DISCUSSION PAPER SERIES

DP13799

TECHNOLOGY GAPS, TRADE AND INCOME

Thomas Sampson

INTERNATIONAL TRADE AND REGIONAL ECONOMICS AND MACROECONOMICS AND GROWTH

ISSN 0265-8003

TECHNOLOGY GAPS, TRADE AND INCOME

Thomas Sampson

Discussion Paper DP13799
Published 14 June 2019
Submitted 10 June 2019

Centre for Economic Policy Research
33 Great Sutton Street, London EC1V 0DX, UK
Tel: +44 (0)20 7183 8801
www.cepr.org

TECHNOLOGY GAPS, TRADE AND INCOME

Abstract

This paper studies the origins and consequences of international technology gaps. I develop an endogenous growth model where R&D efficiency varies across countries and productivity differences emerge from firm-level technology investments. The theory characterizes how innovation and learning determine technology gaps, trade and global income inequality. Countries with higher R&D efficiency are richer and have comparative advantage in more innovation-dependent industries where the advantage of backwardness is lower and knowledge spillovers are more localized. I estimate R&D efficiency by country and innovation-dependence by industry from R&D and bilateral trade data. Calibrating the model implies technology gaps, due to cross-country differences in R&D efficiency, account for around one-quarter to one-third of nominal wage variation within the OECD.

JEL Classification: F11, F43, O14, O41

Keywords: Technology gaps, Trade, Technology investment, Ricardian comparative advantage, International wage inequality

Thomas Sampson - t.a.sampson@lse.ac.uk
LSE and CEPR

Technology Gaps, Trade and Income*

Thomas Sampson†

London School of Economics

June 2019

Abstract

This paper studies the origins and consequences of international technology gaps. I develop an endogenous growth model where R&D efficiency varies across countries and productivity differences emerge from firm-level technology investments. The theory characterizes how innovation and learning determine technology gaps, trade and global income inequality. Countries with higher R&D efficiency are richer and have comparative advantage in more innovation-dependent industries where the advantage of backwardness is lower and knowledge spillovers are more localized. I estimate R&D efficiency by country and innovation-dependence by industry from R&D and bilateral trade data. Calibrating the model implies technology gaps, due to cross-country differences in R&D efficiency, account for around one-quarter to one-third of nominal wage variation within the OECD.

*I am grateful to Pol Antràs, Arnaud Costinot, Ha Nguyen, Andrés Rodríguez-Clare, Pau Roldan-Blanco, Christopher Tonetti and seminar participants at Barcelona, Bologna, Harvard-MIT, Lancaster, Bank of Lithuania, Paris School of Economics, Tilburg, the Barcelona GSE Summer Forum, the CEGAP Economic Growth and Policy conference in Durham, the CESifo Global Economy conference in Munich, ERWIT London, ESSIM Tarragona, ETSG Florence, the ISGEP workshop in Pescara, the Princeton IES Summer Workshop, the Public Economic Policy Responses to International Trade Consequences conference in Munich, the St. Louis Advances in Research conference and the World Bank Globalization: Contents and Discontents conference in Kuala Lumpur for helpful comments. 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.

1 Introduction

International productivity differences lead to variation in living standards (Caselli 2005) and determine Ricardian comparative advantage and the pattern of trade (Ricardo 1817; Dornbusch, Fischer and Samuelson 1977; Eaton and Kortum 2002; Costinot, Donaldson and Komunjer 2012). Productivity differences may result either from variation in the efficiency with which technologies and factor inputs are utilised (Acemoglu and Zilibotti 2001; Hsieh and Klenow 2009) or from technology gaps that occur when countries have access to different production technologies. This paper studies technology gaps. What determines the size of international technology gaps? How do technology gaps differ across industries? And how important are technology gaps in explaining international variation in wages and incomes?

Technology gaps arise because some countries are more innovative than others. Nelson (1993) describes how differences in the institutions and infrastructure that support knowledge creation, which he terms the "national innovation system", lead to large differences in innovative performance. Or, as Ohlin (1933, p.86) puts it "Nations vary much in inventive ability". For example, in 2015 the US and Japan together produced 30% of world GDP, but accounted for 47% of applications filed under the World Intellectual Property Organization's Patent Cooperation Treaty.[1]

But innovation is not the only source of technological improvements. Firms that are behind the technology frontier can raise their productivity by learning about and adopting existing technologies. Consequently, technology gaps also depend upon the rate at which ideas diffuse within countries (Lucas and Moll 2014; Perla and Tonetti 2014) and, in open economies, upon international knowledge diffusion (Grossman and Helpman 1991; Parente and Prescott 1994; Barro and Sala-i-Martín 1997; Eaton and Kortum 1999; Howitt 2000; Klenow and Rodríguez-Clare 2005; Buera and Oberfield 2016). If diffusion is fast, the gap between innovators and imitators will be small, whereas slow diffusion will increase the advantage that accrues to knowledge creators.

To understand the origins and consequences of technology gaps, this paper develops a model of how innovation and learning jointly shape the international productivity distribution. Crucially, I allow the innovation and learning technologies to differ across sectors. Consequently, differences in national innovation systems generate comparative advantage. This has two implications. First, the model gives a new theory of

[1] Patent application data obtained from http://www.wipo.int/pressroom/en/articles/2016/article_0002.html on 27 September 2016. GDP shares at market exchange rates calculated from the World Bank's World Development Indicators.

the origins of Ricardian comparative advantage in which productivity differences arise endogenously from firm-level technology investments. Second, trade data can be used to estimate the parameters that determine the size of international technology gaps. Using these estimates to calibrate the model, the paper quantifies how differences in national innovation systems affect wages and income levels. This provides a novel method for addressing the importance of innovation in explaining international variation in living standards.

The theory has three key components, which are introduced in Section 2. First, the efficiency of R&D varies across countries due to differences in national innovation systems. Countries with better national innovation systems have absolute advantages 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. The threshold is lower in countries with higher R&D efficiency.

Third, there are knowledge spillovers within and across countries. Knowledge is used as an input for 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 R&D or adoption (Gerschenkron 1962; Griffith, Redding and Van Reenen 2004). I assume all knowledge spillovers occur within industries and allow both the localization of knowledge spillovers and the advantage of backwardness to be industry-specific. Peri (2005), Acharya and Keller (2009) and Malerba, Mancusi and Montobbio (2013) show 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.

Because of international knowledge spillovers and the advantage of backwardness, on a balanced growth path technology gaps (i.e. relative productivity levels) are stable, both between domestic firms and across countries. Section 3 characterizes balanced growth in a global economy with many countries and industries and studies the equilibrium intranational and international technology gaps. Unsurprisingly, 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 upgrade productivity through technology adoption and productivity is increasing in R&D capability among innovative firms.

However, the degree of productivity dispersion is endogenous and industry-specific. The model shows how the size of equilibrium technology gaps is determined by the relative strengths of the dispersion and concentration forces in the global economy. The dispersion force results from local knowledge spillovers and differences in R&D efficiency between countries and R&D capability across firms. 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 dispersion.

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 strengthen the dispersion force. I show that industries can be characterized by their *innovation-dependence*, which is decreasing in the advantage of backwardness and increasing in the localization of knowledge spillovers. On any balanced growth path, technology gaps are greater in more innovation-dependent industries. Consequently, countries with higher R&D efficiency have a comparative advantage in industries with a lower advantage of backwardness and more localized knowledge spillovers.

While cross-industry variation in innovation-dependence determines the pattern of comparative advantage, the level of innovation-dependence affects the size of aggregate wage and income gaps. In the simple case with no trade costs, the elasticity of wages to R&D efficiency is proportional to a weighted average of industries' innovation-dependence levels. With trade costs, the relationship is more complex, but the intuition 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 gaps in wages and income per capita.

To illustrate the empirical applicability of the theory, Section 4 estimates and calibrates a first-order approximation to the model using data on 25 OECD economies. There are two key sets of parameters that I obtain using the structure of the model: country R&D efficiency levels and industry innovation-dependences. Because the threshold for selection into R&D is decreasing in R&D efficiency, the ratio of R&D expenditure to value-added at the industry level is larger in countries with higher R&D efficiency. Using this relationship, I infer R&D efficiency from cross-country, within-industry variation in R&D intensity.

Given R&D efficiency, I estimate innovation-dependence for 22 goods industries using the trade equation implied by the model. Innovation-dependence is estimated to be positive and significant in all industries except Mining, where it is insignificantly different from zero. It is largest in the Computers, Electrical equipment and Chemicals industries and smallest in the Petroleum, Agriculture, Food and Wood industries. The estimates imply that moving from the 25th to the 75th percentile of the R&D efficiency distribution increases a country's exports by 57% more for an industry at the 75th percentile of the innovation-dependence distribution than for an industry at the 25th percentile.

I use the calibrated model to quantify the extent to which R&D efficiency can account for cross-country differences in wages and incomes. Development accounting studies show that productivity differences are a major determinant of international variation in income per capita (Caselli 2005) and recent work finds that misallocation is a quantitatively important source of productivity differences (Hsieh and Klenow 2009). I assess the quantitative relevance of an alternative cause of productivity variation, namely that more innovative countries have access to better technologies. I find that R&D efficiency differences account for around one-quarter to one-third of observed variation in nominal wages within the OECD. Assuming innovation-dependence is zero in the services sector, the model also implies that R&D efficiency accounts for around one-sixth of observed income per capita variation. These results provide new evidence on how technology gaps, caused by differences in R&D efficiency, contribute to wage and income differences across OECD countries.

The calibration builds upon a small existing quantitative literature on R&D and international knowledge diffusion. Parente and Prescott (1994) calibrate a single industry exogenous growth model and argue 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) show quantitatively that variation in R&D investment may be sufficient to generate observed international productivity gaps. They also hypothesize that differences in R&D investment rates are caused by policies and institutions that affect R&D costs or returns. 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, this paper's empirical contribution is to show how R&D efficiency and innovation-dependence can be estimated from R&D and trade data without using information on income differences, and to quantify the share of international wage and income variation attributable to technology gaps resulting from differences in R&D efficiency.

Eaton and Kortum (1999) estimate a model of innovation and diffusion in five leading research countries and quantify the degree of international technology diffusion. More recently, Buera and Oberfield (2016) incorporate diffusion into the trade theory of Bernard et al. (2003) and calibrate the model to examine how changes in trade costs affect productivity. These papers shed light on the extent and effects of international diffusion, but do not address comparative advantage or the consequences of differences in national innovation systems.

The theory developed in this paper contributes to several strands of the trade and growth literatures. Existing dynamic models analyze how innovation determines comparative advantage in high-tech production (Grossman 1990; Grossman and Helpman 1991, chs.7-8). Yet innovation is highly concentrated in a few advanced countries and, within countries, at a few high productivity firms. For most firms, and throughout much of the world, learning also contributes to productivity growth. I study how innovation and learning jointly shape comparative advantage and identify the importance of the advantage of backwardness and the localization of knowledge spillovers in determining an industry's innovation-dependence.

Analysis of innovation and the pattern of trade in endogenous growth models can also be found in Grossman and Helpman (1990), Taylor (1993) and Durkin Jr. (1997). In these studies exogenous variation in the productivity of R&D relative to output production determines comparative advantage. Similarly, Somale (2016) and Cai, Li and Santacreu (2019) study trade liberalization in many industry versions of Eaton and Kortum (2001) where comparative advantage in production is shaped by exogenous comparative advantage in innovation. Unlike these papers, I do not assume any exogenous variation in comparative advantage across countries.

