-\alpha} \right)^{\frac{(\sigma-1)(1+\kappa_j)}{\gamma_j}}}{\sum_{\hat{s}=1}^{S} \tau_{j\hat{s}\tilde{s}}^{1-\sigma} w_{\hat{s}}^{1-\sigma} \left(B_{\hat{s}} \Psi_{j\hat{s}}^{\frac{\gamma_j(1-\beta)}{1+\kappa_j}-\alpha} \right)^{\frac{(\sigma-1)(1+\kappa_j)}{\gamma_j}}}. \tag{19}$$

Equations (16)-(18), together with the definition of Z_{js} in (19), comprise a system of equations in the $2S + J$ unknown wage levels w_s, asset holdings a_s and industry growth rates g_j. Any solution to this system of equations gives a balanced growth path. I prove in Appendix A.6 that there exists a unique balanced growth path in the case where $J = 1$ and there are no trade costs. More generally, I assume existence and derive results that must hold on any balanced growth path.

The equilibrium conditions show that, conditional on industry growth rates, R&D efficiency affects wages and asset holdings only through the Z_{js} terms. Z_{js} can be interpreted as a measure of industry size since $\frac{Z_{js}}{Z_{j\tilde{s}}} = \frac{L_{js}}{L_{j\tilde{s}}}$ where L_{js} denotes total employment in industry j in country s. Z_{js} depends upon R&D efficiency B_s both directly and indirectly through Ψ_{js}. In an economy without adoption this indirect effect is absent. Z_{js} also depends upon labor costs and upon market access, which is a function of trade costs and real demand in each destination.[17]

[17] Characterizing the equilibrium industry growth rates g_j given by equation (18) is not the focus of this paper and

2.3 Technology Gaps and Comparative Advantage

On a balanced growth path relative productivity levels within each industry are stationary. However, the location of the productivity distribution in each country depends upon its R&D efficiency. Consequently, variation in R&D efficiency generates technology gaps. This section characterizes the technology gaps that support a balanced growth path equilibrium and analyzes how technology gaps affect comparative advantage.

Let $\overline{\theta}^*_{js} \equiv \left[\mathbb{E}\left(\theta^*_{js}\right)^{\frac{1}{1-\beta}}\right]^{1-\beta}$ denote the average steady state productivity of firms in country s and industry j. The technology gap between countries s and $\tilde{s}$ in industry j is given by:

$$\frac{\overline{\theta}^*_{js}}{\overline{\theta}^*_{j\tilde{s}}} = \left[\frac{B_s}{B_{\tilde{s}}}\left(\frac{\Psi_{js}}{\Psi_{j\tilde{s}}}\right)^{\frac{\gamma_j(1-\beta)}{1+\kappa_j}-\alpha}\right]^{\frac{1+\kappa_j}{\gamma_j}}, \tag{20}$$

which shows that B_s has a direct positive effect on productivity, as well as an indirect negative effect through Ψ_{js}.[18] The direct effect results from R&D being more productive, all else equal, when B_s is higher. The indirect effect occurs because countries with higher B_s have a lower R&D threshold ψ^*_{js}, which reduces average effective capability Ψ_{js}. However, the direct effect is always stronger than the indirect effect, meaning that the net effect of B_s on average productivity is strictly positive.

Equation (20) also implies that R&D efficiency differences are the only source of international technology gaps in this model. Technology gaps do not depend upon trade costs because trade costs do not affect the price-wage ratio p_{js}/w_s, which determines relative technology investment rates and, consequently, relative productivity levels. This ratio is pinned down by the free entry condition independently of trade costs (see Appendix A.5), meaning that any potential impact of trade cost variation on technology investment is offset by adjustments in net entry. The finding that technology gaps are independent of trade costs relies on the assumption that trade does not affect international knowledge spillovers. Linking knowledge spillovers to trade, as in Baldwin and Robert-Nicoud (2008) or Buera and Oberfield (2020), would introduce an additional source of variation in technology gaps.

---

the counterfactual analysis in Section 3 does not require solving for g_j. However, to offer insight into the determinants of growth in this economy, Appendix A.6 shows that in a single sector version of the model growth is increasing in the R&D spillovers $\lambda_s(\cdot)$, the size of each country L_s and the R&D efficiency of each country B_s, but decreasing in the adoption knowledge premium η and adoption efficiency B^A. Growth is also higher in the open economy than autarky due to the existence of global knowledge spillovers, but does not depend upon the localization of knowledge spillovers κ or the level of trade costs $\tau_{s\tilde{s}}$. Lower trade costs increase the effective size of export markets, but also expose domestic firms to increased import competition. In the single sector version of the model, as in Grossman and Helpman (1991, ch.9) and Eaton and Kortum (2001), these effects exactly offset, leaving R&D employment and growth unchanged.

[18]To see that the indirect effect is negative note that $\gamma_j(1-\beta) > \alpha(1+\kappa_j)$ by Assumption 1 and that Ψ_{js} is decreasing in B_s by (11).

The magnitude of international technology gaps is determined by the elasticity of productivity to R&D efficiency, which I call innovation-dependence ID_{js} since it controls the extent to which countries benefit from being more innovative. Formally, define:

$$ID_{js} \equiv \frac{\partial \log \left(B_s \Psi_{js}^{\frac{\gamma_j(1-\beta)}{1+\kappa_j}-\alpha} \right)^{\frac{1+\kappa_j}{\gamma_j}}}{\partial \log B_s} = \frac{1+\kappa_j}{\gamma_j} + \left[(1-\beta) - \frac{\alpha\,(1+\kappa_j)}{\gamma_j} \right] \frac{\partial \log \Psi_{js}}{\partial \log B_s}. \tag{21}$$

In general, innovation-dependence may vary across industries due to differences in γ_j and κ_j and across countries due to differences in the elasticity of Ψ_{js} to B_s. However, in Section 3 I calibrate a first-order approximation of the model in which $\frac{\partial \log \Psi_{js}}{\partial \log B_s}$ is constant across countries, implying that innovation-dependence only varies by industry.

Innovation-dependence is decreasing in the advantage of backwardness γ_j and increasing in the localization of knowledge spillovers κ_j.[19] A higher advantage of backwardness raises the relative efficiency of technology investment at less productive firms and decreases the share of firms that undertake R&D. Both these effects reduce innovation-dependence. By contrast, more localized knowledge spillovers increase innovation-dependence by making technology investment more reliant on domestically generated knowledge.

Differences in innovation-dependence across industries give rise to Ricardian comparative advantage. To see this, let $EX_{js\tilde{s}} = \tau_{js\tilde{s}} p_{js} x_{js\tilde{s}}$ denote the value of exports from s to $\tilde{s}$ in industry j inclusive of trade costs. On a balanced growth path:

$$\log EX_{js\tilde{s}} = \upsilon^1_{j\tilde{s}} + (\sigma - 1) \left(\log \overline{\theta}^*_{js} - \log w_s - \log \tau_{js\tilde{s}} \right), \tag{22}$$

where $\upsilon^1_{j\tilde{s}}$ is a destination-industry specific term defined in Appendix A.7. Equation (22) implies exports are increasing in average productivity and decreasing in the wage level. An increase in average productivity raises exports by reducing the output price p_{js}, whereas higher wages increase labor costs and raise the output price. Equation (22) also implies that the pattern of comparative advantage is stable on a balanced growth path because productivity and wage growth do not vary by country.

By substituting for $\overline{\theta}^*_{js}$ in (22) we obtain:

$$\log EX_{js\tilde{s}} = \upsilon^2_{j\tilde{s}} + (\sigma - 1) \left(\frac{1+\kappa_j}{\gamma_j} \log B_s + \frac{\gamma_j(1-\beta) - \alpha(1+\kappa_j)}{\gamma_j} \log \Psi_{js} - \log w_s - \log \tau_{js\tilde{s}} \right), \tag{23}$$

[19] See the proof of Proposition 2 for details.

showing that R&D efficiency affects exports both directly and indirectly through Ψ_{js} and w_s. In addition, conditional on the wage, the elasticity of exports to R&D efficiency equals $\sigma - 1$ times innovation-dependence ID_{js}. This observation motivates the calibration strategy in Section 3.

Using equation (23), we can characterize the pattern of comparative advantage on a balanced growth path. We have:

$$\frac{\partial^2 \log EX_{js\tilde{s}}}{\partial \gamma_j \partial \log B_s} = (\sigma - 1)\frac{\partial ID_{js}}{\partial \gamma_j} < 0, \qquad \frac{\partial^2 \log EX_{js\tilde{s}}}{\partial \kappa_j \partial \log B_s} = (\sigma - 1)\frac{\partial ID_{js}}{\partial \kappa_j} > 0.$$

Thus, countries with higher R&D efficiency have a comparative advantage in more innovation-dependent industries where γ_j is lower and κ_j is higher. Proposition 2 summarizes these results.

Proposition 2. *Suppose Assumption 1 holds. On a balanced growth path equilibrium:*
(i) Countries with higher R&D efficiency have greater average productivity in each industry;
(ii) Countries with higher R&D efficiency have a comparative advantage in more innovation-dependent industries where the advantage of backwardness is smaller and the localization of knowledge spillovers is greater.

It is worth noting that the proof of Proposition 2 does not rely on the labor, output or asset market clearing conditions. This implies that the pattern of comparative advantage on a balanced growth path is independent of how the market clearing conditions are specified.

Proposition 2 characterizes comparative advantage assuming γ_j and κ_j are the only parameters that vary across industries. But it is straightforward to check that Proposition 2 continues to hold if there is also industry-level heterogeneity in the Armington elasticity σ_j, the capability distribution $G_j(\psi)$, the returns to scale in production β_j, the returns to scale in R&D α_j, the knowledge depreciation rate δ_j, the adoption knowledge advantage η_j, the exit rate ζ_j and the entry cost f_j^E.

In this case, countries with higher R&D efficiency also have a comparative advantage in industries with higher returns to scale in production β_j and R&D α_j and in industries with a lower adoption knowledge advantage η_j. Higher returns to scale in production and R&D increase the average technology gap between innovators and adopters within countries as shown in Proposition 1, which gives a comparative advantage to countries where a higher proportion of firms invest in R&D. A higher η_j raises the R&D threshold by (11), which shrinks international technology gaps since the adoption technology is independent of R&D efficiency.

2.4 International Inequality

How do wages, income and consumption differ across countries on a balanced growth path? The simplest case to consider is a single sector economy with free trade. In this case equations (16), (19) and (20) yield:

$$\frac{w_s}{w_{\tilde{s}}}\left(\frac{L_s}{L_{\tilde{s}}}\right)^{\frac{1}{\sigma}}=\left(\frac{\overline{\theta}_s^*}{\overline{\theta}_{\tilde{s}}^*}\right)^{\frac{\sigma-1}{\sigma}}=\left[\frac{B_s}{B_{\tilde{s}}}\left(\frac{\Psi_s}{\Psi_{\tilde{s}}}\right)^{\frac{\gamma(1-\beta)}{1+\kappa}-\alpha}\right]^{\frac{1+\kappa}{\gamma}\frac{\sigma-1}{\sigma}},$$

which shows that the relative wage of country s is increasing in its relative average productivity and, consequently, in its R&D efficiency.[20] Moreover, differentiating this expression gives that the elasticity of the relative wage to R&D efficiency equals $(\sigma-1)/\sigma$ times innovation-dependence. From Proposition 2 innovation-dependence is decreasing in the advantage of backwardness and increasing in the localization of knowledge spillovers. Thus, wage inequality caused by differences in R&D efficiency is higher when the advantage of backwardness is smaller and when knowledge spillovers are more localized.

The intertemporal budget constraint implies consumption per capita depends upon assets per capita, wages and the consumption price through:

$$c_s=\frac{\rho a_s+w_s}{z_s},$$

implying that consumption per capita equals real income per capita. With a single industry, assets per capita a_s are proportional to w_s by (16) and (17). Because of free trade all countries also face the same consumption price z_s, meaning that consumption per capita c_s is proportional to w_s. It follows that international inequality in incomes and consumption is the same as inequality in wages and is increasing in the degree of innovation-dependence. Proposition 3 summarizes these results.

Proposition 3. *Suppose Assumption 1 holds, the economy has a single industry and there is free trade. On a balanced growth path equilibrium:*
(i) Each country's wage, income per capita and consumption per capita relative to other countries is strictly increasing in its R&D efficiency;
(ii) International inequality in wages, income per capita and consumption per capita due to differences in R&D efficiency is greater when innovation-dependence is higher. Consequently, inequality is strictly decreasing in the advantage of backwardness and strictly increasing in the localization of knowledge spillovers.