Learning and imitation shape trade flows in product cycle models, but the product cycle literature focuses on intraindustry trade (Vernon 1966; Krugman 1979; Grossman and Helpman 1991, chs.11-12).[2] Likewise, learning plays an important role in recent work that incorporates firm heterogeneity into open economy growth models (Perla, Tonetti and Waugh 2015; Impullitti and Licandro 2016; Sampson 2016a; Akcigit, Ates and Impullitti 2018) and in studies of the relationship between technology diffusion, growth and the spatial distribution of economic activity (Desmet and Rossi-Hansberg 2014; Desmet, Nagy and Rossi-Hansberg 2016). However, none of these papers analyze comparative advantage. Learning-by-doing models show how initial conditions can shape long-run comparative advantage in the presence of within-country, within-industry knowledge spillovers (Krugman 1987; Redding 1999). In contrast to the learning-by-doing literature, in this paper the existence of global knowledge spillovers implies steady state comparative advantage is not path dependent.

The methodology I use to estimate innovation-dependence from bilateral trade data is related to empiri-

[2] A notable exception is Lu (2007) who use a quality ladder product cycle model to study how interindustry variation in the size of the quality step determines North-South comparative advantage.

cal studies of comparative advantage that interact country and industry characteristics (Romalis 2004; Nunn 2007; Chor 2010; Manova 2013) and to work by Hanson, Lind and Muendler (2013) and Levchenko and Zhang (2016) that uses structural gravity models to infer productivity differences from trade flows. 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 this paper, 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.

In addition to shedding new light on the origins of comparative advantage, the theory also makes a methodological contribution to the endogenous growth literature by developing a new model where growth results from technology investment by incumbent firms. Foster, Haltiwanger and Krizan (2001) and Garcia-Macia, Hsieh and Klenow (2015) estimate that most growth in US manufacturing comes from incumbent firms, rather than creative destruction or the introduction of new varieties. This paper's model of incumbent firm R&D complements recent work on incumbent innovation and imitation in closed economy, quality ladder models (Klette and Kortum 2004; Akcigit and Kerr 2016; König, Lorenz and Zilibotti 2016) and in symmetric country trade models (Atkeson and Burstein 2010; Perla, Tonetti and Waugh 2015). An appealing feature of the framework is that it remains tractable even with many industries and many asymmetric countries. This tractability facilitates estimation and calibration of the model.

2 A Model of the Global Productivity Distribution

The global economy comprises S countries indexed by s. Time t is continuous and all markets are competitive. All parameters are assumed to be time invariant.

2.1 Preferences

Let L_s be the population of country s and assume there is no population growth. Within each country all individuals are identical and have intertemporal preferences given by:

$$U(t) = \int_t^{\infty} e^{-\rho(\tilde{t}-t)} \log c(\tilde{t}) d\tilde{t},$$

where $\rho > 0$ is the discount rate and c denotes consumption per capita. With these preferences the elasticity of intertemporal substitution equals one. Individuals can lend or borrow at interest rate ι_s. An individual in country s with initial assets $a(t)$ chooses a consumption path to maximize utility subject to the budget constraint:

$$\dot{a}(t) = \iota_s(t)a(t) + w_s(t) - z_s(t)c(t), \tag{1}$$

where w_s is the wage and z_s is the consumption price in country s.

Solving the individual's intertemporal optimization problem gives the Euler equation:

$$\frac{\dot{c}}{c} = \iota_s - \rho - \frac{\dot{z}_s}{z_s}, \tag{2}$$

where to simplify notation I have suppressed the dependence of the endogenous variables on time. For the remainder of the paper I will not write variables as an explicit function of time unless it is necessary to avoid confusion. Since all agents within a country are identical, we can write per capita consumption and assets as country-specific variables c_s and a_s, respectively. Aggregate consumption in country s is $c_s L_s$ and the aggregate value of asset holdings equals $a_s L_s$. The transversality condition for intertemporal optimization in country s is:

$$\lim_{\tilde{t}\to\infty} \left\{ a_s(\tilde{t}) \exp\left[-\int_t^{\tilde{t}} \iota_s(\hat{t}) d\hat{t} \right] \right\} = 0. \tag{3}$$

There are J industries, indexed by j. Consumer demand is Cobb-Douglas across industries and within industries output is differentiated by country of origin following Armington (1969). To be specific, aggregate consumption in country s is given by:

$$c_s L_s = \prod_{j=1}^{J} \left(\frac{X_{js}}{\mu_j} \right)^{\mu_j}, \qquad \text{with } \sum_{j=1}^{J} \mu_j = 1,$$

where X_{js} denotes consumption of industry j output in country s and μ_j equals the share of industry j in consumption expenditure. Let $x_{j\tilde{s}s}$ be industry j output from country $\tilde{s}$ that is consumed in country s. Then:

$$X_{js} = \left(\sum_{\tilde{s}=1}^{S} x_{j\tilde{s}s}^{\frac{\sigma-1}{\sigma}} \right)^{\frac{\sigma}{\sigma-1}},$$

where $\sigma > 1$ is the Armington elasticity.

There are iceberg trade costs $\tau_{j\tilde{s}s}$ of shipping industry j output from country $\tilde{s}$ to country s. This implies that, if $p_{j\tilde{s}}$ is the price received by producers in country $\tilde{s}$, then the price in country s is $\tau_{j\tilde{s}s} p_{j\tilde{s}}$. Solving consumers' intratemporal optimization problem yields:

$$P_{js} X_{js} = \mu_j z_s c_s L_s, \tag{4}$$

$$z_s = \prod_{j=1}^{J} P_{js}^{\mu_j}, \tag{5}$$

$$x_{j\tilde{s}s} = \left(\tau_{j\tilde{s}s} \frac{p_{j\tilde{s}}}{P_{js}} \right)^{-\sigma} X_{js}, \tag{6}$$

$$P_{js} = \left(\sum_{\tilde{s}=1}^{S} \tau_{j\tilde{s}s}^{1-\sigma} p_{j\tilde{s}}^{1-\sigma} \right)^{\frac{1}{1-\sigma}}, \tag{7}$$

where P_{js} denotes the price index for industry j in country s.

2.2 Production

Within each country-industry pair, all firms produce the same homogeneous output, but firms differ along two dimensions: R&D capability ψ and productivity θ. R&D capability is a time-invariant firm characteristic that affects a firm's R&D efficiency and, consequently, the evolution of its productivity. Section 2.3 describes technology investment and productivity dynamics. I assume the distribution of R&D capabilities in each industry has support $\left[\psi^{\min}, \psi^{\max}\right]$ and a continuous cumulative distribution function $G(\psi)$ that does not vary by country.

Productivity is a time-varying, firm-level state variable that determines a firm's production efficiency. Labor is the only factor of production. A firm in industry j and country s with productivity θ that employs l^P production workers produces output:

$$y = \theta \left(l^P\right)^{\beta}, \qquad \text{with } 0 < \beta < 1. \tag{8}$$

The production technology is independent of the firm's R&D capability ψ.

Each firm chooses production employment to maximize production profits $\pi^P = p_{js}y - w_s l^P$ taking the output price p_{js}, the wage w_s and its productivity θ as given.[3] The assumption $\beta < 1$ implies the firm's revenue function is strictly concave in employment. Solving the profit maximization problem implies the firm's production employment l_{js}^P, output y_{js} and production profits π_{js}^P are given by:

$$l_{js}^P(\theta) = \left(\frac{\beta p_{js}\theta}{w_s}\right)^{\frac{1}{1-\beta}}, \tag{9}$$

$$y_{js}(\theta) = \left(\frac{\beta p_{js}}{w_s}\right)^{\frac{\beta}{1-\beta}} \theta^{\frac{1}{1-\beta}}, \tag{10}$$

$$\pi_{js}^P(\theta) = (1-\beta)\left(\frac{\beta}{w_s}\right)^{\frac{\beta}{1-\beta}} (p_{js}\theta)^{\frac{1}{1-\beta}}. \tag{11}$$

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

In country s, good j is produced by a mass M_{js} of firms and the cumulative distribution function of firm productivity is $H_{js}(\theta)$. Both M_{js} and H_{js} are endogenous. M_{js} is determined by the free entry condition described in Section 2.5, while H_{js} depends upon firms' technology investment choices. Summing up across firms we have that aggregate production employment L_{js}^P in industry j and country s is:

$$L_{js}^P = M_{js}\left(\frac{\beta p_{js}}{w_s}\right)^{\frac{1}{1-\beta}} \int_\theta \theta^{\frac{1}{1-\beta}} dH_{js}(\theta), \tag{12}$$

and aggregate output is:

$$Y_{js} = M_{js}\left(\frac{\beta p_{js}}{w_s}\right)^{\frac{\beta}{1-\beta}} \int_\theta \theta^{\frac{1}{1-\beta}} dH_{js}(\theta). \tag{13}$$

2.3 Technology Investment

Firms can increase their productivity by investing in technology upgrading. Firms choose between two forms of technology investment: R&D and adoption. R&D investment seeks to create new ideas and technologies, while firms that choose adoption aim to learn existing production techniques and adapt these techniques to

[3] Technology investment affects future productivity growth (see Section 2.3), but not the current value of θ. Consequently, the firm's static production decision is separable from its dynamic technology investment decision.

improve their production methods. I assume the R&D technology is such that a firm with capability ψ and productivity θ that employs l^R workers in R&D has productivity growth:

$$\frac{\dot{\theta}}{\theta} = \psi B_s \left(\frac{\theta}{\chi_{js}^R}\right)^{-\gamma_j} \left(l^R\right)^{\alpha} - \delta, \tag{14}$$

where $\gamma_j > 0$, $\delta > 0$ and $\alpha \in (0, 1)$.

Firms with a higher capability ψ are better at R&D. In equilibrium, this will lead to cross-firm heterogeneity in R&D investment and productivity. R&D efficiency also differs across countries due to variation in B_s which captures differences in the quality of a country's national innovation system. Countries with a higher B_s have an absolute advantage in R&D. Knowledge spillovers allow firms that perform R&D to build upon the knowledge created by past innovations at home and abroad and χ_{js}^R denotes the R&D knowledge level in industry j and country s. A higher knowledge level raises the efficiency of R&D investment. χ_{js}^R is defined in Section 2.4 and depends upon the domestic productivity frontier and global knowledge capital. The returns to scale in R&D are given by α, while δ is the rate at which a firm's technical knowledge depreciates causing its productivity to decline.

Conditional on a firm's capability and R&D employment, the R&D technology implies productivity growth is decreasing in the firm's current productivity (relative to the knowledge level) with elasticity γ_j. This means there exists an advantage of backwardness that benefits firms further from the technology frontier (Gerschenkron 1962; Nelson and Phelps 1966). 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. 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 requirements for successful technology adoption are closely related to those needed for innovation. For example, when discussing Japan's industrialization Rosenberg (1990, p.152) argues that "Although the contrast between imitation and innovation is often sharply drawn ... It is probably a mistake to believe that the skills required for successful borrowing and imitation are qualitatively drastically different from those required for innovation." Consequently, I let the functional form of the adoption technology be the same as that of the R&D technology. However, adoption does differ 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.[4] Second, adoption draws more heavily on existing knowledge than R&D, which I model by assuming the adoption knowledge level χ_{js}^A is greater than the R&D knowledge level χ_{js}^R. Given these assumptions the adoption technology is:

$$\frac{\dot{\theta}}{\theta} = B^A \left(\frac{\theta}{\chi_{js}^A}\right)^{-\gamma_j} \left(l^A\right)^{\alpha} - \delta, \tag{15}$$

where l^A denotes adoption employment and:

$$\chi_{js}^A = \eta \chi_{js}^R, \qquad \text{with } \eta > 1. \tag{16}$$

I also assume that the decreasing returns to scale in technology investment implied by $\alpha < 1$ applies to the sum of R&D employment and adoption employment. It follows that no firm will ever choose to invest in both R&D and adoption simultaneously.

Each firm faces a constant instantaneous probability $\zeta > 0$ of suffering a shock that leads to the death of the firm. Taking this risk into account, the firm chooses paths for employment in R&D and adoption to maximize its value subject to the R&D technology (14) and the adoption technology (15). 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:

$$V_{js}(\psi, \theta) = \int_t^{\infty} \exp\left[-\int_t^{\tilde{t}} (\iota_s + \zeta)\, d\hat{t}\right] \left[\pi_{js}^P(\theta) - w_s\left(l^R + l^A\right)\right] d\tilde{t}, \tag{17}$$

where $\pi_{js}^P(\theta)$ is given by (11). All endogenous variables in this expression, including the firm's value function, are time dependent.

2.4 Knowledge Spillovers

Knowledge is non-rival and at least partially non-excludable. Consequently, firms learn from the R&D undertaken by their domestic and foreign competitors. However, knowledge spillovers are geographically localized (Jaffe, Trajtenberg and Henderson 1993; Branstetter 2001; Keller 2002) implying that domestic

[4] Appendix B shows that the paper's findings are robust to allowing for the efficiency of adoption investment to depend upon B_s provided the elasticity of adoption efficiency to B_s is below one.

spillovers are weaker than international spillovers. The localization of spillovers can also vary by industry due to differences in the importance of tacit knowledge, the degree of communication between firms within and across countries, and the extent to which production techniques are "circumstantially sensitive" and must be adapted to local requirements (Evenson and Westphal 1995). Peri (2005) and Malerba, Mancusi and Montobbio (2013) find that knowledge flows in Electronics are more global than in other sectors.

The knowledge levels χ_{js}^{R} and χ_{js}^{A} capture knowledge spillovers. I assume knowledge is a function of the domestic productivity frontier and global knowledge capital accumulated through past R&D investment. I focus on intra-industry knowledge flows and do not allow for interindustry spillovers. To be specific, let ω index firms and let Ω_{js} denote the set of firms operating in industry j in country s. Define:

$$\theta_{js}^{\max}(\omega) = \sup_{\tilde{\omega} \in \Omega_{js}, \tilde{\omega} \neq \omega} \left\{ \theta(\tilde{\omega}) \right\}.$$

$\theta_{js}^{\max}(\omega)$ is the supremum of the productivity of all firms in industry j in country s excluding firm ω. Within each country-industry pair there will always be either zero or a continuum of firms with any given productivity level.[5] 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 given by:

$$\chi_{js}^{R} = \left(\theta_{js}^{\max}\right)^{\frac{\kappa_j}{1+\kappa_j}} \chi_j^{\frac{1}{1+\kappa_j}}, \tag{18}$$

where χ_j denotes the global knowledge capital in industry j. Each firm takes the current and future values of χ_{js}^{R} and the adoption knowledge level $\chi_{js}^{A} = \eta \chi_{js}^{R}$ as given when choosing its technology investment.

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.

The knowledge level is homogeneous of degree one in the pair $\left(\theta_{js}^{\max}, \chi_j\right)$. Together with the assumption introduced in equations (14) and (15) that productivity growth is a function of current productivity

[5] In steady state this follows from the assumption that there exist a continuum of firms with each capability ψ in the support of G, see Section 3.2. 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.

relative to the R&D and adoption knowledge levels, this homogeneity restriction is necessary for the existence of a balanced growth path. It ensures knowledge spillovers are sufficiently strong to sustain ongoing productivity growth.

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:

$$\frac{\dot{\chi}_j}{\chi_j} = \sum_{s=1}^{S} M_{js} \int_{\psi^{\min}}^{\psi^{\max}} \lambda_{js}(\psi) l_{js}^{R}(\psi) dG(\psi). \tag{19}$$

where $\lambda_{js}(\psi) \geq 0$ gives the strength of R&D spillovers, which I allow to vary by industry j, country s and the firm's R&D capability ψ. Investment in adoption does not affect the global knowledge capital because it does not generate any new ideas.

2.5 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 (Luttmer 2007; Sampson 2016a) I assume the capability ψ and initial productivity θ of each entrant are determined by a random draw from the joint distribution of ψ and θ in the entrants' country and industry when the new firm is created. This implies the existence of spillovers from incumbents to entrants within a country-industry pair.[6]

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

$$f^E w_s = \int_{(\psi,\theta)} V_{js}(\psi, \theta) d\tilde{H}_{js}(\psi, \theta), \tag{20}$$

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

Let L_{js}^E 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_{js}^E / f^E. Since firms die at rate ζ this means the mass of firms M_{js} evolves according to:

[6] In Sampson (2016a) spillovers from incumbents to entrants lead to endogenous growth through a dynamic selection mechanism. By contrast, in this paper there is no selection and R&D investment by incumbent firms is the source of growth.

$$\dot{M}_{js} = -\zeta M_{js} + \frac{L_{js}^E}{f^E}. \tag{21}$$

2.6 Market Clearing

To complete the specification of the model, we need to impose market clearing for labor, goods and assets. The labor market clearing condition in each country s is:

$$L_s = \sum_{j=1}^{J} \left(L_{js}^P + L_{js}^R + L_{js}^A + L_{js}^E \right), \tag{22}$$

where production employment L_{js}^P in industry j is given by (12), L_{js}^R denotes aggregate R&D employment, L_{js}^A denotes aggregate adoption employment and L_{js}^E is aggregate employment in entry.

Industry output markets clear country-by-country implying domestic output Y_{js} equals the sum of sales to all countries inclusive of the iceberg trade costs:

$$Y_{js} = \sum_{\tilde{s}=1}^{S} \tau_{js\tilde{s}} x_{js\tilde{s}}. \tag{23}$$

I assume there is no international lending and asset markets clear at the national level. Therefore, for each country s total asset holdings equal the aggregate value of all domestic firms:

$$a_s L_s = \sum_{j=1}^{J} M_{js} \int_{(\psi,\theta)} V_{js}(\psi,\theta) d\tilde{H}_{js}(\psi,\theta). \tag{24}$$

I also let global consumption expenditure be the numeraire implying:

$$\sum_{s=1}^{S} z_s c_s L_s = 1. \tag{25}$$

Finally, I assume the parameters that govern 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:* $\frac{1}{1-\beta} > \gamma_j > \frac{\alpha}{1-\beta} + \frac{\kappa_j \gamma_j}{1+\kappa_j}$.

Assumption 1 is sufficient to ensure concavity in firms' intertemporal optimization problems.

2.7 Equilibrium

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 (1) giving the Euler equation (2) and the transversality condition (3); (ii) individuals' intratemporal consumption choices imply consumption levels and prices satisfy (4)-(7); (iii) firms choose production employment to maximize production profits implying industry level production employment and output are given by (12) and (13), respectively; (iv) firms' productivity levels evolve according to the R&D technology (14) and the adoption technology (15) and firms choose R&D and adoption employment to maximize their value (17); (v) the R&D and adoption knowledge levels are given by (16) and (18); (vi) global knowledge capital evolves according to (19); (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 (20) holds and the mass of firms evolves according to (21), and; (viii) labor, output and asset market clearing imply (22)-(24) hold.

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 in this model are differences in R&D efficiency B_s, population L_s and trade costs $\tau_{j\tilde{s}s}$.

3 Balanced Growth Path

This section analyzes a balanced growth path equilibrium of the global economy. I define a balanced growth path as an equilibrium in which all aggregate country and industry level variables grow at constant rates and the productivity distributions $H_{js}(\theta)$ shift outwards at constant rates. In what follows I outline how to solve for a balanced growth path and characterize its properties. Full details of the solution together with proofs of the propositions can be found in Appendix A.

3.1 Growth Rates

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 (16) and (18) 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.[7] Consequently, the productivity distribution $H_{js}(\theta)$ shifts outwards at rate g_j for all s. This means $H_{js}(\theta, t) = H_{js}\left(e^{g_j(\tilde{t}-t)}\theta, \tilde{t}\right)$ for all times t, $\tilde{t}$ and productivity levels θ. The productivity growth rate of each industry is constant across countries because $\kappa_j < \infty$ ensures the existence of some global knowledge spillovers.

Now let q_s be the growth rate of consumption per capita c_s. On a balanced growth path $q_s = q$ 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:

$$q = \sum_{j=1}^{J} \mu_j g_j. \tag{26}$$

In this economy rising productivity is the only source of growth and since the productivity growth rate in each industry does not vary by country, all countries have the same consumption per capita growth rate. It follows that, on a balanced growth path, cross-country heterogeneity leads to differences in the levels, not the growth rates, of endogenous variables.

On a balanced growth path we also have that 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. Finally, from the Euler equation (2) we obtain 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 (3) is satisfied.

[7] To see this, note that the R&D technology (14) implies balanced growth is possible only if the productivity frontier and the R&D knowledge level grow at the same rate in each country.

3.2 Firm-level Technology Investment and Productivity Dynamics

For the economy to be on a balanced growth path, the productivity distribution $H_{js}(\theta)$ must shift outwards at a constant rate g_j. Is optimal firm behavior consistent with this requirement? The next step in solving the model is to characterize firms' technology investment choices and productivity growth rates taking the time paths of w_s, p_{js}, χ^R_{js}, χ^A_{js} 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^R_{js} and χ^A_{js} both grow at rate g_j, and $\iota_s = \rho$.

To solve for firms technology investment choices, we must start by determining whether firms invest in R&D or adoption. A higher capability ψ increases R&D efficiency, but not adoption efficiency. 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 (14)-(16) we have:

$$\psi^*_{js} = \eta^{\gamma_j} \frac{B^A}{B_s}, \tag{27}$$

implying the R&D threshold 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.[8]

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. Taking the time derivative of ϕ and using the R&D technology (14) implies:

$$\frac{\dot{\phi}}{\phi} = \frac{1}{1-\beta}\left[\psi B_s \phi^{-\gamma_j(1-\beta)} \left(l^R\right)^{\alpha} - (\delta + g_j)\right]. \tag{28}$$

Substituting the production profits function (11) into the value function (17), using $\iota_s = \rho$ and changing variables from θ to ϕ, the optimization problem of a firm with capability ψ can be written as:

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

subject to the growth of ϕ being given by (28) and an initial value for ϕ at time t. Since w_s is constant, p_{js} declines at rate g_j and χ^R_{js} grows at rate g_j, the payoff function depends upon time only through exponential

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

discounting meaning the firm faces a discounted infinite-horizon optimal control problem with state variable ϕ and control variable l^R. I prove in Appendix A that any solution to the firm's problem must satisfy:

$$\frac{\dot{l}^R}{l^R} = \frac{1}{1-\alpha}\left[\rho + \zeta + \gamma_j\left(\delta + g_j\right) - \alpha\beta^{\frac{\beta}{1-\beta}}\psi B_s\left(\frac{p_{js}\chi_{js}^R}{w_s}\right)^{\frac{1}{1-\beta}}\phi^{1-\gamma_j(1-\beta)}\left(l^R\right)^{\alpha-1}\right]. \tag{29}$$

Equations (28) and (29) are a system of differential equations for ϕ and l^R. Setting $\dot{\phi} = 0$ and $\dot{l}^R = 0$ shows the system has a unique steady state ϕ_{js}^*, l_{js}^{R*} given by:

$$\phi_{js}^* = \left[\alpha\beta^{\frac{\beta}{1-\beta}}\left(\psi B_s\right)^{\frac{1}{\alpha}}\left(\frac{p_{js}\chi_{js}^R}{w_s}\right)^{\frac{1}{1-\beta}}\frac{\left(\delta + g_j\right)^{\frac{\alpha-1}{\alpha}}}{\rho + \zeta + \gamma_j\left(\delta + g_j\right)}\right]^{\frac{\alpha}{\gamma_j(1-\beta)-\alpha}}, \tag{30}$$

$$l_{js}^{R*} = \left[\alpha\beta^{\frac{\beta}{1-\beta}}\left(\psi B_s\right)^{\frac{1}{\gamma_j(1-\beta)}}\left(\frac{p_{js}\chi_{js}^R}{w_s}\right)^{\frac{1}{1-\beta}}\frac{\left(\delta + g_j\right)^{\frac{\gamma_j(1-\beta)-1}{\gamma_j(1-\beta)}}}{\rho + \zeta + \gamma_j\left(\delta + g_j\right)}\right]^{\frac{\gamma_j(1-\beta)}{\gamma_j(1-\beta)-\alpha}}. \tag{31}$$

Under Assumption 1 there exists a neighborhood of the steady state within which the firm's R&D problem has a unique solution which, conditional on the firm's survival, converges to the steady state. Thus, the steady state is locally saddle-path stable. 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 holding all else constant. 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 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. 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. Likewise, global knowledge spillovers ensure that level differences across countries in the location of the productivity distribution offset the growth effects resulting from all other sources of cross-country heterogeneity, meaning the industry productivity growth rate is the same in all countries.