In the general case with trade costs and many industries, innovation-dependence continues to be the key determinant of the mapping from R&D efficiency differences to international inequality. In particular, equations (16), (17) and (19) show that, conditional on industry growth rates, B_s enters the balanced growth path equations for w_s, a_s, z_s and, therefore, c_s only through the term $B_s\Psi_{js}^{\frac{\gamma_j(1-\beta)}{1+\kappa_j}-\alpha}$.[21] It follows that the elasticities of w_s, a_s, z_s and c_s to B_s can all be expressed in terms of the innovation-dependence levels in the J industries.

[20] The relative wage is also decreasing in relative population $L_s/L_{\tilde{s}}$ due to the assumption of Armington demand.
[21] See Appendix A.8 for the derivation of equilibrium consumption prices z_s.

A simple example arises when all industries are non-tradable. In this case, equilibrium consumption per capita satisfies:

$$\frac{c_s}{c_{\tilde{s}}} = \prod_{j=1}^{J} \left[\frac{B_s}{B_{\tilde{s}}} \left(\frac{\Psi_{js}}{\Psi_{j\tilde{s}}} \right)^{\frac{\gamma_j(1-\beta)}{1+\kappa_j} - \alpha} \right]^{\frac{\mu_j(1+\kappa_j)}{\gamma_j}},$$

which shows that relative consumption per capita in country s is increasing in R&D efficiency B_s with an elasticity $\sum_{j=1}^{J} \mu_j ID_{js}$ that equals the expenditure share weighted average of industry nnovation-dependence levels. Consequently, a higher innovation-dependence in any industry raises the elasticity of relative consumption per capita to R&D efficiency.

With finite trade costs, exporters' market shares vary by importer meaning that the terms-of-trade effects of productivity differences are country-specific. Consequently, the relationship between R&D efficiency, wages and incomes is more complex and depends upon the entire system of equations (16)-(19). However, by estimating innovation-dependence and calibrating the model, it is possible to quantify the impact of variation in R&D efficiency on wage and income inequality in the general model. The remainder of the paper takes up this challenge.

3 Quantitative Analysis

This section calibrates the model and quantifies the effect of R&D efficiency differences on trade flows and international inequality.

3.1 Model Approximation

Before calibrating the model, I log-linearize the balanced growth path equilibrium conditions by taking a first order approximation to average effective capability Ψ_{js} around an equilibrium where the share of firms that perform R&D vanishes. The approximation makes the equilibrium conditions log-linear in R&D efficiency B_s, which facilitates the calibration.

Suppose the R&D capability distribution $G(\psi)$ is truncated Pareto with lower bound $\psi^{\min} = 1$ and shape parameter k, where $k > \frac{1}{\gamma_j(1-\beta)-\alpha}$ for all industries j.[22] Using this functional form in (15) to compute average effective capability Ψ_{js} and letting $\psi^{\max} \to \infty$ yields:[23]

$$\Psi_{js} \approx \left(\psi_{js}^{*}\right)^{\frac{1}{\gamma_j(1-\beta)-\alpha}} \left[1 + \frac{\left(\psi_{js}^{*}\right)^{-k}}{k\left[\gamma_j(1-\beta)-\alpha\right]-1} \right]. \tag{24}$$

[22]The assumption $\psi^{\min} = 1$ is without loss of generality.

[23]Appendix A.9 provides further details on the derivation of the approximation.

Taking a first order approximation to this expression for large ψ_{js}^* then gives:

$$\Psi_{js} \approx \left(\psi_{js}^*\right)^{\frac{1}{\gamma_j(1-\beta)-\alpha}} = \left(\eta^{\gamma_j}\frac{B^A}{B_s}\right)^{\frac{1}{\gamma_j(1-\beta)-\alpha}}, \tag{25}$$

where the second equality follows from the solution for the R&D threshold in equation (11). Since the approximation drops terms of order $\left(\psi_{js}^*\right)^{-k}$, it is valid provided $\left(\psi_{js}^*\right)^{-k}$ is small. With $\psi^{\max} \to \infty$, $\left(\psi_{js}^*\right)^{-k}$ equals the share of firms that undertake R&D. In UK data for 2008-09, 9.9% of goods firms report performing R&D, which is consistent with $\left(\psi_{js}^*\right)^{-k}$ being small. Section 3.6 computes an upper bound on the approximation error in the counterfactual results and shows that it is not quantitatively important.

With this approximation to Ψ_{js}, the innovation-dependence of each industry is constant across countries. Applying the approximation to $\frac{\partial \log \Psi_{js}}{\partial \log B_s}$ and using equation (21) yields:

$$ID_{js} = ID_j = \frac{(1-\beta)\,\kappa_j}{\gamma_j\,(1-\beta)-\alpha}. \tag{26}$$

Note that ID_j is increasing in κ_j, α and β, and decreasing in γ_j, meaning that the signs of the relationships between these parameters and innovation-dependence, characterized in Section 2.3, are unaffected by taking the approximation.

Now substituting the approximation to Ψ_{js} into the definition of Z_{js} in equation (19) yields:

$$Z_{js} = \sum_{\tilde{s}=1}^{S} \frac{\tau_{js\tilde{s}}^{1-\sigma}\left(\rho a_{\tilde{s}} + w_{\tilde{s}}\right) L_{\tilde{s}} w_s^{-\sigma} B_s^{(\sigma-1)ID_j}}{\sum_{\hat{s}=1}^{S} \tau_{j\hat{s}\tilde{s}}^{1-\sigma} w_{\hat{s}}^{1-\sigma} B_{\hat{s}}^{(\sigma-1)ID_j}}, \tag{27}$$

implying that the industry size measure Z_{js} depends upon R&D efficiency only through $B_s^{ID_j}$. Section 3.4 shows how R&D and bilateral trade data can be used to obtain model-consistent estimates of $B_s^{ID_j}$.

3.2 Calibration Strategy

The goal of the counterfactual analysis is to quantify how R&D efficiency affects trade, wages and incomes, not to analyze growth rates. Focussing on this objective simplifies the calibration by reducing the number of parameters needed to calibrate the model. To see this, first note that equation (14) implies the R&D intensity of innovative firms $FiRD_j$ is given by:

$$FiRD_j = \frac{\alpha(\delta + g_j)}{\rho + \zeta + \gamma_j(\delta + g_j)}.$$

Substituting this equation into the balanced growth path equilibrium conditions (16) and (17) then yields:

$$L_s = \sum_{j=1}^{J} \frac{\mu_j \left(\zeta + \beta\rho + \rho FiRD_j\right)}{\rho + \zeta} Z_{js}, \quad a_s L_s = \sum_{j=1}^{J} \frac{\mu_j \left(1 - \beta - FiRD_j\right)}{\rho + \zeta} w_s Z_{js}. \tag{28}$$

Equation (28) together with the expression for Z_{js} in equation (27) can be used to solve for wages w_s and assets a_s without fully calibrating the model. Inspection of equations (27) and (28) shows that this approach requires calibrating trade costs $\tau_{js\tilde{s}}$, R&D efficiency B_s, innovation-dependence ID_j, firm-level R&D intensity $FiRD_j$, the discount rate ρ, the Armington elasticity σ, industry expenditure shares μ_j, the returns to scale in production β and the exit rate ζ. However, parameters such as the returns to scale in R&D α, the knowledge depreciation rate δ and the strength of R&D spillovers $\lambda_{js}(\psi)$ are not needed, which reduces the information required to calibrate the model. The cost of adopting this calibration strategy is that the quantitative analysis does not address the determinants of growth.

When solving the calibrated model, I assume the economy has $J-1$ tradable industries and one non-tradable services industry. The goods trade data that I use to calibrate ID_j does not provide information on the innovation-dependence of non-tradables. However, taking the limit of equation (27) as $\tau_{js\tilde{s}} \to \infty$ for all $\tilde{s} \neq s$ implies that if industry j is non-tradable then:

$$Z_{js} = \left(\rho a_s + w_s\right) \frac{L_s}{w_s}.$$

It follows that equilibrium wages w_s and assets a_s do not depend upon innovation-dependence in the non-tradable sector. The intuition for this result is related to the Balassa-Samuelson effect: in an open economy nominal wages are determined by productivity in tradable sectors.

The innovation-dependence of non-tradables does affect real variables through the price index. Consequently, the counterfactual analysis focuses on nominal wages as the main outcome of interest. However, I also calculate real incomes under the assumption that the innovation-dependence of non-tradables equals zero. This assumption will lead the model to underestimate variation in real incomes caused by differences in R&D efficiency if the innovation-dependence of non-tradables is positive.

3.3 Data

This section briefly describes the data sources used for the quantitative analysis. Full details can be found in Appendix C.

The primary data constraint is the limited availability of internationally comparable data on R&D expenditure at the industry-level, which is needed to calibrate R&D efficiency. From the

OECD's ANBERD database, I obtain R&D expenditure for 20 ISIC 2 digit manufacturing industries. The OECD defines R&D as "work undertaken in order to increase the stock of knowledge ... and to devise new applications of knowledge" (OECD 2015, p.44). This definition corresponds to the model's conceptualization of R&D as investment that seeks to expand the knowledge stock through discovering new ideas or developing new production techniques. By contrast, the goal of adoption is to learn about existing knowledge and techniques, meaning adoption investment should not be counted in R&D data.

The coverage of ANBERD at the 2 digit level has improved over time, but the annual data has many missing values. Consequently, I pool data for 2010-14 and, for each year, keep countries where R&D intensity is available for at least two-thirds of industries. This gives a baseline sample of 25 OECD countries with R&D intensity data. As an alternative innovation measure, I also use patent data from the OECD's Patents by technology database.

Value-added, output and trade by 2 digit ISIC industry for 2010-14 are taken from the OECD's STAN database. Gravity variables are from the CEPII gravity data set. Additional country-level variables are obtained from the Penn World Tables, the IMF's International Financial Statistics and the World Bank's World Development Indicators, Worldwide Governance Indicators, Financial Structure Database and Doing Business data set.

The analysis also uses firm-level data on R&D investment in the UK. This data comes from two surveys undertaken by the Office for National Statistics: the Business Enterprise R&D Survey, and; the Annual Business Survey.

3.4 Calibration

The main parameters needed for the calibration are the R&D efficiency of each country B_s and the innovation-dependence of each industry ID_j. This section calibrates these parameters using moments derived from the model's equilibrium conditions and then briefly describes how the remaining parameters are calibrated. The calibrated model has 25 countries and 23 industries (22 tradable goods industries and one non-tradable services industry).

R&D efficiency. R&D efficiency differences can be inferred from cross-country variation in innovation intensity. Let industry R&D intensity RD_{js} be the ratio of industry R&D expenditure to industry value-added. At the firm-level, R&D intensity is the same for all firms that innovate by equation (14). However, because firms are heterogeneous and choose between R&D and adoption, industry R&D intensity depends upon the share of firms that select into R&D. Computing RD_{js} from (1), (2), (12) and (13), imposing the first order approximation for large ψ^*_{js} and using (11) to substitute for ψ^*_{js} gives:

$$RD_{js} = \frac{\alpha(\delta + g_j)}{\rho + \zeta + \gamma_j(\delta + g_j)} \frac{k\left[\gamma_j(1-\beta) - \alpha\right]}{k\left[\gamma_j(1-\beta) - \alpha\right] - 1} \eta^{-k\gamma_j} \left(\frac{B_s}{B^A}\right)^k . \tag{29}$$

Equation (29) shows that R&D intensity is higher in countries with greater R&D efficiency. An increase in B_s results in a larger share of firms performing R&D, which raises RD_{js}.[24]

Using equation (29) to take the ratio of RD_{js} for any pair of countries implies:

$$\frac{RD_{js}}{RD_{j\tilde{s}}} = \left(\frac{B_s}{B_{\tilde{s}}}\right)^k , \tag{30}$$

showing that the relative R&D intensity of countries s and $\tilde{s}$ in industry j depends upon their relative R&D efficiency levels. I use equation (30) to calibrate R&D efficiency differences from within-industry, cross-country variation in observed R&D intensity. In particular, let $b_s \equiv k \log B_s$ denote log R&D efficiency in country s and suppose R&D efficiency is normalized to one for the US, i.e. $B_{US} = 1$. Using data on industry-level R&D intensity for each of the 25 sample countries, I calibrate log R&D efficiency b_s^R as the median (across industries) of $\log\left(RD_{js}/RD_{j\tilde{s}}\right)$ where $\tilde{s} = US$.