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 (10) and (31) implies the steady state ratio of R&D investment to sales satisfies:

$$\frac{w_s l^{R*}_{js}}{p_{js} y_{js}\left(\phi^*_{js}\right)} = \frac{\alpha\left(\delta + g_j\right)}{\rho + \zeta + \gamma_j\left(\delta + g_j\right)}, \tag{32}$$

which is constant across firms within an 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.

Next consider a firm that invests in adoption because it has R&D capability below the threshold ψ^*_{js}. Substituting (16) and (27) into (15) implies the firm's productivity growth is given by:

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

Comparing this expression with the R&D technology (14) shows that the firm's adoption investment problem is equivalent to that of an R&D firm with capability ψ^*_{js}. It immediately follows that the firm's steady state relative productivity and adoption employment are given by (30) and (31), respectively, with $\psi = \psi^*_{js}$. Moreover, the steady state is unique and locally saddle-path stable given Assumption 1. Choosing adoption rather than R&D allows firms with capability below the R&D threshold to attain the productivity level of a firm with R&D capability ψ^*_{js}.

Recall that 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 the steady state productivity level corresponding to its capability and the productivity distributions of entering and exiting firms are identical. It follows that entry, exit and firms' optimal R&D and adoption choices generate balanced growth if and only if all incumbent firms are in steady state.[9]

[9]Formally, a balanced growth path only requires a mass M_{js} of firms to be in steady state, which allows for individual firms

On a balanced growth path both the location and the shape of the productivity distribution are endogenous.[10] The location varies by country and industry and depends upon w_s, p_{js}, χ_{js}^R, χ_{js}^A and g_j, which are determined in general equilibrium as described in Section 3.3 below. Within each country-industry pair, the shape of the productivity distribution depends upon the exogenous firm capability distribution $G(\psi)$, the R&D threshold ψ_{js}^* and the parameters α, β and γ_j.

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:

$$\frac{\theta_{js}^*(\psi')}{\theta_{js}^*(\psi)} = \begin{cases} \left(\frac{\psi'}{\psi}\right)^{\frac{1-\beta}{\gamma_j(1-\beta)-\alpha}}, & \psi' \geq \psi \geq \psi_{js}^*, \\ \left(\frac{\psi'}{\psi_{js}^*}\right)^{\frac{1-\beta}{\gamma_j(1-\beta)-\alpha}}, & \psi' \geq \psi_{js}^* \geq \psi, \\ 1, & \psi_{js}^* \geq \psi' \geq \psi. \end{cases} \tag{33}$$

When both firms perform R&D, technology gaps and productivity inequality are strictly increasing in α and β and strictly decreasing in γ_j.[11] 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. 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.

Adopters constitute a fringe of firms with mass $M_{js}G\left(\psi_{js}^*\right)$ that compete with innovators. 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 (9)-(11) inequality in production employment, revenue and profits are also strictly increasing in α, β and B_s and strictly decreasing in γ_j.

with zero mass to deviate from steady state. I overlook this distinction since it does not matter for industry or aggregate outcomes.

[10] By contrast, in most heterogeneous firm models following 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 Perla, Tonetti and Waugh (2015) and Sampson (2016a). An exception is Bonfiglioli, Crinò and Gancia (2015) who allow firms to choose between receiving productivity draws from distributions with different shapes.

[11] All results concerning inequality 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.

Proposition 1 summarizes how balanced growth path productivity varies across firms within the same country and industry.

Proposition 1. *Suppose Assumption 1 holds. On a balanced growth path equilibrium all firms 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;

(ii) Productivity inequality across 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.

3.3 General Equilibrium

To complete the solution for a balanced growth path equilibrium we can now impose the remaining equilibrium conditions. Before doing so, let Ψ_{js} be defined by:

$$\Psi_{js} \equiv \int_{\psi_{js}^*}^{\psi^{\max}} \psi^{\frac{1}{\gamma_j(1-\beta)-\alpha}} dG(\psi) + \left(\psi_{js}^*\right)^{\frac{1}{\gamma_j(1-\beta)-\alpha}} G\left(\psi_{js}^*\right). \tag{34}$$

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 the benefits resulting from selection into adoption, which are larger in countries with lower B_s.

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 I show in Appendix A that on a balanced growth path:

$$L_s = \sum_{j=1}^{J} \frac{\mu_j}{\rho+\zeta}\left(\zeta+\beta\rho+\frac{\alpha\rho\left(\delta+g_j\right)}{\rho+\zeta+\gamma_j\left(\delta+g_j\right)}\right) Z_{js}, \tag{35}$$

$$a_s L_s = \sum_{j=1}^{J} \frac{\mu_j}{\rho+\zeta}\left(1-\beta-\frac{\alpha\left(\delta+g_j\right)}{\rho+\zeta+\gamma_j\left(\delta+g_j\right)}\right) w_s Z_{js}, \tag{36}$$

$$g_j = \sum_{s=1}^{S} \mu_j \frac{\alpha\left(\delta+g_j\right)}{\rho+\zeta+\gamma_j\left(\delta+g_j\right)} \frac{Z_{js}}{\Psi_{js}} \int_{\psi_{js}^*}^{\psi^{\max}} \lambda_{js}(\psi)\psi^{\frac{1}{\gamma_j(1-\beta)-\alpha}} dG(\psi), \tag{37}$$

where:

$$Z_{js} \equiv \sum_{\tilde{s}=1}^{S} \frac{\tau_{js\tilde{s}}^{1-\sigma} \left(\rho a_{\tilde{s}} + w_{\tilde{s}}\right) L_{\tilde{s}} w_s^{-\sigma} \left(B_s \Psi_{js}^{\frac{\gamma_j(1-\beta)}{1+\kappa_j} - \alpha} \right)^{\frac{(\sigma-1)(1+\kappa_j)}{\gamma_j}}}{\sum_{\hat{s}=1}^{S} \tau_{j\hat{s}\tilde{s}}^{1-\sigma} w_{\hat{s}}^{1-\sigma} \left(B_{\hat{s}} \Psi_{j\hat{s}}^{\frac{\gamma_j(1-\beta)}{1+\kappa_j} - \alpha} \right)^{\frac{(\sigma-1)(1+\kappa_j)}{\gamma_j}}}. \tag{38}$$

Equations (35)-(37), together with the definition of Z_{js} in (38), 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 defines a balanced growth path. I prove in Appendix A that there exists a unique balanced growth path in the case where $J = 1$ and there are no trade costs. In the general case, I assume existence and obtain results that must hold on any balanced growth path.

R&D efficiency B_s enters the equilibrium equations not only directly, but also indirectly through ψ_{js}^* and Ψ_{js}. In an economy without adoption these indirect effects do not exist. Sections 3.4 and 3.5 characterize international technology and income gaps on a balanced growth path and show that the indirect effects reduce the dispersion in outcomes across countries.

To gain insight into the determinants of growth in this economy, consider a single sector version of the model. Setting $J = 1$ and substituting (35) into (37) yields:

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

The left hand side of this expression 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 by (26).

Growth is higher when R&D spillovers $\lambda_s(\cdot)$ are stronger and when there is more employment in R&D. As in the canonical growth models of Romer (1990) and Aghion and Howitt (1992) 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.

Raising the number of countries in the global economy increases growth because the R&D spillovers specified in (19) are global in scope. Consequently, global integration generates higher growth than autarky. 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 τ. Lower trade costs increase the effective size of export markets, but also expose domestic firms to increased import competition. In this 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.

3.4 Technology Gaps and Comparative Advantage

On a balanced growth path all countries have the same growth rates regardless of their underlying differences. Consequently, relative productivity levels remain constant over time when comparing countries within the same industry. However, heterogeneity in R&D efficiency does lead to international variation in productivity levels, which generates comparative advantage and cross-country gaps in incomes and consumption. This section studies the technology gaps that support a balanced growth path equilibrium and characterizes cross-sectional variation in exports and productivity.

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

$$\frac{\overline{\theta}^*_{js}}{\overline{\theta}^*_{j\tilde{s}}} = \left[\frac{B_s}{B_{\tilde{s}}}\left(\frac{\Psi_{js}}{\Psi_{j\tilde{s}}}\right)^{\frac{\gamma_j(1-\beta)}{1+\kappa_j}-\alpha}\right]^{\frac{1+\kappa_j}{\gamma_j}}. \tag{39}$$

This expression is the cross-country analogue of equation (33), which gives the ratio of steady state productivity levels for two firms in the same country and industry, but with different capabilities. We see that variation in R&D efficiency is the only source of international technology gaps and that B_s has a direct positive effect on productivity, as well as an indirect negative effect through Ψ_{js}.[12] In addition, technology gaps do not depend on the level of trade costs.

The direct effect of R&D efficiency results from firms in a country with lower B_s being less innovative, all else constant. Consequently, they fall behind foreign firms until their advantage of backwardness is

[12]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 (27).

sufficient to offset poor fundamentals and ensure they grow at the same rate as firms in more innovative countries. The indirect effect is the result of countries with higher B_s having a lower R&D threshold and, therefore, lower average effective capability Ψ_{js}. However, the direct effect is always stronger than the indirect effect and the net effect of higher B_s on average productivity is strictly positive, implying an increase in R&D efficiency shifts the productivity distribution in country s outwards relative to other countries.

The size of international technology gaps depends upon the elasticity of relative average productivity to R&D efficiency. I will refer to this elasticity as the level of innovation-dependence, since it captures the extent to which countries benefit from having higher B_s and being more innovative. Innovation-dependence differs by industry because of variation in the advantage of backwardness γ_j and the localization of knowledge spillovers κ_j.[13] 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 country s to country $\tilde{s}$ in industry j inclusive of trade costs. On a balanced growth path log exports can be written as:

$$\log EX_{js\tilde{s}} = \upsilon^1_{j\tilde{s}} + (\sigma - 1)\left(\log \overline{\theta}^*_{js} - \log w_s - \log \tau_{js\tilde{s}}\right), \tag{40}$$

where $\upsilon^1_{j\tilde{s}}$ is a destination-industry specific term defined in Appendix A. 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.

By substituting for $\overline{\theta}^*_{js}$ in (40) we obtain:

$$\log EX_{js\tilde{s}} = \upsilon^2_{j\tilde{s}} + (\sigma - 1)\left(\frac{1+\kappa_j}{\gamma_j}\log B_s + \frac{\gamma_j(1-\beta) - \alpha(1+\kappa_j)}{\gamma_j}\log \Psi_{js} - \log w_s - \log \tau_{js\tilde{s}}\right), \tag{41}$$

R&D efficiency has a direct effect on exports as well as indirect effects through Ψ_{js} and w_s. Moreover, (41) implies that the elasticity of exports to R&D efficiency equals $\sigma - 1$ times innovation-dependence. Therefore, countries with higher R&D efficiency will have a comparative advantage in more innovation-dependent industries. Differentiating (41) shows that the direct and indirect effects of R&D efficiency on

[13]Formally, I define innovation-dependence by: $\partial \log \left[B_s \Psi_{js}^{\frac{\gamma_j(1-\beta)}{1+\kappa_j} - \alpha}\right]^{\frac{1+\kappa_j}{\gamma_j}} / \partial \log B_s$. In general, this elasticity may vary across countries as well as industries. In Section 4 I estimate a first-order approximation of the model in which innovation-dependence is constant within industries.

comparative advantage reinforce each other and gives:

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

Innovation-dependence is decreasing in the advantage of backwardness and increasing in the localization of knowledge spillovers. Consequently, countries with higher R&D efficiency have a comparative advantage in industries with a lower γ_j and a higher κ_j. Proposition 2 summarizes these results.[14]

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 industries where the advantage of backwardness is smaller and in industries where the localization of knowledge spillovers is greater.

The pattern of Ricardian comparative advantage is stable on a balanced growth path because the combination of international technology gaps, the advantage of backwardness and global knowledge spillovers imply productivity growth does not vary by country. The advantage of backwardness imposes a cost on more productive firms by reducing the efficiency of their technology investment. In industries where the advantage of backwardness is greater, this cost is larger. Therefore, a higher γ_j reduces technology gaps between countries, just as it reduces technology gaps between firms within the same country. The indirect impact operating through Ψ_{js} reinforces this direct effect because the R&D threshold ψ^*_{js} is increasing in γ_j. This means that in industries where the advantage of backwardness is greater, the share of firms that invest in R&D is lower and differences in R&D efficiency matter less. Both effects make industries with higher γ_j less innovation-dependent, generating a comparative advantage for countries with lower B_s.

Countries with higher R&D efficiency have a higher domestic productivity frontier and, therefore, greater R&D and adoption knowledge levels by (16) and (18). When domestic spillovers are stronger relative to global spillovers these differences in knowledge levels are magnified, increasing the relative efficiency of technology investment in more productive countries and raising the industry's innovation-dependence. Consequently, countries with higher B_s have a comparative advantage in industries where knowledge spillovers

[14]The proof of Proposition 2 does not use the labor, output or asset market clearing conditions. This implies the pattern of comparative advantage does not depend upon how the market clearing conditions are specified.

are more localized.

Proposition 2 highlights an important distinction between how the advantage of backwardness and knowledge spillovers affect the global productivity distribution. Although a greater advantage of backwardness reduces the technology gaps between countries, stronger knowledge spillovers benefit those countries that are better able to take advantage of spillovers. When spillovers are more localized, the beneficiaries are the higher productivity countries that use better technologies and generate greater spillovers.

Before moving on, it is worth considering how other sources of industry heterogeneity affect comparative advantage. Suppose 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 are industry specific, but the model is otherwise unchanged. Then it is straightforward to show that all the equilibrium conditions derived above continue to hold after adding industry subscripts to these parameters. This has three immediate implications.

First, Proposition 2 still holds, implying the model's implications for how the advantage of backwardness and the localization of knowledge spillovers affect comparative advantage are robust to allowing for these additional sources of cross-industry heterogeneity. Second, cross-industry variation in δ_j, ζ_j, f_j^E, the industry expenditure shares μ_j or the strength of R&D spillovers $\lambda_{js}(\cdot)$ does not generate comparative advantage. Third, if there is no adoption, meaning Ψ_{js} is constant across countries, then $G_j(\psi)$, β_j, α_j and η_j do not affect comparative advantage. However, with adoption, it can be shown that countries with higher R&D efficiency 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 (27), which shrinks international technology gaps since the adoption technology is independent of R&D efficiency.

3.5 International Inequality

How do wages, income and consumption differ across countries on a balanced growth path? The intertemporal budget constraint (1) implies consumption per capita depends upon assets per capita, wages and the consumption price through:

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

The simplest case to consider is a single sector economy with free trade. In this case equations (35), (38) and (39) yield:

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

The relative wage of country s is increasing in its relative average productivity and, consequently, in its R&D efficiency.[15] Moreover, note that the elasticity of the relative wage to R&D efficiency equals $(\sigma-1)/\sigma$ times innovation-dependence. From Proposition 2 we know that 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.

With a single industry, assets per capita a_s are proportional to w_s by (35) and (36). 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. Therefore, relative income per capita and consumption per capita levels are equal to relative wages. It follows that international inequality in incomes and consumption is increasing in the degree of innovation-dependence. When innovation-dependence is high, the technology gap between leaders and followers is greater and this leads to larger dispersion in incomes and consumption. By contrast, low innovation-dependence reduces the cost of being less innovative and shrinks international gaps. 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 resulting from differences in R&D efficiency is strictly decreasing in the advantage of backwardness and strictly increasing in the localization of knowledge spillovers.

[15] The relative wage is also decreasing in relative population $L_s/L_{\tilde{s}}$ due to the assumption of Armington demand.

In the general case with many industries and trade costs, innovation-dependence continues to be a key determinant of international inequality. In particular, the elasticities of w_s, a_s, z_s and c_s to B_s all depend upon innovation-dependence in the J industries.[16] For example, whenever there is free trade, the elasticity of wages to R&D efficiency is:

$$\frac{\partial \log w_s}{\partial \log B_s} = \frac{\sigma - 1}{\sigma} \frac{\sum_{j=1}^{J} \left(\zeta + \beta\rho + \frac{\alpha\rho(\delta+g_j)}{\rho+\zeta+\gamma_j(\delta+g_j)} \right) \mu_j Z_{js} \left[\partial \log \left(B_s \Psi_{js}^{\frac{\gamma_j(1-\beta)}{1+\kappa_j} - \alpha} \right)^{\frac{1+\kappa_j}{\gamma_j}} / \partial \log B_s \right]}{\sum_{j=1}^{J} \left(\zeta + \beta\rho + \frac{\alpha\rho(\delta+g_j)}{\rho+\zeta+\gamma_j(\delta+g_j)} \right) \mu_j Z_{js}}.$$

This expression shows that wages are increasing in R&D efficiency and that the wage elasticity is proportional to a weighted average of the industry innovation-dependence levels. Consequently, an economy-wide increase in innovation-dependence increases the wage inequality caused by variation in B_s.

With trade costs, there is no simple expression for how relative wages and incomes depend upon R&D efficiency. However, by estimating innovation-dependence it's possible to calibrate the model and quantify the impact of R&D efficiency. The next section takes up this challenge.

4 Estimation and Quantification

To quantify the contribution of R&D efficiency to international income inequality, I assume the economy is on a balanced growth path and estimate the model to obtain countries' R&D efficiency levels and the innovation-dependence of each industry. R&D efficiency levels are inferred from variation in R&D intensity within industries, but across countries. Given R&D efficiency, I then use trade data to estimate innovation-dependence at the industry-level. With these parameters in hand, I calibrate the model and analyze the extent to which observed wage and income inequality across OECD countries can be explained by differences in R&D efficiency.

[16]This follows from the observation that B_s enters the balanced growth path equations for w_s, a_s, z_s and, therefore, c_s only through the term $B_s \Psi_{js}^{\frac{\gamma_j(1-\beta)}{1+\kappa_j} - \alpha}$. See Appendix A for the derivation of equilibrium consumption prices z_s.

4.1 Empirical Strategy

In order to estimate the model, I assume a functional form for the R&D capability distribution and take a first order approximation to the balanced growth path equilibrium under the assumption that the R&D threshold ψ_{js}^* is large. The first order approximation facilitates the empirical implementation by making the equilibrium conditions used for estimation log-linear in R&D efficiency B_s.

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.[17] Using this functional form in (34) to compute average effective capability Ψ_{js}, letting $\psi^{\max} \to \infty$ and taking a first order approximation for large ψ_{js}^* yields:

$$\Psi_{js} \approx \left(\psi_{js}^*\right)^{\frac{1}{\gamma_j(1-\beta)-\alpha}} = \left(\eta^{\gamma_j}\frac{B^A}{B_s}\right)^{\frac{1}{\gamma_j(1-\beta)-\alpha}}, \tag{42}$$

where the second line follows from (27). I estimate and calibrate the model using this approximation. Since the approximation drops terms of order $\left(\psi_{js}^*\right)^{-k}$, it is valid provided $\left(\psi_{js}^*\right)^{-k}$ is small. When $\psi^{\max} \to \infty$, $\left(\psi_{js}^*\right)^{-k}$ gives the fraction of firms that invest in R&D. In UK data for 2008-09, 9.9% of goods firms report performing R&D.

Let industry R&D intensity RD_{js} be the ratio of R&D investment to value-added in industry j and country s. R&D intensity does not vary across firms that perform R&D by (32). However, RD_{js} also depends upon whether firms choose R&D or adoption. Computing R&D intensity from (10), (30) and (31), taking a first order approximation for large ψ_{js}^* and using (27) to substitute for ψ_{js}^* gives:

$$RD_{js} = \frac{\alpha(\delta + g_j)}{\rho + \zeta + \gamma_j(\delta + g_j)} \frac{k\left[\gamma_j(1-\beta)-\alpha\right]}{k\left[\gamma_j(1-\beta)-\alpha\right]-1}\eta^{-k\gamma_j}\left(\frac{B_s}{B^A}\right)^k. \tag{43}$$

Equation (43) shows that R&D intensity is higher in countries with greater R&D efficiency. A higher B_s implies a larger share of firms select into R&D, which increases RD_{js}.[18] I will use equation (43) to infer countries' R&D efficiency levels B_s from within-industry variation in R&D intensity.

Next, we can calculate innovation-dependence using the approximation to Ψ_{js} in (42). This yields

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

[18] The implication that cross-country variation in RD_{js} comes entirely from the extensive margin is not necessary to obtain the industry and aggregate level predictions used for estimation and calibration. For example, if firm output is the sum of production 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} will come from the intensive margin, but the balanced growth path is otherwise unchanged.

that the innovation-dependence of industry j is $ID_j = \frac{(1-\beta)\kappa_j}{\gamma_j(1-\beta)-\alpha}$. Note that ID_j is increasing in κ_j and decreasing in γ_j, meaning the patterns of comparative advantage characterized in Proposition 2 are unaffected by using the approximation. We can also write the exports equation (41) as:

$$\log EX_{js\tilde{s}} = \upsilon^3_{j\tilde{s}} + (\sigma - 1) \left(ID_j \log B_s - \log w_s - \log \tau_{js\tilde{s}} \right), \tag{44}$$

where $\upsilon^3_{j\tilde{s}} = \upsilon^2_{j\tilde{s}} + \frac{\sigma-1}{\gamma_j} \frac{\gamma_j(1-\beta)-\alpha(1+\kappa_j)}{\gamma_j(1-\beta)-\alpha} \log \left(\eta^{\gamma_j} B^A \right)$. I will use this expression to obtain ID_j from industry-level bilateral trade data.[19]

Each industry's innovation-dependence enters the general equilibrium conditions (35)-(37) through Z_{js} defined in (38). Using the approximation to Ψ_{js} in (42) we have:

$$Z_{js} = \sum_{\tilde{s}=1}^{S} \frac{\tau^{1-\sigma}_{js\tilde{s}} \left(\rho a_{\tilde{s}} + w_{\tilde{s}} \right) L_{\tilde{s}} w_s^{-\sigma} B_s^{(\sigma-1)ID_j}}{\sum_{\hat{s}=1}^{S} \tau^{1-\sigma}_{j\hat{s}\tilde{s}} w_{\hat{s}}^{1-\sigma} B_{\hat{s}}^{(\sigma-1)ID_j}}, \tag{45}$$

which shows how the relationship between B_s and Z_{js} depends upon ID_j. After estimating B_s and ID_j, I use (45) to calibrate the model in Section 4.5.

4.2 Data

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

The primary data constraint is the limited availability of internationally comparable R&D data at the industry-level. I obtain R&D expenditure for 20 ISIC 2 digit manufacturing industries from the OECD's ANBERD database. 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

[19] In principle, innovation-dependence could also be estimated using measures of industry-level productivity. However, unlike bilateral trade, productivity is not directly observable.

with R&D intensity data.

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 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: Business Expenditure on R&D, and; the Annual Business Survey.