In the model, R&D expenditure corresponds directly to innovation investment. However, given the difficulties in measuring R&D and obtaining internationally comparable R&D data, I also calibrate R&D efficiency from data on innovation outputs (patents) instead of inputs (R&D). Appendix D.1 shows how patenting intensity, defined analogously to R&D intensity as the industry-level ratio of patents to value-added, can be used to calibrate R&D efficiency. Let b_s^P denote log R&D efficiency calibrated as the median (across industries) of log patenting intensity in country s relative to the US.

Figure 2 plots b_s^P against b_s^R for the 25 sample countries. The two measures of log R&D efficiency have a correlation of 0.88, although b_s^P has higher variance than b_s^R. Both measures are strongly positively correlated with GDP per capita implying that, on average, countries with higher R&D efficiency are richer.[25] The correlation with log GDP per capita in 2012 is 0.73 for b_s^R and 0.78 for b_s^P. Most of the variation in R&D efficiency is between richer and poorer countries. However, there are notable differences in R&D efficiency even within wealthy countries, for example compare Canada and Italy to France and the US in Figure 2.

Innovation-dependence. Countries with higher R&D efficiency have a comparative advantage

[24] In the model cross-country variation in RD_{js} comes entirely from the extensive margin, but this restriction is not necessary to obtain the industry-level equilibrium conditions used in the calibration. For example, if firm output is the sum of output of a unit mass of non-tradeable tasks and R&D capability has distribution $G(\psi)$ across tasks then all international variation in RD_{js} comes from the intensive margin, but the balanced growth path is otherwise unchanged.

[25] For consistency with the model, GDP per capita is measured as GDP per member of the working age population. See Appendix C for details.

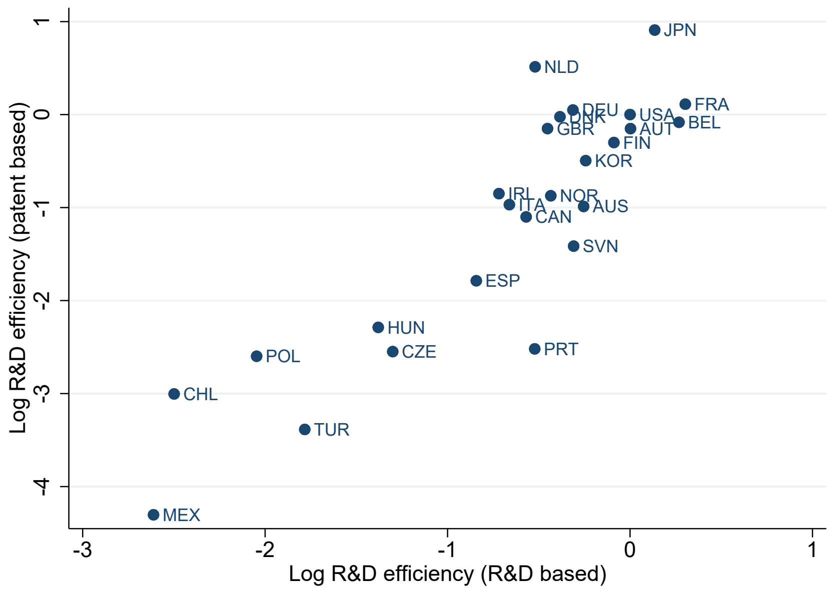

Figure 2: R&D efficiency

Notes: R&D efficiency for 2010-14 calculated using OECD's ANBERD, Patents by technology and STAN databases.

in more innovation-dependent industries, as shown in Proposition 2. Consequently, the correlation between R&D efficiency and bilateral trade flows can be used to estimate each industry's innovation-dependence.[26]

However, a challenge in calibrating innovation-dependence, is that R&D efficiency may be correlated with other country characteristics that affect productivity and trade, such as institutional quality and factor endowments. To allow for this possibility, suppose that instead of equation (1), the production function is given by $y = A_{js}\theta \left(l^P\right)^{\beta}$, where A_{js} is the allocative efficiency of industry j in country s, which is exogenous and time invariant. Otherwise, the model is unchanged. If countries with better national innovation systems also have better economic institutions and policies more broadly, allocative efficiency A_{js} and R&D efficiency B_s will be positively correlated.

It is straightforward to solve the model incorporating A_{js} (see Appendix B.1 for details). Although A_{js} enters the equilibrium conditions, all the theoretical results in Section 2 continue to hold because the effects of allocative efficiency and R&D efficiency on technology gaps and income differences are separable.[27] In particular, using the approximation to Ψ_{js}, the export equation (23) can be written as:

[26] An alternative approach would be to estimate innovation-dependence using productivity data. However, measuring international productivity differences requires comparable cross-country price data, which is less widely available than trade data.

[27] Because of this separability, I use the version of the model without allocative efficiency differences except when estimating innovation-dependence.

$$\log EX_{js\tilde{s}} = \upsilon^3_{j\tilde{s}} + (\sigma - 1)\left(ID_j \log B_s + \log A_{js} - \log w_s - \log \tau_{js\tilde{s}}\right), \tag{31}$$

where $\upsilon^3_{j\tilde{s}} = \upsilon^2_{j\tilde{s}} + \frac{\sigma-1}{\gamma_j}\frac{\gamma_j(1-\beta)-\alpha(1+\kappa_j)}{\gamma_j(1-\beta)-\alpha}\log\left(\eta^{\gamma_j}B^A\right)$. It follows that correlation between allocative efficiency and R&D efficiency could lead to omitted variable bias in estimating the effect of R&D efficiency on trade. In order to alleviate this concern, I use country characteristics known to affect productivity and comparative advantage as proxies for allocative efficiency.

To estimate the exports equation, I also parameterize bilateral trade costs. Following Eaton and Kortum (2002), I model trade costs as a function of gravity variables. In addition, I include exporter-industry fixed effects to capture the possibility that export costs vary by countries as argued by Waugh (2010). Specifically, suppose $\tau_{jss} = 1$ meaning there are no internal trade costs and that international trade costs can be expressed as:

$$\log \tau_{js\tilde{s}} = DIST^i_{js\tilde{s}} + BORD_{js\tilde{s}} + CLANG_{js\tilde{s}} + FTA_{js\tilde{s}} + \delta^1_{js}, \tag{32}$$

where: the impact of bilateral distance on trade costs $DIST^i_{js\tilde{s}}$ depends on which of $i = 1, \ldots, 6$ intervals the distance between countries s and $\tilde{s}$ belongs to: $[0, 375)$, $[375, 750)$, $[750, 1500)$, $[1500, 3000)$, $[3000, 6000)$, or ≥ 6000 miles; $BORD_{js\tilde{s}}$ denotes the effect of sharing a border; $CLANG_{js\tilde{s}}$ gives the effect of sharing a common language; $FTA_{js\tilde{s}}$ is the impact of having a free trade agreement, and; δ^1_{js} is an exporter-industry fixed effect. The impact of all gravity variables on trade costs is allowed to vary by industry.

Using this parameterization of trade costs and rearranging the exports equation (31) yields the specification that I estimate to obtain innovation-dependence:

$$\begin{aligned}\log\left(\frac{EX_{js\tilde{s}}}{EX_{j\tilde{s}\tilde{s}}}\right) - (\sigma-1)\log\left(\frac{w_{\tilde{s}}}{w_s}\right) &= -(\sigma-1)\frac{ID_j}{k}b_{\tilde{s}} - (\sigma-1)A_{j\tilde{s}} \\ &\quad - (\sigma-1)\left(DIST^i_{js\tilde{s}} + BORD_{js\tilde{s}} + CLANG_{js\tilde{s}} + FTA_{js\tilde{s}} + \delta^2_{js}\right) + \epsilon_{js\tilde{s}},\end{aligned} \tag{33}$$

where $\delta^2_{js} = \delta^1_{js} - ID_j b_s/k - A_{js}$ and $\epsilon_{js\tilde{s}}$ captures unmodelled variation in trade costs, productivity and comparative advantage. The left hand side of this expression is observable given a value for $\sigma - 1$. From equation (22), we see that $\sigma - 1$ equals both the trade elasticity and the elasticity of exports to average productivity. Costinot, Donaldson and Komunjer (2012) estimate this elasticity in an Eaton and Kortum (2002) framework. For the baseline calibration, I set $\sigma - 1$ equal to their preferred estimate of 6.53, while Section 3.6 reports robustness checks for alternative values of the trade elasticity.

Equation (33) is estimated including the interaction of industry dummies with $(1-\sigma)b^R_{\tilde{s}}$ on the right hand side. The resulting coefficient estimates give the innovation-dependence of each industry j relative to the shape parameter of the R&D capability distribution ID_j/k. Note that the parameter k cancels out when ID_j/k is multiplied by calibrated log R&D efficiency $b_s = k \log B_s$. When solving the calibrated model, the product of these two terms is sufficient to quantify the impact of R&D efficiency on comparative advantage, wages and income levels (recall equation 27). Consequently, there is no need to calibrate k.

Table 1 reports estimates of ID_j/k obtained from (33) using pooled trade data for 2010-14. The sample includes exports of 117 countries to the 25 OECD importers for which R&D efficiency can be calibrated. It covers 22 ISIC goods industries at the 2 digit level (the 20 manufacturing industries used to calculate R&D efficiency plus the Agriculture and Mining industries). See Appendix C for data details.

Column (a) does not include controls for the importer's allocative efficiency $A_{j\tilde{s}}$. Estimated innovation-dependence is highest in Machinery and equipment, Computers, and Pharmaceuticals, and lowest in Mining and Agriculture. However, these estimates are likely to be biased upwards by correlation between $b_{\tilde{s}}^R$ and $A_{j\tilde{s}}$.

As proxies for allocative efficiency, column (b) adds measures of the importer's institutional quality, business environment and financial development. Institutional quality is measured by the rule of law, control of corruption, government effectiveness, political stability, regulatory quality, and voice and accountability variables from the Worldwide Governance Indicators. Business environment is the country's distance to the frontier in the Doing Business data set. Financial development is measured by the log of private credit as a share of GDP. As expected, including these controls reduces the magnitude of the innovation-dependence estimates, but the pattern of cross-industry variation is similar to column (a).

Column (c) also controls for sources of comparative advantage other than R&D efficiency by including the interaction of industry dummy variables with the importer's rule of law, log private credit to GDP ratio, log physical capital per employee and human capital. This specification allows for comparative advantage due to institutional quality (Nunn 2007), financial development (Manova 2013) and Heckscher-Ohlin effects (Romalis 2004). Adding the comparative advantage controls further reduces the innovation-dependence estimates. Average innovation-dependence in column (c) is 0.28, compared to 0.43 in column (b) and 0.55 in column (a). However, the pattern of variation in innovation-dependence across industries is similar in all three columns.

Finally, column (d) estimates the same specification as column (c), but using the R&D efficiency measure based on patenting intensity $b_{\tilde{s}}^P$. The innovation-dependence estimates in column (d) are smaller than in column (c), reflecting the fact that $b_{\tilde{s}}^P$ exhibits greater dispersion than $b_{\tilde{s}}^R$. But reassuringly the correlation between the estimates in column (c) and those in column (d) is 0.89. In each case, innovation-dependence is largest in the Computers, Machinery and equipment, and Chemicals industries and lowest in Mining. All except two of the 22 innovation-dependence estimates in column (c) and all except six of those in column (d) are positive and significantly different from zero at the 10 percent level.

The baseline calibration uses R&D efficiency calibrated from R&D data and the innovation-dependence estimates in column (c). But I also assess the robustness of the counterfactual results to using patent data to calibrate R&D efficiency together with the innovation-dependence estimates in column (d). In both cases I set innovation-dependence to zero for any industry where the point estimate of innovation-dependence is negative.