4.3 Estimation

The first step in estimating the model is to obtain R&D efficiency. Equation (43) shows that relative R&D efficiency can be recovered from relative R&D intensity levels within industries. Specifically, $RD_{js}/RD_{j\tilde{s}} = (B_s/B_{\tilde{s}})^k$. Therefore, I let:

$$b_s = \frac{1}{N_s} \sum_{t=2010}^{2014} \sum_{j=1}^{20} \log \left(\frac{RD_{jst}}{RD_{j\tilde{s}t}} \right), \tag{46}$$

where the summation omits any years and industries for which R&D intensity is missing, N_s is the number of non-missing observations and $\tilde{s}$ is a reference country. b_s measures $k \log B_s$ up to an unidentified additive constant. As a normalization, I set R&D efficiency in the US equal to one, $B_{US} = 1$. Figure 2 plots b_s against GDP per capita. On average richer countries have better national innovation systems and, therefore, higher b_s, but there are notable differences in R&D efficiency even within the wealthiest group of countries.

Next, to estimate innovation-dependence levels, we need to specify bilateral trade costs. Following Eaton and Kortum (2002), I model trade costs as a function of gravity variables. I also 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:

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

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 $\geq$ 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 j.

Using this specification of trade costs, we can rearrange the bilateral trade equation (44) to obtain the following estimating equation:

$$\log\left(\frac{EX_{js\tilde{s}}}{EX_{j\tilde{s}\tilde{s}}}\right) - (\sigma-1)\log\left(\frac{w_{\tilde{s}}}{w_s}\right) = -(\sigma-1)\frac{ID_j}{k}b_{\tilde{s}} \quad (47)$$
$$-(\sigma-1)\left(DIST^i_{js\tilde{s}} + BORD_{js\tilde{s}} + CLANG_{js\tilde{s}} + FTA_{js\tilde{s}} + \delta^2_{js}\right) + \epsilon_{js\tilde{s}},$$

where $\delta^2_{js} = \delta^1_{js} - ID_j b_s/k$ and the error term $\epsilon_{js\tilde{s}}$ includes unmodelled variation in trade costs, productivity and comparative advantage. Sales of domestic production to the domestic market $EX_{j\tilde{s}\tilde{s}}$ equal the difference between output and total exports. Therefore, the left hand side of (47) is observable given industry-level trade and output, country-level nominal wages and a value for $\sigma - 1$. From (40), 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. I start by setting $\sigma - 1$ equal to their preferred estimate of 6.53, while Section 4.5 analyzes the implications of using alternative trade elasticity estimates.

Table 1, column (a) reports estimates of ID_j/k obtained from (47) using pooled trade data for 2010-14. The sample includes exports of 117 countries to the 25 importers with R&D intensity data and covers 22 ISIC goods industries at the 2 digit level of aggregation (the 20 manufacturing industries used to calculate R&D efficiency plus the Agriculture and Mining industries). Innovation-dependence is estimated relative to the shape parameter of the R&D capability distribution k. However, note that k cancels out when estimated innovation-dependence is multiplied by b_s. When calibrating the model, the product of these two terms is sufficient to quantify the impact of R&D efficiency on comparative advantage, wages and income levels. Consequently, we do not need to calibrate k.

In column (a), estimated innovation-dependence is highest in the Machinery, Computers, Electrical equipment, and Pharmaceutical industries. It is lowest in the Mining, Agriculture, Petroleum, Food and

Leather industries. However, these estimates are likely to be biased upwards. To understand why, suppose there exist productivity differences across countries due to causes other than R&D efficiency. For example, conditional on the technological component of firm-level productivity θ, the efficiency of production may vary across countries because of differences in institutions, governance and infrastructure. We can allow for this possibility by assuming that, instead of (8), the production function is $y = A_{js}\theta \left(l^P\right)^\beta$ 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 for details). Except for the fact that A_{js} is present in the equilibrium conditions, none of the results in Section 3 are affected. In particular, exports are still given by (41) after adding $(\sigma-1)A_{js}$ to the right hand side. This implies the error term in (47) is a function of the importer's allocative efficiency $A_{j\tilde{s}}$ and any positive correlation between $A_{j\tilde{s}}$ and $b_{\tilde{s}}$ will lead to upwards bias in the estimates of innovation-dependence.[20]

To control for other sources of productivity variation, I use 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. Column (b) shows the results from estimating (47) including these variables. As expected, the magnitude of the innovation-dependence estimates declines, but the pattern of cross-industry variation is similar to column (a).

Column (c) also controls for sources of comparative advantage beyond R&D efficiency. I include the interactions 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 0.31, compared to 0.46 in column (b) and 0.62 in column (a). However, inter-industry variation in innovation-dependence is similar to columns (a) and (b). The correlation between the innovation-dependence estimates in column (c) and those in either column (a)

[20] The right hand side of (47) also depends upon A_{js}, but this is captured by the exporter-industry fixed effect δ^2_{js}.

or column (b) is 0.79.

Going forward, I will work with the estimates in column (c). According to these estimates, innovation-dependence is largest in the Computers, Electrical equipment and Chemicals industries and is positive and significant at the 5% level in all industries except Mining. The estimates imply that moving from the 25th to the 75th percentile of the R&D efficiency distribution increases a country's exports by 57% more for an industry at the 75th percentile of the innovation-dependence distribution than for an industry at the 25th percentile.

4.4 Model Validation

Before calibrating the model, I undertake two validation exercises to assess the model's empirical credibility. First, I examine cross-industry variation in firm-level R&D investment choices. Second, I conduct an out-of-sample test of the model's predictions for comparative advantage.

Firm-level R&D investment. The first validation exercise analyzes whether cross-industry variation in firms' intensive and extensive margin R&D investment choices is consistent with the model. The share of firms that invest in R&D $ShRD_{js}$ equals $\left(\psi^*_{js}\right)^{-k}$, which is decreasing in the advantage of backwardness γ_j by (27). Among firms with positive R&D investment, R&D intensity $FiRD_j$ is also decreasing in the advantage of backwardness by (32). Manipulating (27) and (32) yields:

$$\frac{1}{FiRD_j} = -\frac{1}{\alpha k \log \eta} \log ShRD_{js} - \frac{\log\left(B^A/B_s\right)}{\alpha \log \eta} + \frac{\rho + \zeta}{\alpha(\delta + g_j)}. \tag{48}$$

Figure 3 uses UK data to plot the inverse of $FiRD_j$ against negative $\log ShRD_{jUK}$. Consistent with (48), a positive linear relationship fits the data well for most industries. Moreover, departures from linearity are negatively correlated with industry growth rates g_j as predicted by (48). 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.[21] This shows the model can account for how firm-level R&D decisions differ across industries.

Data on the share of firms that perform R&D can also be used to assess whether cross-industry variation in estimated innovation-dependence is consistent with the model. Innovation-dependence is decreasing in the advantage of backwardness γ_j. Therefore, conditional on the localization of knowledge spillovers κ_j,

[21] Industry growth rates g_j are estimated using value-added volume per worker data for OECD countries. See Appendix C for details.

the model predicts that innovation-dependence is positively correlated with the share of firms that invest in R&D. The model does not provide an observable proxy for κ_j, but Figure 4 plots the innovation-dependence estimates from Table 1, column (c) against $\log ShRD_{jUK}$. The two variables have a positive relationship with a correlation coefficient of 0.41.[22] Assuming this correlation is not driven by unobserved variation in κ_j, Figure 4 supports the model's empirical credibility.

Out-of-sample comparative advantage test. The second validation exercise uses the innovation-dependence estimates from Section 4.3 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 9 European countries (see Appendix C for details). I then estimate the following variant of equation (47):

$$\log\left(\frac{EX_{js\tilde{s}}}{EX_{j\tilde{s}\tilde{s}}}\right) - (\sigma - 1)\log\left(\frac{w_{\tilde{s}}}{w_s}\right) = \xi CompAdv_{j\tilde{s}} + Controls_{js\tilde{s}} + \epsilon_{js\tilde{s}}, \tag{49}$$

where $\tilde{s}$ indexes the 9 out-of-sample countries, s indexes their trading partners, $CompAdv_{j\tilde{s}} = -(\sigma - 1)\frac{ID_j}{k}b_{\tilde{s}}$ equals the model-implied impact of R&D efficiency on comparative advantage and $Controls_{js\tilde{s}}$ denotes the same set of trade cost, productivity and comparative advantage controls used in column (c) of Table 1. I calculate $CompAdv_{j\tilde{s}}$ using the innovation-dependence estimates from column (c) of Table 1 and setting $\sigma - 1 = 6.53$. The model predicts the coefficient ξ of $CompAdv_{j\tilde{s}}$ should be positive and equal to one.

To estimate (49), I use pooled trade data from 2010-14 for 20 ISIC 2 digit manufacturing industries and 117 partner countries. The estimated value of ξ equals 0.74 with a standard error, clustered by importer-industry, of 0.060. Subject to the caveats that the estimation only uses R&D efficiency for 9 countries, and that the estimate is a little below one, this result shows 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.

[22]The Mining industry is not shown in Figure 4 because it is an outlier with a negative (though insignificant) estimated innovation-dependence. Including the Mining industry reduces the correlation with $\log ShRD_{jUK}$ to 0.39.

4.5 Calibration

The goal of the calibration is to quantify the variation in wages and incomes due to differences in R&D efficiency within the OECD. The calibration uses the 25 OECD countries from the baseline estimation sample and is based on 2012 data. In addition to the 22 goods industries for which Section 4.3 estimated innovation-dependence, I assume there is a non-tradable services industry. If industry j is non-tradable, then (45) yields:

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

Consequently, equations (35)-(37) imply 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. Since the innovation-dependence of services is unknown, I will primarily focus on nominal wage variation when discussing the calibration results. However, I also calculate real incomes under the assumption that the innovation-dependence of services equals zero. This assumption will underestimate variation in real incomes caused by differences in R&D efficiency if the innovation-dependence of services is positive.

Equilibrium wage and income levels can be calculated without calibrating all the model's parameters. To see this, recall from (32) that the R&D intensity of firms with positive R&D investment is given by:

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

Given $FiRD_j$ and the definition of Z_{js} in (45), equations (35) and (36) can be solved for w_s and a_s without knowing g_j. Therefore, 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 required for the calibration.

The parameters used in the calibration are obtained as follows (see Appendix C for further details). R&D efficiency, innovation-dependence and trade costs are taken from the estimates in Section 4.3. I use the innovation-dependence and trade costs estimates from column (c) of Table 1 and set the innovation-dependence of the Mining industry to zero. $FiRD_j$ is computed from UK firm-level data as in Section 4.4. Population levels L_s are taken from the Penn World Tables. I assume $\sigma - 1$ equals 6.53, consistent with the value of the trade elasticity used to estimate the model. 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.

I start by calibrating the model using observed R&D efficiency levels. Then, I calibrate a counterfactual economy in which R&D efficiency is constant across countries, but the other parameters are unchanged. To quantify the effect of R&D efficiency on nominal wages, I compute the log difference between calibrated nominal wages in these two cases. I will refer to this difference as the calibrated log wage. Similarly, calibrated log real income per capita is computed as the log difference between real income per capita in the two calibrations.[23]

Figure 5 plots the calibrated log wage against log nominal wages from the Penn World Tables, with both variables normalized to zero for the US. On average, countries where R&D efficiency is predicted to generate higher wages have larger observed wages. For example, the wage gap relative to the US due to differences in R&D efficiency is 32% for the Czech Republic and 14% for Italy. However, comparing the scales of the two axes, shows that R&D efficiency explains only a fraction of observed wage variation. Table 2 reports two statistics that formalize this observation (column a, row i). The estimated elasticity of calibrated to observed wages is 0.30. The ratio of the standard deviation of calibrated log wages to the standard deviation of observed log nominal wages is 0.36. These results imply R&D efficiency accounts for around one-third of observed wage variation within the OECD.

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 variables than in nominal outcomes. Figure 6 plots calibrated log real income per capita against observed log GDP per capita. In this case the estimated elasticity is 0.14 and the standard deviation ratio is 0.19 (column a, row ii). Thus, R&D efficiency can account for around one-sixth of observed income per capita variation in the OECD, even assuming it does not effect productivity in the services industry.