Other parameters. The remaining parameters are calibrated as follows using data from 2012 or the nearest available year (see Appendix C for further details). Bilateral trade costs $\tau_{js\tilde{s}}$ are calculated from equation (32) using the coefficient estimates obtained when estimating innovation-dependence. I use the trade cost estimates from the specification estimated in column (c) of Table 1 for the baseline R&D data calibration and those corresponding to column (d) for the patent data calibration. $FiRD_j$ is computed from UK firm-level data. Population L_s is calibrated to the working age population from the World Development Indicators. I assume $\sigma - 1 = 6.53$, consistent with the value of the trade elasticity used to estimate innovation-dependence. Expenditure shares μ_j are calibrated to the average across OECD countries of each industry's share of domestic absorption. The exit rate ζ is set to 0.103, which is the average OECD death rate of employer enterprises in the business economy excluding holding companies. The share of profits in firm revenue before accounting for R&D investment is $1 - \beta$. I set $\beta = 0.85$ implying a profit share of 15% as in Gabler and Poschke (2013) and close to Barkai's (2017, Figure 2b) estimate of the aggregate US profit share in 2012. Finally, I let the discount rate $\rho = 0.04$, which implies a risk free interest rate of 4% per annum. Numerically, solving the model using these parameters with different initial guesses for wages and assets, delivers a unique equilibrium.

3.5 Model Validation

Before undertaking counterfactual analysis, I perform three validation exercises to assess the model's empirical credibility under the assumption that sample countries are on a balanced growth path. First, I compare wages and incomes implied by the calibrated model to their observed values. Second, I examine cross-industry differences in firm-level R&D investment choices. Third, I conduct an out-of-sample test of the model's predictions for comparative advantage.

Wages and incomes. The left hand panel of Figure 3 plots calibrated against observed nominal wages for the calibration using R&D data. There is an upwards sloping, approximately log-linear relationship between the two variables with a correlation of 0.89. However, there is more variation in observed than calibrated wages. The standard deviation of log wages is 2.3 times higher for observed wages. The right hand panel shows that a similar log-linear relationship holds for calibrated and observed income per capita, except that the standard deviation is now 4.7 times higher for observed than for calibrated log income per capita. The calibrated model explains less of the observed variation in income per capita than in wages because assuming that innovation-dependence equals zero in non-tradables reduces cross-country dispersion in consumption prices without affecting equilibrium wages.

Although wage and income data were not targeted by the calibration, Figure 3 shows that the pattern of cross-country wage and income differences is comparable in the model and the data, except that calibrated wages and incomes exhibit less dispersion. Similar results hold for the patent data calibration.

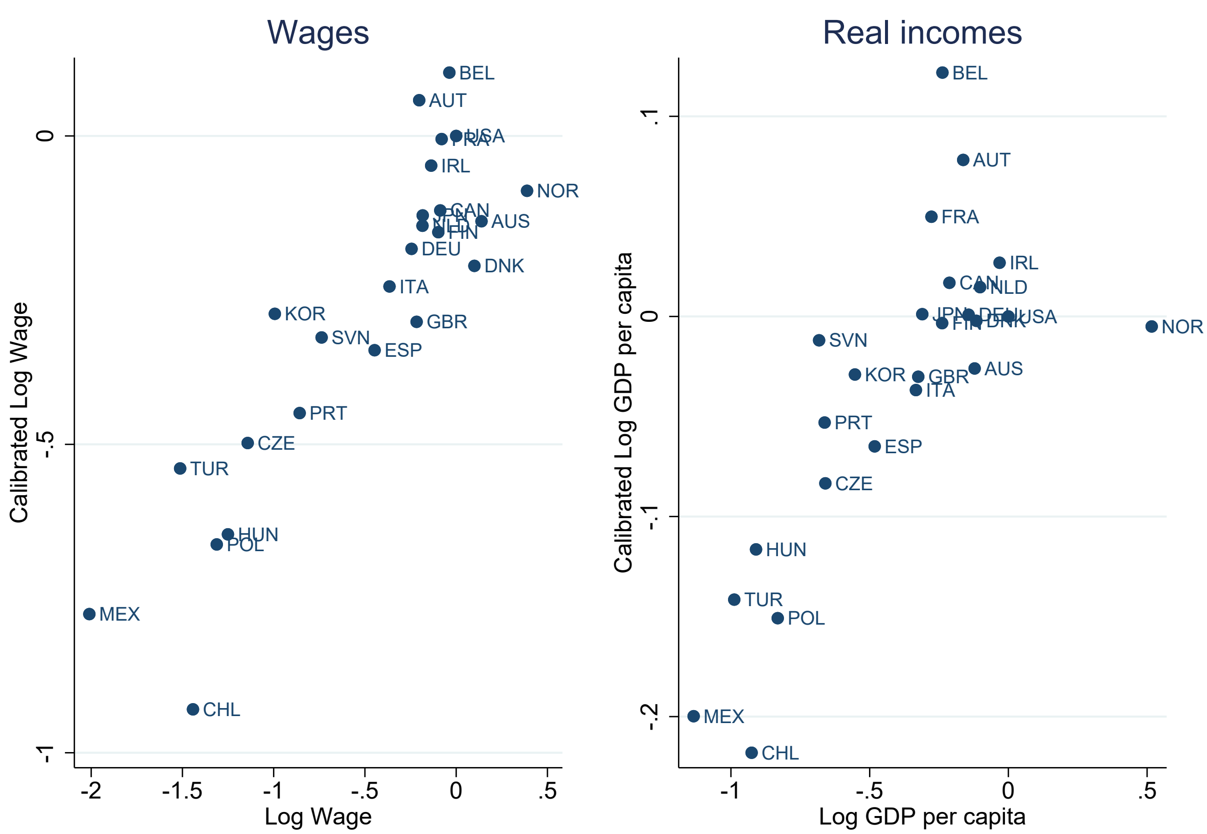

Figure 3: Calibrated Versus Observed Wages and Real Incomes

Notes: Model calibrated in 2012 using R&D data to measure R&D efficiency. Observed nominal wages in 2012 from Penn World Tables 9.0. Observed real GDP per capita, defined as GDP per member of the working age population, calculated from Penn World Tables 9.0 and the World Development Indicators. Variables normalized to zero for the US.

Firm-level R&D investment. The second validation exercise analyzes cross-industry variation in the intensive and extensive margins of firm-level R&D investment. Let $ShRD_{js}$ denote the share of firms that perform R&D and $FiRD_j$ be firm-level R&D intensity conditional on performing R&D. All else equal, the model predicts that $ShRD_{js}$, $FiRD_j$ and innovation-dependence ID_j are each decreasing in the advantage of backwardness γ_j.[28] Consistent with this prediction, $FiRD_j$ and $ShRD_{js}$ calculated from UK data are positively correlated with the innovation-dependence estimates in columns (c) and (d) of Table 1; the correlations range between 0.32 and 0.48.

The model also delivers a prediction for the functional form relationship between $FiRD_j$ and $ShRD_{js}$. Equations (11) and (14) together yield:

$$\frac{1}{FiRD_j} = -\frac{1}{\alpha k \log \eta} \log ShRD_{js} - \frac{\log\left(B^A/B_s\right)}{\alpha \log \eta} + \frac{\rho + \zeta}{\alpha(\delta + g_j)}. \tag{34}$$

This equation implies that the inverse of $FiRD_j$ is linearly decreasing in $\log ShRD_{js}$. Figure 4 plots $1/FiRD_j$ against negative $\log ShRD_{jUK}$ using UK data. A linear relationship fits the data well for most industries, as shown by the solid line, which plots the line of best fit excluding two outliers (industries 0103 and 17). Moreover, departures from linearity are negatively correlated with industry growth rates g_j as predicted by equation (34).[29] These results show that the model is consistent with the observed variation in $FiRD_j$ and $ShRD_{js}$ in the UK.

Out-of-sample comparative advantage test. The final validation exercise uses the innovation-dependence estimates from Section 3.4 to perform an out-of-sample test of the proposition that countries with higher R&D efficiency have a comparative advantage in more innovation-dependent industries. The baseline estimation covered 25 countries for which R&D efficiency could be computed from OECD data. To conduct the out-of-sample test I use Eurostat data to calculate R&D efficiency for an additional nine European countries (see Appendix C for details). I then estimate the following variant of equation (33):

$$\log\left(\frac{EX_{js\tilde{s}}}{EX_{j\tilde{s}\tilde{s}}}\right) - (\sigma - 1)\log\left(\frac{w_{\tilde{s}}}{w_s}\right) = \xi CompAdv_{j\tilde{s}} + Controls_{js\tilde{s}} + \epsilon_{js\tilde{s}}, \tag{35}$$

where $\tilde{s}$ indexes the nine out-of-sample countries, s indexes their trading partners, $CompAdv_{j\tilde{s}} = -\frac{ID_j}{k} b_{\tilde{s}}$ equals the interaction of R&D efficiency and innovation-dependence and $Controls_{js\tilde{s}}$ denotes the same set of trade cost, productivity and comparative advantage controls included in columns (c) and (d) of Table 1.

Table 2 reports the results from estimating (35) using pooled trade data from 2010-14 for 20

[28] For $\psi^{\max} \to \infty$, $ShRD_{js} = \left(\psi^*_{js}\right)^{-k} = \eta^{-k\gamma_j}\left(B_s/B^A\right)^k$, which is decreasing in γ_j since $\eta > 1$. $FiRD_j$ is decreasing in γ_j by equation (14).

[29] A robust regression of $\frac{1}{FiRD_j}$ on $-\log ShRD_{jUK}$ and g_j yields a positive coefficient on $-\log ShRD_{jUK}$ with p-value 0.00 and a negative coefficient on g_j with p-value 0.13 (see Appendix C for details on how industry growth rates g_j are calculated). As the slope of the relationship between $\frac{1}{FiRD_j}$ and $-\log ShRD_{jUK}$ depends upon the product $\alpha k \log \eta$, it could be used to jointly calibrate the returns to scale in technology investment, dispersion in R&D capability across firms and the extent to which existing knowledge is more useful for adoption than innovation. However, the quantitative analysis does not require calibrating this bundle of parameters.

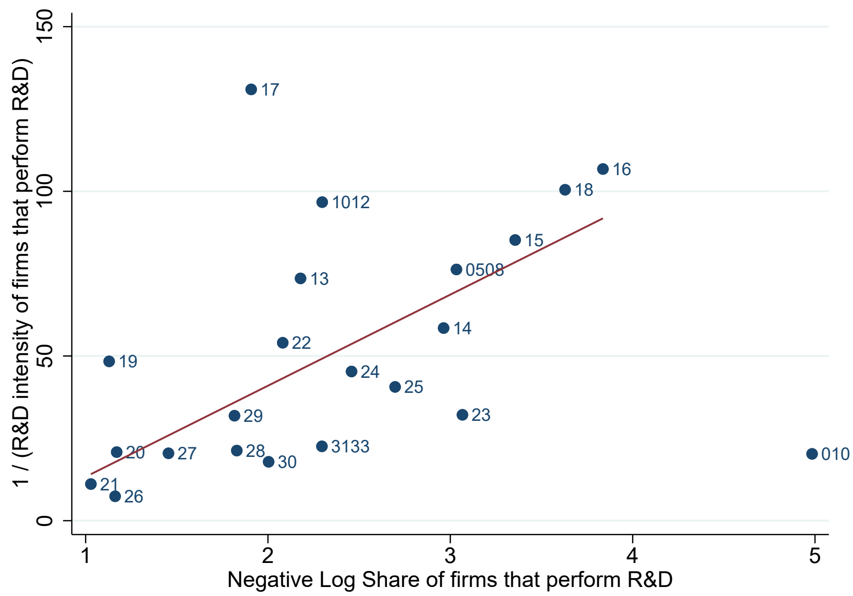

Figure 4: Firm-level R&D investment

Notes: Share of firms that invest in R&D and median R&D intensity of firms with positive R&D investment computed for UK industries from Office for National Statistics' Annual Business Survey and Business Expenditure on Research and Development data set. Both variables calculated for 2008-09 using 2 digit ISIC Revision 4 goods industries. Solid line shows predicted relationship from linear regression excluding industries 0103 and 17.

ISIC 2 digit manufacturing industries and 117 partner countries. In column (a), $CompAdv_{j\tilde{s}}$ is calculated using $b^R_{\tilde{s}}$ and the innovation-dependence estimates from column (c) of Table 1, while column (b) uses $b^P_{\tilde{s}}$ and the corresponding innovation-dependence estimates. The model predicts the coefficient ξ of $CompAdv_{j\tilde{s}}$ should be positive and equal to the trade elasticity. The estimated value of ξ equals 5.3 in column (a) and 13.1 in column (b). Subject to the caveats that the estimation only uses R&D efficiency for nine countries, and that the estimated trade elasticity differs across specifications, these results show that the relationship between R&D efficiency and comparative advantage that exists in the baseline sample is also present out-of-sample. This supports the proposition that countries with higher R&D efficiency have a comparative advantage in more innovation-dependent industries.

3.6 Counterfactual Analysis

The goal of the counterfactual analysis is to quantify the extent to which R&D efficiency differences explain cross-country variation in comparative advantage, wages and incomes. For this purpose, I compare the calibrated equilibrium to a counterfactual economy where R&D efficiency is the same in all 25 sample countries, but the other calibrated parameters are unchanged.[30] For any variable x, let x^o be the observed value of x in the data and x^c be the difference between the value of x in the counterfactual economy and in the calibrated model. The counterfactual change x^c quantifies the effect of eliminating R&D efficiency differences on x.

R&D efficiency differences are quantitatively important in determining comparative advantage. Let CA_{js}^c be the counterfactual change in the comparative advantage of country s relative to the US defined by:

$$CA_{js}^c = -(\sigma - 1)\log B_s \left(ID_j - \frac{1}{J-1}\sum_j ID_j \right),$$

where the summation only includes tradable industries. This definition implies that the cross-industry average of each country's CA_{js}^c equals zero. Moreover, since log exports satisfy equation (31), CA_{js}^c maps one-to-one into counterfactual changes in log exports (defined relative to the US and to the average change across industries in country s).

Figure 5 plots the cross-country average of CA_{js}^c by industry for the calibration using R&D data (similar results are obtained with patent data). Industries are ordered with innovation-dependence increasing from left to right. Because the US has higher than sample average R&D efficiency, the counterfactual mechanically decreases average comparative advantage relative to the US in low innovation-dependence industries and increases average comparative advantage relative to the US in high innovation-dependence industries. These effects are evident in Figure 5.

More importantly, the figure shows that eliminating differences in R&D efficiency has large effects on comparative advantage and trade. At one extreme, average comparative advantage in Computers relative to the US increases by 143 log points, implying that average exports of Computers (relative to the US and compared to the average industry) rise by the same amount. Likewise, in the counterfactual with no R&D efficiency differences, average exports relative to the US increase by 94 log points more for Chemicals at the 90th percentile of the innovation-dependence distribution, than for Agriculture at the 10th percentile.

Next, consider wages. The left hand panel of Figure 6 plots the counterfactual change in nominal log wages $\log w_s^c$ against observed log wages $\log w_s^o$ for the R&D calibration. Both variables are normalized to zero for the US. Eliminating R&D efficiency differences raises relative wages in countries with lower R&D efficiency and the figure shows that this results in higher wage gains for countries with smaller observed wages. The counterfactual wage changes are economically significant. For example, compared to the calibrated equilibrium, wages relative to the US increase

[30] The counterfactual analysis holds productivity growth g_j and, consequently, firm-level R&D intensity $FiRD_j$ constant. This is equivalent to assuming that any growth effects of changes in R&D efficiency are offset by variation in other parameters that impact growth, but do not effect the calibrated equilibrium conditional on $FiRD_j$, for example changes in the strength of R&D spillovers.

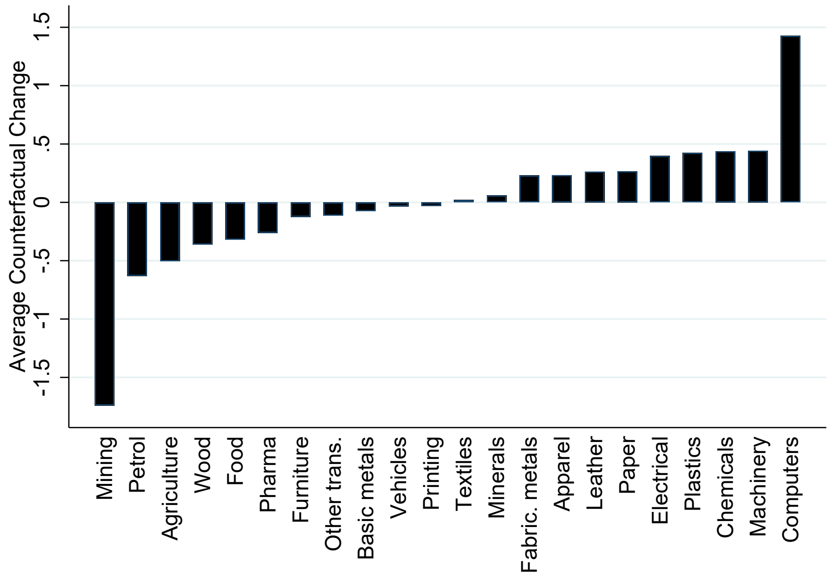

Figure 5: Counterfactual Change in Comparative Advantage

Notes: Average counterfactual change in comparative advantage by industry from author's calculations. Sample includes 25 OECD countries. Comparative advantage defined relative to US. Counterfactual sets R&D efficiency equal across countries. Model calibrated in 2012 using R&D data to measure R&D efficiency. Industries are 2 digit ISIC Revision 4 industries.

by 56% for Turkey, 23% for Spain and 15% for Canada. On average, eliminating R&D efficiency differences raises wages relative to the US by 18 log points or 22% (see Table 3, column a). For comparison, the average observed wage gap relative to the US is 52 log points.

What determines the magnitude of the counterfactual wage changes? Section 2.4 showed that, in a single sector economy with free trade, the elasticity of relative wages to R&D efficiency equals $(\sigma - 1)/\sigma$ times innovation-dependence. The counterfactual results imply that this finding generalizes to an economy with many sectors. Let $Q_s \equiv -\frac{\sigma-1}{\sigma} b_s \sum_j \mu_j ID_j$, where the summation computes the expenditure share weighted average innovation-dependence across tradable industries. Q_s accounts for essentially all of the variation in counterfactual wage changes. The correlation between $\log w_s^c$ and Q_s is 0.99 and the R-squared from regressing $\log w_s^c$ on Q_s equals 0.98. It follows that the magnitude of wage gaps due to R&D efficiency differences is determined by the extent to which R&D efficiency varies across countries together with the average level of innovation-dependence.

By eliminating wage differences due to variation in R&D efficiency, the counterfactual reduces wage dispersion. The standard deviation of model-implied log wages is 55% lower in the counterfactual economy than in the calibrated equilibrium. To quantify the contribution of R&D efficiency differences to international wage inequality, I follow the development accounting literature and

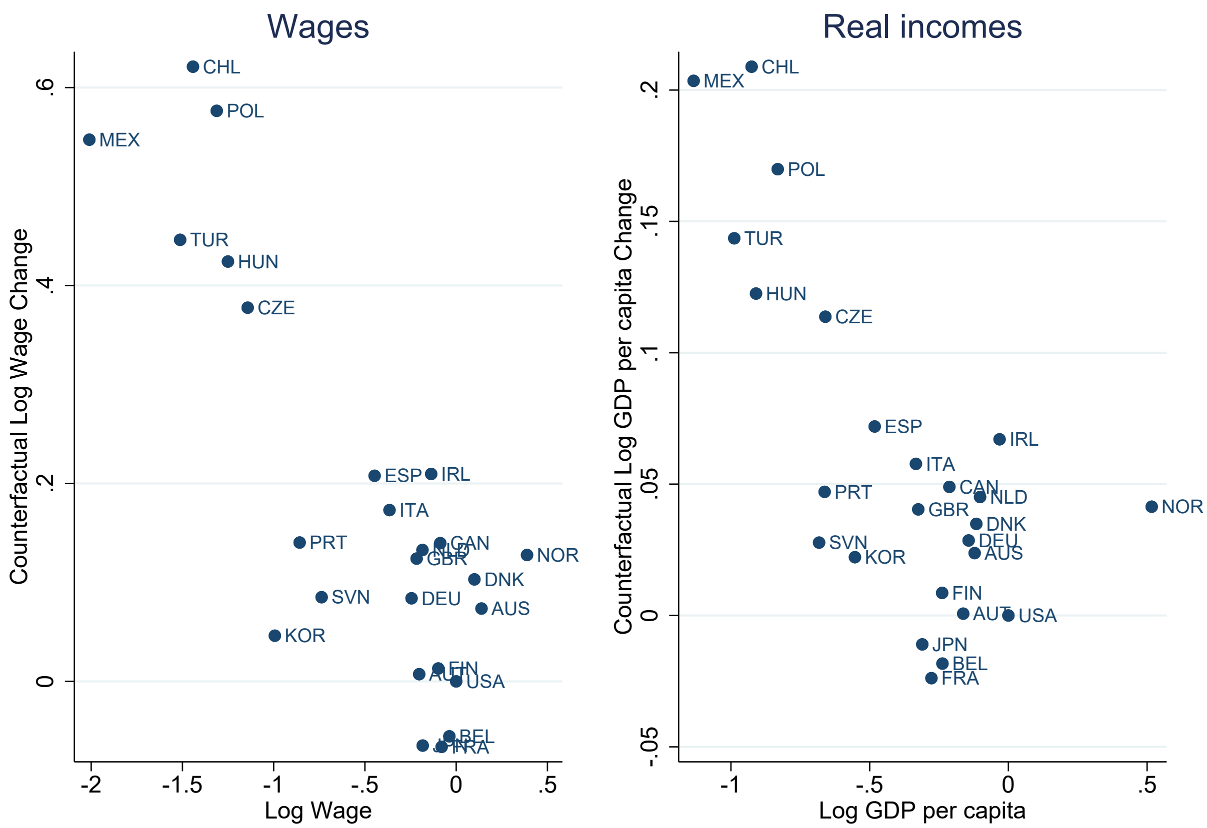

Figure 6: Counterfactual Log Wage and Real Income Changes

Notes: Counterfactual changes calculated by setting R&D efficiency equal across countries. Observed nominal wages in 2012 from Penn World Tables 9.0. Observed real GDP per capita, defined as GDP per member of the working age population, calculated from Penn World Tables 9.0 and the World Development Indicators. Variables normalized to zero for the US. Model calibrated in 2012 using R&D data to measure R&D efficiency.

compare wage gaps caused by variation in R&D efficiency with observed wage dispersion. Specifically, I compute the ratio of the standard deviation of the log wage change caused by eliminating R&D efficiency differences to the standard deviation of observed log wages:

$$\text{Wage dispersion ratio } = \frac{StdDev\left(\log w_s^c\right)}{StdDev\left(\log w_s^o\right)}.$$

The wage dispersion ratio measures the share of observed wage differences that can be explained by variation in R&D efficiency. Column (a) of Table 3 reports that the wage dispersion ratio equals 0.32.[31] Thus, R&D efficiency differences account for just under one-third of observed

[31]The standard deviation of log wages is used to measure dispersion because Figure 6 shows a linear relationship between $\log w_s^c$ and $\log w_s^o$. Suppose $\log w_s^c = -\xi^w \log w_s^o$. Then the wage dispersion ratio equals ξ^w. Indeed, regressions of $\log w_s^c$ on negative $\log w_s^o$ give coefficient estimates close to the wage dispersion ratios in Table 3.

wage dispersion within the OECD. This result implies that technology gaps are quantitatively important in explaining wage gaps within the OECD, but also that other sources of cross-country heterogeneity, such as factor endowment differences and misallocation, account for the majority of wage differences.

Wages are the primary outcome of interest for the reasons discussed in Section 3.2. However, I also calculate the counterfactual changes in real income per capita when R&D efficiency is equalized across countries. The right hand panel of Figure 6 plots the counterfactual change $\log GDPPC_s^c$ against observed log income per capita $\log GDPPC_s^o$. And column (a) of Table 3 reports the average counterfactual change in income per capita relative to the US and the income dispersion ratio, defined as the standard deviation of $\log GDPPC_s^c$ relative to the standard deviation of $\log GDPPC_s^o$.

The counterfactual increases real incomes relative to the US in poorer countries, leading to a decline in real income dispersion. But as expected, given the assumption that the innovation-dependence of services is zero, R&D efficiency accounts for a smaller share of variation in real incomes than in nominal wages. Average real income per capita relative to the US increases by 5.9 log points, compared to an average observed income per capita gap of 40 log points. The income dispersion ratio equals 0.17, implying that technology gaps due to R&D efficiency differences explain around one-sixth of real income dispersion in the OECD.[32]

When interpreting these results, it is worth noting that the share of wage and income variation accounted for by R&D efficiency differences may depend upon the sample of countries studied. In particular, R&D efficiency differences may be more important between countries at different stages of development than within groups of similar countries. In Figure 6 there are six countries that benefit noticeably more from the elimination of R&D efficiency differences than the rest of the OECD.[33] Dropping these countries when calculating the wage and income dispersion ratios yields a wage dispersion ration of 0.25 and an income dispersion ratio of 0.10. Thus, R&D efficiency differences explain a slightly lower share of wage and income variation within higher income OECD countries than in the full sample, but remain quantitatively important.