Columns (b)-(g) of Table 2 report a series of robustness checks on the baseline calibration results. For

[23] An alternative way to quantify the impact of R&D efficiency differences is to calibrate a counterfactual economy in which R&D efficiency is the only parameter that varies across countries. Implementing this approach leads to similar quantitative results.

each robustness check, I first estimate innovation-dependence levels and trade costs using a specification based on column (c) of Table 1. Given these estimates I then calibrate the model and calculate the effect of R&D efficiency on nominal wages and real incomes per capita using the same counterfactual method as before. When calibrating the model, I set innovation-dependence equal to zero in all industries where innovation-dependence is estimated to be either negative or insignificantly different from zero at the 10% level.[24]

The first robustness check in column (b) includes an additional 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. Average innovation-dependence declines slightly in this case, leading to a fall in the explanatory power of R&D efficiency, but the difference is small.

In column (c) I restrict innovation-dependence to be homogeneous across all industries (except services). In this case the estimated innovation-dependence is 0.29, a little below the average innovation-dependence of 0.32 from column (a), but the nominal wage and real income variation accounted for by R&D efficiency is effectively unchanged.

Section 3.5 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 and, consequently, is increasing in σ holding all else constant. Allowing for multiple sectors and trade costs enriches the analysis and necessitates the calibration of additional parameters. However, it does not change the intuition that, in addition to R&D efficiency and innovation-dependence, the trade elasticity is a key parameter for the quantification. Consequently, the remaining columns of Table 2 consider the impact of estimating and calibrating the model using alternative values of the trade elasticity $\sigma - 1$.

Column (d) reduces the trade elasticity to 2.5, column (e) uses an elasticity of 4.5, which is close to the aggregate elasticity estimated by Caliendo and Parro (2015), and column (f) increases the elasticity to 8.5. The results show that increasing the trade elasticity slightly reduces the average estimated innovation-dependence and this offsets the direct effect of increasing σ, implying that the explanatory power of R&D efficiency changes little.

[24] The only case where innovation-dependence is estimated to be significantly negative at the 10% level is for the Mining industry when I set the trade elasticity equal to 2.5.

Finally, column (g) uses the industry-specific trade elasticities estimated by Caliendo and Parro (2015). With these elasticities, average innovation-dependence increases to 0.38, but the elasticity of calibrated to observed wages declines to 0.24. Similarly, the standard deviation ratio for calibrated relative to observed log wages falls to 0.30. These declines occur because the innovation-dependence estimates are negatively correlated with the industry trade elasticities and because the changes in estimated trade costs compared to the baseline specification serve to reduce nominal wage variation caused by R&D efficiency differences. Combined, these effects more than offset the consequences of higher innovation-dependence. However, despite this, R&D efficiency still explains around one-quarter of nominal wage variation within the OECD. Looking at real income per capita, the offsetting forces are more closely balanced and R&D efficiency continues to explain around one-sixth of observed variation. Collectively, the robustness checks reinforce the conclusion that R&D efficiency accounts for a non-trivial share, although not the majority, of wage and income variation within the OECD.

5 Conclusions

There are persistent productivity differences across firms, industries and countries. These differences give rise to both Ricardian comparative advantage and international income inequality. To understand the origins and magnitude of productivity differences, this paper studies how innovation and learning determine equilibrium technology gaps. Some firms and countries are more innovative than others, which gives them a technological advantage, but the size of their advantage depends upon the advantage of backwardness and the geographic scope of knowledge spillovers. Countries with an absolute advantage in R&D have a comparative advantage in innovation-dependent industries, where the advantage of backwardness is lower and knowledge spillovers are more localized. The degree of innovation-dependence also determines the size of international wage and income gaps due to differences in R&D efficiency.

I estimate the model using R&D investment by country-industry and bilateral trade data. There is substantial heterogeneity in innovation-dependence across industries and, consistent with the model's predictions, innovation-dependence is positively correlated with R&D intensity. Using the structural estimates to calibrate the model, I find that differences in R&D efficiency account for around one-quarter to one-third of nominal wage variation within the OECD. This shows that, although spillovers and learning do reduce productivity dispersion, technology gaps still play an important role in explaining cross-country productivity and income differences.

This paper focuses on trade and technology gaps, but the theory and empirical methodology it develops are applicable to other important questions. For example, industrial policy is frequently justified as a tool to promote exports (or reduce imports) in particular industries. The model could be used to analyze how the effects of industrial policy depend upon R&D efficiency, knowledge spillovers, the advantage of backwardness and trade costs. It would also be interesting to extend the model to allow firms to invest in international technology transfer through foreign direct investment or technology licensing. Alternatively, embedding the R&D and adoption technologies in a model with differentiated products and monopolistic competition would enable an analysis of the interactions between technology investment and selection in a dynamic Melitz (2003) environment.

Finally, the framework is suitable for analyzing not only cross-sectional variation on a balanced growth path, but also changes in the global economy. Numerical solutions could be used to study the transition dynamics that occur when, for example, a country's R&D efficiency improves. The empirical strategy could also be adapted to estimate whether R&D efficiency and innovation-dependence have changed over time. This would shed light on whether recent advances in information and communication technologies have reduced international technology gaps and cross-country income dispersion. Through such applications, the methods developed in this paper could further contribute to understanding the global economy.

References

Acemoglu, Daron. 2009. "Introduction to Modern Economic Growth." Princeton University Press.

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

Acharya, Ram C., and Wolfgang Keller. 2009. "Technology Transfer through Imports." *Canadian Journal of Economics* 42 (4): 1411-1448.

Aghion, Philippe, and Peter Howitt. 1992. "A Model of Growth through Creative Destruction." *Econometrica* 60: 323-351.

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. 2016. "Growth through Heterogeneous Innovations." *Journal of Political*

Economy, forthcoming.

Allen, Treb, Costas Arkolakis, and Xiangliang Li. 2015. "On the Existence and Uniqueness of Trade Equilibria." mimeo, Yale 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.

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

Barro, Robert J., and Xavier Sala-i-Martín. 1997. "Technological Diffusion, Convergence and Growth." *Journal of Economic Growth* 2 (1): 1-26.

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.

Basu, Susanto, and David N. Weil. 1998. "Appropriate Technology and Growth." *Quarterly Journal of Economics* 113 (4): 1025-1054.

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.

Bernard, Andrew B., Jonathan Eaton, J. Bradford Jensen, and Samuel Kortum. 2003. "Plants and Productivity in International Trade." *American Economic Review* 93 (4): 1268-1290.

Bonfiglioli, Alessandra, Rosario Crinò, and Gino A. Gancia. 2015. "Betting on Exports: Trade and Endogenous Heterogeneity." CESifo Working Paper No. 5597.

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. 2016. "The Global Diffusion of Ideas." NBER Working Paper No. 21844.

Cai, Jie, Nan Li, and Ana Maria Santacreu. 2019. "Knowledge Diffusion, Trade, and Innovation across Countries and Sectors." mimeo, Federal Reserve Bank of St. Louis.

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.

Chor, Davin. 2010. "Unpacking Sources of Comparative Advantage: A Quantitative Approach." *Journal of International Economics* 82 (2): 152-167.

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, 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.

Desmet, Klaus, and Esteban Rossi-Hansberg. 2014. "Spatial Development." *American Economic Review* 104 (4): 1211-1243.

Desmet, Klaus, Dávid Krisztián Nagy, and Esteban Rossi-Hansberg. 2016. "The Geography of Development." *Journal of Political Economy*, forthcoming.

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

Dornbusch, Rudiger, Stanley Fischer, and Paul Anthony Samuelson. 1977. "Comparative Advantage, Trade, and Payments in a Ricardian Model with a Continuum of Goods." *American Economic Review* 67 (5): 823-839.

Durkin Jr., John T. 1997. "Perfect Competition and Endogenous Comparative Advantage." *Review of International Economics* 5 (3): 401-411.

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.

Foster, Lucia, John C. Haltiwanger, and Cornell John Krizan. 2001. "Aggregate Productivity Growth.

Lessons from Microeconomic Evidence," in *New Developments in Productivity Analysis*, Charles R. Hulten, Edwin R. Dean, and Michael J. Harper, eds. (Chicago, IL: University of Chicago Press), 303-72.

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. 2015. "How Destructive is Innovation?" mimeo, Stanford University.

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. 1990. "Explaining Japan's Innovation and Trade: A Model of Quality Competition and Dynamic Comparative Advantage." *Monetary and Economic Studies* 8 (2): 75-100.

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, 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.

Impullitti, Giammario, and Omar Licandro. 2016. "Trade, Firm Selection, and Innovation: The Competition Channel." mimeo, University of Nottingham.

Jaffe, Adam B., Manuel Trajtenberg, and Rebecca Henderson. 1993. "Geographic Localization of Knowledge Spillovers as Evidenced by Patent Citations." *Quarterly Journal of Economics* 108 (3): 577-598.

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.

Klette, Tor Jakob, and Samuel Kortum. 2004. "Innovating Firms and Aggregate Innovation." *Journal of Political Economy* 112 (5): 986-1018.

König, Michael, Jan Lorenz, and Fabrizio Zilibotti. 2016. "Innovation vs Imitation and the Evolution of Productivity Distributions," *Theoretical Economics*, forthcoming.

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. 1987 "The Narrow Moving Band, the Dutch Disease, and the Competitive Consequences of Mrs. Thatcher: Notes on Trade in the Presence of Dynamic Scale Economies." *Journal of Development Economics* 27 (1-2): 41-55.

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

Lu, Chia-Hui. 2007. "Moving Up or Moving Out? A Unified Theory of R&D, FDI, and Trade." *Journal of International Economics* 71 (2): 324-343.

Lucas, Robert E., and Benjamin Moll. 2014. "Knowledge Growth and the Allocation of Time." *Journal of Political Economy*, 122, 1-51.

Luttmer, Erzo G.J. 2007. "Selection, Growth, and the Size Distribution of Firms." *Quarterly Journal Of Economics* 122 (3): 1103-1144.

Malerba, Franco, Maria Luisa Mancusi, and Fabio Montobbio. 2013. "Innovation, International R&D Spillovers and the Sectoral Heterogeneity of Knowledge Flows." *Review of World Economics* 149 (4): 697-722.

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.

Nelson, Richard R., and Edmund S. Phelps. 1966. "Investment in Humans, Technological Diffusion, and Economic Growth." *American Economic Review* 56 (1/2): 69-75.

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.

Ohlin, Bertil. 1933. "Interregional and International Trade." Cambridge, MA: Harvard University Press.

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. 2015. "Equilibrium Technology Diffusion, Trade, and Growth." NBER Working Paper No. 20881.

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

Ricardo, David. 1817. "Principles of Political Economy and Taxation."

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. 2016. “Comparative Advantage in Innovation and Production.” mimeo, Federal Reserve Board.

Taylor, M. Scott. 1993. “‘Quality Ladders’ and Ricardian Trade.” *Journal of International Economics* 34 (3): 225-243.

Vernon, Raymond. 1966. “International Investment and International Trade in the Product Cycle.” *Quarterly Journal of Economics* 80 (2): 190-207.

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

Appendices

A Proofs and Derivations

Analysis of growth rates in Section 3.1

On a balanced growth path the individual's budget constraint (1) 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 (20) into the asset market clearing condition (24) 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 (12). 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 (7) 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 (5) 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 and equation (26) holds. The numeraire condition (25) 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.

To obtain $\iota_s = \rho$, substitute (53) into the Euler equation (2). Note also that using (4) to substitute for X_{js} in (6) 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 (23) that industry output Y_{js} also grows at rate g_j.