It is also informative to compare the baseline results with an alternative counterfactual that eliminates goods trade by assuming all industries are non-tradable, but does not change R&D efficiency levels. In this case nominal wages are not comparable across countries, but counterfactual real incomes are still of interest. Eliminating trade reduces average real income per capita relative to the US by 2.7 log points, with more open economies experiencing larger relative income de-

Using the Gini coefficient as an alternative wage dispersion measure and taking the ratio of the counterfactual to observed Gini coefficients gives a wage dispersion ratio of 0.42.

[32] When the Gini coefficient is used to measure income dispersion, the income dispersion ratio equals 0.18.

[33] These six countries are Chile, Czech Republic, Hungary, Mexico, Poland and Turkey. They are the poorest sample countries (as measured by observed nominal wages) and also have the lowest calibrated R&D efficiency levels.

clines. The income dispersion ratio for this autarky counterfactual is 0.080. These results imply that goods trade plays a less important role than R&D efficiency differences in explaining income variation within the OECD.

Robustness. The baseline results in column (a) of Table 3 are robust to a series of alternative calibrations. Column (b) of Table 3 reports the impact of eliminating R&D efficiency differences when data on patenting intensity is used to calibrate R&D efficiency. For the patent data calibration, the wage dispersion ratio is 0.27 and the income dispersion ratio is 0.13. Comparing column (b) to column (a) shows that the counterfactual changes are similar regardless of whether R&D or patent data is used to calibrate the model. Indeed, the correlation of $\log w_s^c$ for the two calibrations is 0.86, while the correlation of $\log GDPPC_s^c$ is 0.87. Given the challenges inherent in measuring innovation, this consistency alleviates potential concerns that the results are driven by measurement error in R&D or patent data.

Appendix D.2 describes a series of additional robustness checks that: (i) modify how innovation-dependence is estimated; (ii) reduce the elasticity of patenting to R&D expenditure in the patent data calibration, and; (iii) vary the trade elasticity $\sigma - 1$ between 2.5 and 8.5 and allow for cross-industry heterogeneity in trade elasticities using estimates from Caliendo and Parro (2015).

These alternative calibrations do not make a substantial difference to the quantitative results (see Table A1). For example, increasing the trade elasticity dampens counterfactual changes in wages and incomes. However, the difference is small because of two countervailing effects. On the one hand, an increase in σ reduces estimated innovation-dependence by inflating the size of the independent variable $(\sigma - 1)\, b_{\bar{s}}$ in the exports equation (33). On the other hand, increasing σ raises the elasticity of relative wages and incomes to R&D efficiency conditional on innovation-dependence, which depends upon $(\sigma - 1)/\sigma$ as shown in Section 2.4. Overall, the robustness checks reinforce the conclusion that R&D efficiency accounts for an economically significant fraction of wage and income variation within the OECD.

Finally, Appendix D.3 studies how using a first order approximation to the model affects the counterfactual results. The appendix shows that the approximation reduces cross-country inequality due to R&D efficiency differences and computes an upper bound on the size of the approximation error in the baseline R&D calibration. Performing the counterfactual analysis without approximating the model increases the wage dispersion ratio from 0.32 in column (a) of Table 3 to at most 0.36, and the income dispersion ratio from 0.17 to at most 0.19. These comparisons imply that the approximation error is small.

4 Generalizations

This section analyzes how relaxing some of the simplifying assumptions made in the baseline model affects the quantitative results. It starts by generalizing how national innovation systems affect the R&D and adoption technologies, and then introduces inter-industry knowledge spillovers.

4.1 R&D and Adoption Technologies

In the baseline model, R&D efficiency is homogeneous across industries and the efficiency of technology adoption is the same for all firms and countries. Relaxing these assumptions allows for cross-industry heterogeneity in how national innovation systems determine R&D efficiency and for the possibility that more innovative firms and countries also have an advantage in adoption. Suppose industry-level R&D efficiency is given by $B_{js} = B_s^{\nu_{0j}}$ where $\nu_{0j} > 0$ is the elasticity of B_{js} to country-level R&D efficiency B_s. National innovation systems matter more in industries with higher ν_{0j}. In addition, assume that instead of equation (7), the adoption technology is given by:

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

where $\nu_{1j}, \nu_{2j} \in [0, 1)$. This specification allows the returns to adoption to be greater, all else equal, for firms with higher R&D capability ψ and countries with higher R&D efficiency B_s. However, since ν_{1j} and ν_{2j} are below one, the efficiency of R&D relative to adoption is increasing in ψ and B_s as in the baseline model.

Appendix B.1 solves the generalized model and shows that the structure of the balanced growth path equilibrium is unchanged. Moreover, Propositions 1, 2 and 3 continue to hold. It also shows that, conditional on B_s, the calibration strategy developed in Section 3 remains valid. In particular, B_s and ID_j are still sufficient statistics for quantifying international technology gaps. And although innovation-dependence ID_j depends upon ν_{0j} and ν_{2j} in the generalized model, it can be calibrated from trade data as before.

However, the generalization does affect the calibration of R&D efficiency B_s. Cross-industry variation in ν_{0j}, ν_{1j} and ν_{2j} means that the ratio of industry-level R&D intensities in any county pair differs by industry. Consequently, a double differences approach is required to calibrate R&D efficiency (see Appendix D.4 for details). Log R&D efficiency calibrated from R&D data using this approach has a correlation of 0.99 with b_s^R and 0.88 with b_s^P.

The counterfactual results for the generalized model are reported in column (c) of Table 3. Eliminating R&D efficiency differences has similar effects on wages and real incomes as in the baseline R&D calibration. The wage dispersion ratio in column (c) is 0.31 and the income dispersion ratio is 0.16. These findings demonstrate that the baseline quantitative results are robust to introducing greater flexibility in how national innovation systems and firm capabilities affect the returns to technology investment.

Both the baseline model and this generalization assume that technology gaps depend upon a single dimension of cross-country heterogeneity B_s. An interesting avenue for future work would be to allow for country-specific adoption efficiency B_s^A and estimate the correlation between adoption efficiency and R&D efficiency. Implementing this idea would require using an additional observable moment to calibrate B_s^A.

4.2 Inter-industry Spillovers

In the baseline model, knowledge spillovers occur exclusively within industries. To introduce inter-industry spillovers, assume that, instead of equation (4), the knowledge level χ_{js}^R satisfies:

$$\chi_{js}^R = \chi_j \left[\prod_{i=1}^{J} \left(\frac{\theta_{is}^{\max}}{\chi_{is}^R} \right)^{d_{ij}} \right]^{\kappa_j}, \qquad \text{with } \sum_{i=1}^{J} d_{ij} = 1. \tag{36}$$

This expression implies that domestic spillovers are a weighted average across industries of the ratio of frontier productivity $\theta_{is}^{\max}$ to R&D knowledge χ_{is}^R. The parameter $d_{ij} \in [0,1]$ determines the strength of domestic knowledge spillovers from industry i to industry j. Note that setting $d_{jj} = 1$ and $d_{ij} = 0$ for $i \neq j$ gives the baseline specification with only intra-industry spillovers.[34]

Appendix B.2 solves this extension of the model and shows that the calibration and counterfactual analysis in Section 3 are unchanged, meaning that the quantitative results in Table 3 continue to hold even with inter-industry spillovers. However, inter-industry spillovers do affect the determinants of innovation-dependence, which is now given by:

$$ID_j = \kappa_j \sum_{i=1}^{J} \frac{d_{ij}(1-\beta)}{\gamma_i(1-\beta) - \alpha}. \tag{37}$$

The key difference from the baseline model (compare equation 37 with equation 26) is that the innovation-dependence of industry j now depends upon the advantage of backwardness in all other industries, with weights given by the spillover parameters d_{ij}. Innovation-dependence is high in industries that receive strong spillovers from industries with a low advantage of backwardness.[35]

The conclusion that inter-industry spillovers do not affect the baseline results holds for the

[34] See Huang and Zenou (2020) for a related formulation of inter-industry spillovers in a version of the Romer (1990) growth model.

[35] Appendix B.2 also discusses the case where inter-industry spillovers impact the growth of global knowledge capital χ_j. Global inter-industry spillovers affect equilibrium growth rates g_j, but leave innovation-dependence levels, equilibrium technology gaps, and the quantitative results unchanged.

specification of spillovers in equation (36). Future research should provide additional evidence on the nature of cross-border and cross-industry spillovers and facilitate a better understanding of how the localization of knowledge spillovers, the advantage of backwardness, and the strength of inter-industry spillovers determine innovation-dependence levels.

5 Conclusions

Understanding the origins of cross-country differences in income per capita is one of the central concerns of economics. The extent to which the geography of innovation contributes to international inequality depends upon the rate at which technologies diffuse across national borders. When diffusion is fast, the technology gap between innovators and imitators is small, whereas slow diffusion increases the advantage that accrues to knowledge creators. Yet evidence on the quantitative importance of technology gaps is scarce.

By building a quantifiable model of innovation and adoption in open economies, this paper develops a new methodology to estimate the importance of international technology gaps in explaining cross-country income differences. Rather than treating productivity as a residual, the paper shows how the size of technology gaps can be inferred from data on innovation intensity and bilateral trade. More innovative countries have a comparative advantage in more innovation-dependent industries, and, when innovation-dependence is higher, international wage and income gaps due to differences in R&D efficiency are larger.

Counterfactual analysis implies that eliminating R&D efficiency differences within the OECD would increase wages relative to the US by around 20% for the average country. Moreover, R&D efficiency accounts for one-quarter to one-third of nominal wage dispersion and approximately 15% of real income per capita dispersion in the OECD. The empirical analysis also finds that there is substantial heterogeneity in innovation-dependence across industries, implying that differences in R&D efficiency are an important source of Ricardian comparative advantage.

While the quantitative analysis in this paper focuses on cross-sectional technology gaps, the modelling framework has many potential applications. It could be used to decompose the sources of growth in open economies with endogenous innovation and adoption, to characterize transition dynamics following changes in R&D efficiency, or to estimate changes in technology gaps over time. For example, the analysis could be adapted to study whether recent advances in information and communication technologies have promoted global convergence by shrinking international technology gaps. Addressing such questions would provide further insight into how access to technologies shapes global economic outcomes.

References

Acemoglu, Daron, and Fabrizio Zilibotti. 2001. "Productivity Differences." *Quarterly Journal of Economics* 116 (2): 563-606.

Akcigit, Ufuk, Sina Ates, and Giammario Impullitti. 2018. "Innovation and Trade Policy in a Globalized World." NBER Working Paper No. 24543.

Akcigit, Ufuk, and William Kerr. 2018. "Growth through Heterogeneous Innovations." *Journal of Political Economy* 126 (4): 1374-1443.

Alviarez, Vanessa, Javier Cravino, and Natalia Ramondo. 2021. "Firm-embedded Productivity and Cross-country Income Differences." mimeo, Boston University.

Armington, Paul S. 1969. "A Theory of Demand for Products Distinguished by Place of Production." *IMF Economic Review* 16 (1): 159-178.

Atkeson, Andrew, and Ariel Tomas Burstein. 2010. "Innovation, Firm Dynamics, and International Trade." *Journal of Political Economy* 118 (3): 433-484.

Baldwin, Richard E., and Frédéric Robert-Nicoud. 2008. "Trade and Growth with Heterogeneous Firms." *Journal of International Economics* 74 (1): 21-34.

Barkai, Simcha. 2017. "Declining Labor and Capital Shares." mimeo, London Business School.

Bartelsman, Eric J., Jonathan Haskel, and Ralf Martin. 2008. "Distance to Which Frontier? Evidence on Productivity Convergence from International Firm-level Data." CEPR Discussion Paper No. 7032.

Benhabib, Jess, Jesse Perla, and Christopher Tonetti. 2014. "Catch-up and Fall-back through Innovation and Imitation." *Journal of Economic Growth* 19 (1): 1-35.

Bonfiglioli, Alessandra, Rosario Crinò, and Gino A. Gancia. 2018. "Betting on Exports: Trade and Endogenous Heterogeneity." *Economic Journal* 128 (609): 612-651.

Branstetter, Lee G. 2001. "Are Knowledge Spillovers International or Intranational in Scope?: Microeconometric Evidence from the US and Japan." *Journal of International Economics* 53 (1): 53-79.

Buera, Francisco J., and Ezra Oberfield. 2020. "The Global Diffusion of Ideas." *Econometrica* 88 (1): 83-114.

Cai, Jie, Nan Li, and Ana Maria Santacreu. 2022. "Knowledge Diffusion, Trade, and Innovation across Countries and Sectors." *American Economic Journal: Macroeconomics* 14 (1): 104-45.

Caliendo, Lorenzo, and Fernando Parro. 2015. "Estimates of the Trade and Welfare Effects of NAFTA." *Review of Economic Studies* 82 (1): 1-44.

Caselli, Francesco, and Wilbur John Coleman. 2001. "Cross-Country Technology Diffusion: The Case of Computers." *American Economic Review Papers and Proceedings* 91 (2): 328-335.

Caselli, Francesco. 2005. "Accounting for Cross-country Income Differences." *Handbook of*

Economic Growth 1: 679-741.

Comin, Diego, Bart Hobijn, and Emilie Rovito. 2008. "Technology Usage Lags." *Journal of Economic Growth* 13 (4): 237-256.

Comin, Diego, and Martí Mestieri. 2018. "If Technology has Arrived Everywhere, Why has Income Diverged?" *American Economic Journal: Macroeconomics* 10 (3): 137-78.

Costinot, Arnaud. 2009. "On the Origins of Comparative Advantage." *Journal of International Economics* 77 (2): 255-264.

Costinot, Arnaud, Dave Donaldson, and Ivana Komunjer. 2012. "What Goods do Countries Trade? A Quantitative Exploration of Ricardo's Ideas." *Review of Economic Studies* 79 (2): 581-608.

Doraszelski, Ulrich, and Jordi Jaumandreu. 2013. "R&D and Productivity: Estimating Endogenous Productivity." *Review of Economic Studies* 80 (4): 1338-1383.

Eaton, Jonathan, and Samuel Kortum. 1999. "International Technology Diffusion: Theory and Measurement." *International Economic Review* 40 (3): 537-570.

Eaton, Jonathan, and Samuel Kortum. 2001. "Technology, Trade, and Growth: A Unified Framework." *European Economic Review* 45 (4): 742-755.

Eaton, Jonathan, and Samuel Kortum. 2002. "Technology, Geography and Trade." *Econometrica* 70 (5): 1741-1779.

Evenson, Robert E., and Larry E. Westphal. 1995. "Technological Change and Technology Strategy." *Handbook of Development Economics* 3: 2209-2299.

Gabler, Alain, and Markus Poschke. 2013. "Experimentation by Firms, Distortions, and Aggregate Productivity." *Review of Economic Dynamics* 16 (1): 26-38.

Gancia, Gino, Andreas Müller, and Fabrizio Zilibotti. 2013. "Structural Development Accounting." *Advances in Economics and Econometrics* 2: 373-418.

Garcia-Macia, Daniel, Chang-Tai Hsieh, and Peter J. Klenow. 2019 "How Destructive is Innovation?" *Econometrica* 87 (5): 1507-1541.

Gerschenkron, Alexander. 1962. "Economic Backwardness in Historical Perspective: A Book of Essays." Cambridge, MA: Belknap Press of Harvard University Press.

Griffith, Rachel, Stephen Redding, and John Van Reenen. 2004. "Mapping the Two Faces of R&D: Productivity Growth in a Panel of OECD Industries." *Review of Economics and Statistics* 86 (4): 883-895.

Griffith, Rachel, Stephen Redding, and Helen Simpson. 2009. "Technological Catch-up and Geographic Proximity." *Journal of Regional Science* 49 (4): 689-720.

Grossman, Gene M., and Elhanan Helpman. 1990. "Trade, Innovation, and Growth." *American Economic Review* 80 (2): 86-91.

Grossman, Gene M., and Elhanan Helpman. 1991. "Innovation and Growth in the Global Economy." The MIT Press.

Hanson, Gordon H., Nelson Lind, and Marc-Andreas Muendler. 2013. "The Dynamics of Comparative Advantage: Hyperspecialization and Evanescence." mimeo, University of California, San Diego.

Hsieh, Chang-Tai, Erik Hurst, Charles I. Jones, and Peter J. Klenow. "The Allocation of Talent and US Economic Growth." *Econometrica* 87 (5): 1439-1474.

Hsieh, Chang-Tai, and Peter J. Klenow. 2009. "Misallocation and Manufacturing TFP in China and India." *Quarterly Journal of Economics* 124 (4): 1403-1448.

Howitt, Peter. 2000. "Endogenous Growth and Cross-Country Income Differences." *American Economic Review* 90 (4): 829-46.

Huang, Jingong, and Yves Zenou. 2020. "Key Sectors in Endogeneous Growth." CEPR Discussion Paper No. 15281.

Keller, Wolfgang. 2002. "Geographic Localization of International Technology Diffusion." *American Economic Review* 92 (1): 120-142.

Klenow, Peter J., and Andrés Rodríguez-Clare. 2005. "Externalities and Growth." *Handbook of Economic Growth* 1: 817-861.

König, Michael, Jan Lorenz, and Fabrizio Zilibotti. 2016. "Innovation vs Imitation and the Evolution of Productivity Distributions," *Theoretical Economics* 11 (3): 1053-1102.

Krugman, Paul. 1979. "A Model of Innovation, Technology Transfer, and the World Distribution of Income." *Journal of Political Economy* 87 (2): 253-66.

Krugman, Paul. 1985. "A 'Technology Gap' Model of International Trade." in *Structural Adjustment in Developed Open Economies*, Palgrave Macmillan, London, 35-61.

Levchenko, Andrei A., and Jing Zhang. 2016. "The Evolution of Comparative Advantage: Measurement and Welfare Implications." *Journal of Monetary Economics* 78: 96-111.

Lind, Nelson, and Natalia Ramondo. 2022. "Global Innovation and Knowledge Diffusion." NBER Working Paper No. 29629.

Manova, Kalina. 2013. "Credit Constraints, Heterogeneous Firms, and International Trade." *Review of Economic Studies* 80 (2): 711-744.

Melitz, Marc J. 2003. "The Impact of Trade on Intra-industry Reallocations and Aggregate Industry Productivity." *Econometrica* 71 (6): 1695-1725.

Nelson, Richard R., ed. 1993. "National Innovation Systems: A Comparative Analysis." Oxford University Press.

Nunn, Nathan. 2007. "Relationship-specificity, Incomplete Contracts, and the Pattern of Trade." *Quarterly Journal of Economics* 122 (2): 569-600.

OECD. 2015. "Frascati Manual 2015: Guidelines for Collecting and Reporting Data on Research and Experimental Development, The Measurement of Scientific, Technological and Innovation Activities." OECD Publishing, Paris.

Parente, Stephen L., and Edward C. Prescott. 1994. "Barriers to Technology Adoption and Development." *Journal of Political Economy* 102 (2): 298-321.

Peri, Giovanni. 2005. "Determinants of Knowledge Flows and their Effect on Innovation." *Review of Economics and Statistics* 87 (2): 308-322.

Perla, Jesse, and Christopher Tonetti. 2014. "Equilibrium Imitation and Growth." *Journal of Political Economy* 122 (1): 52-76.

Perla, Jesse, Christopher Tonetti, and Michael E. Waugh. 2021. "Equilibrium Technology Diffusion, Trade, and Growth." *American Economic Review* 111 (1): 73-128.

Redding, Stephen. 1999. "Dynamic Comparative Advantage and the Welfare Effects of Trade." *Oxford Economic Papers* 51 (1): 15-39.

Romalis, John. 2004. "Factor Proportions and the Structure of Commodity Trade." *American Economic Review* 94 (1): 67-97.

Romer, Paul M. 1990. "Endogenous Technological Change." *Journal of Political Economy* 98: S71-S102.

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.

Sampson, Thomas. 2016a. "Dynamic Selection: An Idea Flows Theory of Entry, Trade, and Growth." *Quarterly Journal of Economics* 131 (1): 315-380.

Sampson, Thomas. 2016b. "Assignment Reversals: Trade, Skill Allocation and Wage Inequality." *Journal of Economic Theory* 163: 365-409.

Somale, Mariano. 2021. "Comparative Advantage in Innovation and Production." *American Economic Journal: Macroeconomics* 13 (3): 357-96.

Waugh, Michael E. 2010. "International Trade and Income Differences." *American Economic Review* 100 (5): 2093-2124.

Table 1: Innovation-dependence by industry

R&D efficiency measure	R&D intensity			Patenting intensity
	(a)	(b)	(c)	(d)
Agriculture, forestry and fishing (0103)	0.454	0.333	0.169	0.00742
	(0.0561)	(0.0468)	(0.0898)	(0.0621)
Mining and quarrying (0508)	0.371	0.250	-0.105	-0.140
	(0.0860)	(0.0658)	(0.133)	(0.0780)
Food products, beverages and tobacco (1012)	0.480	0.359	0.210	0.0608
	(0.0463)	(0.0434)	(0.0782)	(0.0579)
Textiles (13)	0.507	0.417	0.286	0.117
	(0.0465)	(0.0505)	(0.0640)	(0.0516)
Wearing apparel (14)	0.473	0.370	0.334	0.131
	(0.0566)	(0.0606)	(0.0547)	(0.0428)
Leather and related products (15)	0.479	0.386	0.340	0.122
	(0.0590)	(0.0679)	(0.0835)	(0.0699)
Wood and products of wood and cork, except furniture (16)	0.520	0.397	0.201	0.0282
	(0.0627)	(0.0394)	(0.0678)	(0.0451)
Paper and paper products (17)	0.580	0.451	0.341	0.127
	(0.0549)	(0.0388)	(0.0668)	(0.0535)
Printing and reproduction of recorded media (18)	0.579	0.460	0.274	0.109
	(0.0572)	(0.0383)	(0.0611)	(0.0396)
Coke and refined petroleum products (19)	0.479	0.359	0.141	0.0530
	(0.0452)	(0.0424)	(0.0782)	(0.0363)
Chemicals and chemical products (20)	0.587	0.474	0.379	0.189
	(0.0513)	(0.0519)	(0.0948)	(0.0576)
Basic pharmaceutical products and pharmaceutical preparations (21)	0.622	0.496	0.223	0.168
	(0.0735)	(0.0610)	(0.141)	(0.0964)
Rubber and plastics products (22)	0.603	0.478	0.376	0.183
	(0.0518)	(0.0362)	(0.0507)	(0.0373)
Other non-metallic mineral products (23)	0.568	0.447	0.295	0.118
	(0.0520)	(0.0384)	(0.0583)	(0.0384)
Basic metals (24)	0.577	0.425	0.265	0.178
	(0.0458)	(0.0487)	(0.0745)	(0.0338)
Fabricated metal products, except machinery and equipment (25)	0.598	0.475	0.333	0.138
	(0.0548)	(0.0360)	(0.0561)	(0.0387)
Computer, electronic and optical products (26)	0.653	0.487	0.599	0.295
	(0.0580)	(0.0408)	(0.119)	(0.0542)
Electrical equipment (27)	0.606	0.530	0.370	0.185
	(0.0871)	(0.0688)	(0.0958)	(0.0383)
Machinery and equipment n.e.c. (28)	0.712	0.599	0.380	0.213
	(0.0780)	(0.0522)	(0.107)	(0.0618)
Motor vehicles, trailers and semi-trailers (29)	0.552	0.387	0.274	0.188
	(0.0480)	(0.0363)	(0.0839)	(0.0328)
Other transport equipment (30)	0.563	0.380	0.256	-0.00330
	(0.0969)	(0.0633)	(0.128)	(0.0545)
Furniture, other manufacturing (3133)	0.547	0.424	0.254	0.103
	(0.0659)	(0.0397)	(0.0651)	(0.0520)
Observations	171,152	171,152	171,152	171,152
R-squared	0.524	0.652	0.697	0.694
Trade cost controls	Yes	Yes	Yes	Yes
Productivity level controls	No	Yes	Yes	Yes
Comparative advantage controls	No	No	Yes	Yes
F test innovation-dependence equal across industries	0.133	0.001	0.067	0.000
Average innovation-dependence	0.551	0.427	0.282	0.117

Innovation-dependence estimated relative to the shape parameter of the R&D capability distribution. Standard errors clustered by importer-industry in parentheses. R&D efficiency is measured from data on R&D intensity in columns (a)-(c) and patenting intensity in column (d). Trade cost controls are: exporter-industry fixed effects; interaction of industry dummy variables with six bilateral distance intervals, and; whether the countries share a border, common language or free trade agreement. Productivity level controls are the importer's rule of law, control of corruption, government effectiveness, political stability, regulatory quality, voice and accountability, ease of doing business and log private credit as a share of GDP. Comparative advantage controls are the interaction of industry dummy variables with the importer's rule of law, log private credit as a share of GDP, log physical capital per employee and human capital. Sample includes 25 importers and 117 exporters and uses data for 2010-14.

Table 2: Out-of-sample comparative advantage test

R&D efficiency measure	R&D intensity	Patenting intensity
	(a)	(b)
CompAdv	5.27	13.13
	(0.48)	(1.27)
Observations	31,996	31,996
R-squared	0.85	0.86
Trade cost controls	Yes	Yes
Productivity level controls	Yes	Yes
Comparative advantage controls	Yes	Yes

Dependent variable is bilateral imports relative to domestic trade adjusted for efficiency wage differences. CompAdv is interaction of country R&D efficiency with industry innovation-dependence. Standard errors clustered by importer-industry in parentheses. Column (a) uses measures of R&D efficiency and innovation-dependence from R&D intensity data. Column (b) uses measures of R&D efficiency and innovation-dependence from patenting intensity data. Trade cost controls are exporter-industry fixed effects and the interaction of industry dummy variables with six bilateral distance intervals and whether the countries share a border, a common language or a free trade agreement. Productivity level controls are the importer's rule of law, control of corruption, government effectiveness, political stability, regulatory quality, voice and accountability, ease of doing business and log private credit as a share of GDP. Comparative advantage controls are the interaction of industry dummy variables with the importer's rule of law, log private credit as a share of GDP, log physical capital per employee and human capital. Sample includes 9 importers and 117 exporters and uses data for 2010-14.

Table 3: Counterfactual results

R&D efficiency measure		R&D intensity	Patenting intensity	R&D intensity: Generalized model
		(a)	(b)	(c)
(i) Nominal wage	Average change relative to US	0.18	0.14	0.18
	Dispersion ratio	0.32	0.27	0.31
(ii) Real income per capita	Average change relative to US	0.059	0.042	0.059
	Dispersion ratio	0.17	0.13	0.16

Row (i) reports the average log wage change relative to the US between the counterfactual economy and the calibrated model, and the ratio of the standard deviation of the 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 for 2012. For column (a) the model is calibrated using R&D data. For column (b) the calibration uses patent data. Column (c) uses the generalized model in Section 5.1 calibrated with R&D data.

Online Appendix for
"Technology Gaps, Trade and Income"

Thomas Sampson

September 2022

A Proofs and Derivations

A.1 Equilibrium definition from Section 1.4

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

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

Solving the intertemporal optimization problem gives the Euler equation:

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

The transversality condition for intertemporal optimization in country s is:

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

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:

$$P_{js}X_{js} = \mu_j z_s c_s L_s, \tag{41}$$

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

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

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

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

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

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

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

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:

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

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:

$$Y_{js} = \sum_{\tilde{s}=1}^{S} \tau_{js\tilde{s}} x_{js\tilde{s}}. \tag{48}$$

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

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

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 $x_{j\tilde{s}s}$, 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 $s, \tilde{s} = 1, \ldots, S$ and all industries $j = 1, \ldots J$ 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 2

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 $\theta_{js}^{\max}$, 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.

[36]To 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.

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:

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

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:

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

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

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

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:

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

for all $s, \tilde{s} = 1, \ldots, S$. 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:

$$\frac{\dot{z}_s}{z_s} = -q, \tag{53}$$

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 $x_{js\tilde{s}}$ 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 2.1

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:

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

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:

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

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

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

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

$$0=\lim_{\tilde{t}\to\infty}\left[e^{-(\rho+\zeta)(\tilde{t}-t)}\mathcal{H}(\phi,l^R,\lambda)\right], \tag{56}$$

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

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

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

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

Equations (54) and (58) are an autonomous nonlinear system of differential equations in (ϕ, l^R) whose unique steady state $(\phi^*_{js}, l^{R*}_{js})$ 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:

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

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

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

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 $\dot{\lambda} \to 0$ 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 $\lim_{\tilde{t}\to\infty}\left[e^{-(\rho+\zeta)(\tilde{t}-t)}\lambda\phi\right] \geq 0$. Now define:

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

where the second line follows from solving the maximization problem in the first line. Assumption 1 implies $\overline{\mathcal{H}}(\phi, \lambda)$ 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 2.1. 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 follow 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:

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

where $\theta_{js}^* = \chi_{js}^R \left(\phi_{js}^*\right)^{1-\beta}$ 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 $V_{js}\left(\psi, \theta_{js}^*\right)$ is positive.

Section 2.1 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 $V_{js}\left(\psi, \theta_{js}^*\right)$ into the free entry condition (9) yields:

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

Next, observe that on a balanced growth path:

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

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:

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

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

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

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

$$z_s c_s = \rho a_s + w_s, \tag{62}$$

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

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

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 $j = 1, \ldots, J$ and all countries $s = 1, \ldots, S$.

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

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

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:

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

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 $\tau_{js\tilde{s}} = 1$ for all $j, s, \tilde{s}$ gives:

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

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

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

Suppose the growth rates g_j for $j = 1, \ldots, J$ 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 $s = 1, \ldots, S$ define the scaffold function $\mathbf{F} : \mathbb{R}^{S+1}_{++} \to \mathbb{R}^S$ by:

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

Note that $f_s\left(\mathbf{w}\right) = F_s\left(\mathbf{w}, w_s\right)$ 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 $\tilde{\mathbf{w}}$, $F_s\left(\tilde{\mathbf{w}}, w_s\right)$ 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 $F_s\left(\tilde{\mathbf{w}}, w_s\right)$ 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 $F_s\left[\lambda\tilde{\mathbf{w}}, w_s(\lambda)\right]=0$. Let $u\in(0,1)$ be such that $-1+\sigma u<0$. Then $F_s\left[\lambda\tilde{\mathbf{w}}, \lambda^{1-u}w_s(1)\right]$ is strictly negative if $\lambda>1$ and strictly positive if $\lambda<1$. Since $F_s\left(\tilde{\mathbf{w}}, w_s\right)$ 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 $\lambda\to\infty$, $\frac{\lambda}{w_s(\lambda)}\to\infty$ and when $\lambda\to 0$, $\frac{\lambda}{w_s(\lambda)}\to 0$ implying condition (iii) holds. Thus, a solution exists.

To prove uniqueness I use Theorem 2 in Allen, Arkolakis and Li (2015). Since $f_s\left(\mathbf{w}\right)$ is strictly increasing in $w_{\hat{s}}$ whenever $\hat{s}\neq s$, $\mathbf{f}\left(\mathbf{w}\right)$ satisfies gross substitution. Also, $f_s\left(\mathbf{w}\right)$ can be written as $f_s\left(\mathbf{w}\right)=\tilde{f}_s\left(\mathbf{w}\right)-L_s$ where $\tilde{f}_s\left(\mathbf{w}\right)$ 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:

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

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 $\theta_{js}^* = \chi_{js}^R \left(\phi_{js}^*\right)^{1-\beta}$ to obtain:

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

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:

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

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

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

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:

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

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

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

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

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

and differentiating (69) then implies:

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

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

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

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:

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

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

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

A.8 Derivation of balanced growth path consumption prices from Section 2.4

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:

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

$$\times\left[\frac{\alpha(\delta+g_j)^{\frac{\alpha-1}{\alpha}}}{\rho+\zeta+\gamma_j(\delta+g_j)}\right]^{\frac{-\alpha(1+\kappa_j)}{\gamma_j}}\frac{w_s}{\chi_j}\left(B_s\Psi_{js}^{\frac{\gamma_j(1-\beta)}{1+\kappa_j}-\alpha}\right)^{-\frac{1+\kappa_j}{\gamma_j}}.$$

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 3.1 and 3.4

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:

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

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

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

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

From equation (11) we have $\frac{\partial \log \psi_{js}^{*}}{\partial \log B_s} = -1$. Consequently, letting $\psi^{\max} \to \infty$ 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:

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

Letting $\psi^{\max} \to \infty$, 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:

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

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:

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

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 2. The main differences from the baseline model are as follows. The R&D threshold (11) is now given by:

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

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 $\psi^{\nu_{1j}} \left(\psi_{js}^*\right)^{1-\nu_{1j}}$. 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 $\psi^{\nu_{1j}}\left(\psi^*_{js}\right)^{1-\nu_{1j}}$ that chooses to invest in R&D. In addition, the average effective capability in industry j and country s is:

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

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:

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

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:

$$\begin{aligned}\log EX_{js\tilde{s}} = \upsilon^4_{j\tilde{s}} + (\sigma-1)\Bigg(&\frac{1+\kappa_j}{\gamma_j}\nu_{0j}\log B_s \\ &+\frac{\gamma_j(1-\beta)-\alpha(1+\kappa_j)}{\gamma_j}\log\Psi_{js} + \log A_{js} - \log w_s - \log\tau_{js\tilde{s}}\Bigg),\end{aligned}$$

where $\upsilon^4_{j\tilde{s}}$ 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:

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

and using this approximation to calculate innovation-dependence yields:

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

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:

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

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 3.4. 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:

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

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:

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

which implies that equation (19) becomes:

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

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

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

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 3 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:

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

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:

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

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:

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

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

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

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 3.6

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 (c) and (d) of Table 1, except in the importer fixed effects calibrations in columns (c) and (d) 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 (a) reports the results when R&D data is used to measure R&D efficiency, while patent data is used in column (b). 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 (c) 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 (d) 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 (e) and (f) show the counterfactual results differ little from the baseline results in Table 3.

The baseline patent data calibration in column (b) 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 (g) 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 (h) 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 (i)-(p) 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 (i) and (j) reduce the trade elasticity to 2.5 for the R&D intensity and patenting intensity calibrations, respectively. Columns (k) and (l) use an elasticity of 4.5, which is close to the aggregate elasticity estimated by Caliendo and Parro (2015). Columns (m) and (n) 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 (o) and (p) use the 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:

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

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:

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

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:

$$Z_{js} = \sum_{\tilde{s}=1}^{S} \frac{\tau_{js\tilde{s}}^{1-\sigma} \left(\rho a_{\tilde{s}} + w_{\tilde{s}}\right) L_{\tilde{s}} w_s^{-\sigma} B_s^{(\sigma-1)ID_j} E_{js}^{(\sigma-1)(1-\beta)}}{\sum_{\hat{s}=1}^{S} \tau_{j\hat{s}\tilde{s}}^{1-\sigma} w_{\hat{s}}^{1-\sigma} B_{\hat{s}}^{(\sigma-1)ID_j} E_{j\hat{s}}^{(\sigma-1)(1-\beta)}}. \tag{75}$$

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:

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

where the first equality is obtained by substituting equation (11) into equation (73), and the second equality uses $ShRD_{js} = \eta^{-k\gamma_j}\left(B_s/B^A\right)^k$, 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 (h) 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 4.1

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:

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

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 $b_{\tilde{s}}^G = 0$. 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 3.4. 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 Demirgüc-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.ggdc.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: Counterfacutal robustness checks

Robustness check		GDP per capita		Importer fixed effects		Significant innovation-dependence estimates		Patenting elasticity = 0.5	Approximation error
R&D efficiency measure		R&D intensity	Patenting intensity	R&D intensity	Patenting intensity	R&D intensity	Patenting intensity	Patenting intensity	R&D intensity
		(a)	(b)	(c)	(d)	(e)	(f)	(g)	(h)
(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 efficiency measure		R&D intensity	Patenting intensity	R&D intensity	Patenting intensity	R&D intensity	Patenting intensity	R&D intensity	Patenting intensity
		(i)	(j)	(k)	(l)	(m)	(n)	(o)	(p)
(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 (a) and (b) 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 (c) and (d) 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 (e) and (f) all innovation-dependence estimates that are insignificant at the 10 percent level are set equal to zero. For column (g) R&D efficiency is calculated from patenting data assuming that the elasticity of patenting to R&D expenditure equals 0.5. Column (h) 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 (o) and (p) from Caliendo and Parro (2015).