Solution to firm's R&D problem in Section 3.2

The firm faces a discounted infinite-horizon optimal control problem of the type studied in Section 7.5 of Acemoglu (2009). The current-value Hamiltonian is:

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

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

$$0 = \frac{\partial \mathcal{H}}{\partial l^R} = -w_s + \lambda \frac{\alpha}{1-\beta} \psi B_s \phi^{1-\gamma_j(1-\beta)} \left(l^R\right)^{\alpha-1}, \tag{54}$$

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

$$0 = \lim_{\tilde{t}\to\infty} \left[e^{-(\rho+\zeta)(\tilde{t}-t)} \mathcal{H}(\phi, l^R, \lambda) \right], \tag{55}$$

where equation (55) is the transversality condition. Differentiating the upper expression with respect to τ gives:

$$(1-\alpha) \frac{\dot{l}^R}{l^R} = [1-\gamma_j(1-\beta)] \frac{\dot{\phi}}{\phi} + \frac{\dot{\lambda}}{\lambda}, \tag{56}$$

and using the first order conditions of the Hamiltonian to substitute for λ and $\dot{\lambda}$, and (28) to substitute for $\dot{\phi}$ we obtain equation (29).

Equations (28) and (29) are an autonomous nonlinear system of differential equations in (ϕ, l^R) whose unique steady state $(\phi^*_{js}, l^{R*}_{js})$ is given by (30) and (31). 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:

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

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

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

which is negative by Assumption 1. This means the Jacobian has one strictly negative and one strictly positive eigenvalue. Therefore, by Theorem 7.19 in Acemoglu (2009), the steady state is locally saddle-path stable. There exists an open neighborhood of the steady state such that if the firm's initial ϕ lies within this neighborhood, the system of differential equations given by (28) and (29) 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 (56) 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 (55).

The solution to (28) and (29) 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 (28) and (29). Equation (54) 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:

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

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 (28) and (29) is the unique solution to the firm's optimal control problem.

Derivation of balanced growth path equilibrium equations (35)-(37)

Suppose the global economy is on a balanced growth path. Using (11), (17), (30) and (31) implies that on a balanced growth path the steady state value of a firm with capability $\psi \geq \psi_{js}^*$ is:

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

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 3.2 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 (20) yields:

$$f^E = \left(1-\beta-\frac{\alpha\left(\delta+g_j\right)}{\rho+\zeta+\gamma_j\left(\delta+g_j\right)}\right)\frac{\Psi_{js}}{\rho+\zeta}\left[\alpha^\alpha\beta^{\gamma_j\beta}B_s\left(\frac{p_{js}\chi^R_{js}}{w_s}\right)^{\gamma_j}\frac{(\delta+g_j)^{\alpha-1}}{\left[\rho+\zeta+\gamma_j(\delta+g_j)\right]^\alpha}\right]^{\frac{1}{\gamma_j(1-\beta)-\alpha}}. \tag{57}$$

Next, observe that on a balanced growth path:

$$\int_\theta \theta^{\frac{1}{1-\beta}}dH_{js}(\theta) = \int_{\psi^{\min}}^{\psi^{\max}}\left(\chi^R_{js}\right)^{\frac{1}{1-\beta}}\phi^*_{js}dG(\psi),$$

where ϕ^*_{js} is given by (30) for R&D firms and by (30) with $\psi = \psi^*_{js}$ for adopters. Thus, by substituting (12) and (31) into the labor market clearing condition (22) and using (21) with $\dot{M}_{js} = 0$ to solve for L^E_{js} we obtain:

$$\begin{aligned} L_s = \sum_{j=1}^{J} M_{js}\Bigg\{&\left(1+\frac{\alpha}{\beta}\frac{\delta+g_j}{\rho+\zeta+\gamma_j(\delta+g_j)}\right)\Psi_{js} \\ &\times\left[\alpha^\alpha\beta^{\gamma_j-\alpha}B_s\left(\frac{p_{js}\chi^R_{js}}{w_s}\right)^{\gamma_j}\frac{(\delta+g_j)^{\alpha-1}}{\left[\rho+\zeta+\gamma_j(\delta+g_j\right]^\alpha}\right]^{\frac{1}{\gamma_j(1-\beta)-\alpha}} + f^E\zeta\Bigg\}. \end{aligned} \tag{58}$$

Similarly, substituting (4), (6), (13) and (30) into the goods market clearing condition (23) and using (57) we obtain:

$$\sum_{\tilde{s}=1}^{S}\left(\frac{\tau_{js\tilde{s}}p_{js}}{P_{j\tilde{s}}}\right)^{1-\sigma}\mu_j z_{\tilde{s}}c_{\tilde{s}}L_{\tilde{s}} = f^E(\rho+\zeta)\left(1-\beta-\frac{\alpha(\delta+g_j)}{\rho+\zeta+\gamma_j(\delta+g_j)}\right)^{-1}M_{js}w_s. \tag{59}$$

We showed in Section (3.1) that on a balanced growth path $\dot{a}_s = 0$ and $\iota_s = \rho$. Therefore, the individual's budget constraint implies:

$$z_s c_s = \rho a_s + w_s, \tag{60}$$

while substituting the free entry condition (20) into the asset market clearing condition (24) gives:

$$a_s L_s = \sum_{j=1}^{J} M_{js} w_s f^E. \tag{61}$$

Equations (57)-(61) together with consumption prices (5), industry price indices (7), R&D knowledge levels (18) and knowledge capital growth rates (19) form a system of $4JS + 4S + J$ equations. Together with the numeraire condition (25), the steady state relative productivity levels in (30) 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 (57) and (59) into (58) giving:

$$L_s = \sum_{j=1}^{J} \frac{\mu_j}{\rho + \zeta} \left(\zeta + \beta\rho + \frac{\alpha\rho(\delta + g_j)}{\rho + \zeta + \gamma_j(\delta + g_j)} \right) \sum_{\tilde{s}=1}^{S} \left(\frac{\tau_{js\tilde{s}} p_{js}}{P_{j\tilde{s}}} \right)^{1-\sigma} \frac{z_{\tilde{s}} c_{\tilde{s}} L_{\tilde{s}}}{w_s}. \tag{62}$$

Using (7) to obtain the industry price index, (57) to substitute for p_{js}, (18) to give χ_{js}^R and (30) to solve for relative steady state productivity levels then implies:

$$\left(\frac{p_{js}}{P_{j\tilde{s}}} \right)^{1-\sigma} = \frac{w_s^{1-\sigma} \left(B_s \Psi_{js}^{\frac{\gamma_j(1-\beta)}{1+\kappa_j} - \alpha} \right)^{\frac{(\sigma-1)(1+\kappa_j)}{\gamma_j}}}{\sum_{\hat{s}=1}^{S} \tau_{j\hat{s}\tilde{s}}^{1-\sigma} w_{\hat{s}}^{1-\sigma} \left(B_{\hat{s}} \Psi_{j\hat{s}}^{\frac{\gamma_j(1-\beta)}{1+\kappa_j} - \alpha} \right)^{\frac{(\sigma-1)(1+\kappa_j)}{\gamma_j}}}. \tag{63}$$

Substituting this expression into (62) and using (60) yields equation (35). Equation (36) can be derived in a similar manner by substituting (59) and (63) into the asset market clearing condition (61). Finally, substituting steady state R&D employment (31) together with (57), (59) and (63) into (19) yields equation (37).

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

I will start by proving that, under free trade, equations (35) and (36) yield a unique solution for w_s and a_s given growth rates g_j.

Free trade implies $\tau_{js\tilde{s}} = 1$ for all $j, s, \tilde{s}$ and using the numeraire condition (25) and equation (60) we have:

$$\sum_{\tilde{s}=1}^{S} \tau_{js\tilde{s}}^{1-\sigma} \left(\rho a_{\tilde{s}} + w_{\tilde{s}}\right) L_{\tilde{s}} = 1.$$

Using this result in (38) gives:

$$Z_{js} = \frac{w_s^{-\sigma} \left(B_s \Psi_{js}^{\frac{\gamma_j(1-\beta)}{1+\kappa_j} - \alpha} \right)^{\frac{(\sigma-1)(1+\kappa_j)}{\gamma_j}}}{\sum_{\hat{s}=1}^{S} w_{\hat{s}}^{1-\sigma} \left(B_{\hat{s}} \Psi_{j\hat{s}}^{\frac{\gamma_j(1-\beta)}{1+\kappa_j} - \alpha} \right)^{\frac{(\sigma-1)(1+\kappa_j)}{\gamma_j}}}, \tag{64}$$

and substituting this expression into (35) 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:

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

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 $F_s(\tilde{\mathbf{w}}, w_s)$ is strictly increasing in $\tilde{w}_{\hat{s}}$ for all $\hat{s}$.

Now, given $\lambda>0$ and $\tilde{\mathbf{w}} \in \mathbb{R}_{++}^S$ define $w_s(\lambda)$ by $F_s[\lambda\tilde{\mathbf{w}}, w_s(\lambda)]=0$. Let $u \in (0,1)$ be such that $-1+\sigma u<0$. Then $F_s\left[\lambda\tilde{\mathbf{w}}, \lambda^{1-u}w_s(1)\right]$ is strictly negative if $\lambda>1$ and strictly positive if $\lambda<1$. Since $F_s(\tilde{\mathbf{w}}, w_s)$ is strictly decreasing in w_s it follows that $w_s(\lambda)<\lambda^{1-u}w_s(1)$ if $\lambda>1$ and $w_s(\lambda)>\lambda^{1-u}w_s(1)$ if $\lambda<1$. Therefore, when $\lambda\to\infty$, $\frac{\lambda}{w_s(\lambda)}\to\infty$ 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 $f_s(\mathbf{w})$ is strictly increasing in $w_{\hat{s}}$ whenever $\hat{s}\neq s$, $\mathbf{f}(\mathbf{w})$ satisfies gross substitution. Also, $f_s(\mathbf{w})$ can be written as $f_s(\mathbf{w})=\tilde{f}_s(\mathbf{w})-L_s$ where $\tilde{f}_s(\mathbf{w})$ is positive and homogeneous of degree minus one, while L_s is positive and homogeneous of degree zero in $\mathbf{w}$. Consequently, Theorem 2 in Allen, Arkolakis and Li (2015) implies the solution is unique.

Using the solution for wages and equation (64) for Z_{js}, assets a_s are given immediately by (36). 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. The discussion in the main text established that when $J=1$ there exists a unique equilibrium growth rate g. It follows immediately that, if $J=1$ and there are no trade costs, the global economy has a unique balanced growth path.

Proof of Proposition 2

To derive (39) start by substituting the free entry condition (57) into (30) and using $\theta_{js}^*=\chi_{js}^R\left(\phi_{js}^*\right)^{1-\beta}$ to obtain:

$$\left(\theta_{js}^*\right)^{\frac{1}{1-\beta}}=\left[f^E\frac{\rho+\zeta}{(\delta+g_j)^{\frac{1}{\alpha}}}\left(\frac{1-\beta}{\alpha}\frac{\rho+\zeta+\gamma_j(\delta+g_j)}{\delta+g_j}-1\right)^{-1}\frac{B_s^{\frac{1}{\alpha}}}{\Psi_{js}}\right]^{\frac{\alpha}{\gamma_j(1-\beta)}}\psi^{\frac{1}{\gamma_j(1-\beta)-\alpha}}\left(\chi_{js}^R\right)^{\frac{1}{1-\beta}}, \tag{65}$$

where $\psi=\psi_{js}^*$ for firms that choose adoption. Setting $\psi=\psi^{\max}$ in this expression and using (18) 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 (18) back into (65) and integrating over the capability distribution yields:

$$\overline{\theta}_{js}^{*} = \left[f^E \frac{\rho+\zeta}{(\delta+g_j)^{\frac{1}{\alpha}}} \left(\frac{1-\beta}{\alpha} \frac{\rho+\zeta+\gamma_j(\delta+g_j)}{\delta+g_j} - 1 \right)^{-1} \right]^{\frac{\alpha(1+\kappa_j)}{\gamma_j}} \left(\psi^{\max}\right)^{\frac{(1-\beta)\kappa_j}{\gamma_j(1-\beta)-\alpha}} \chi_j \left(B_s \Psi_{js}^{\frac{\gamma_j(1-\beta)}{1+\kappa_j}-\alpha} \right)^{\frac{1+\kappa_j}{\gamma_j}}, \tag{66}$$

and dividing this equation by the equivalent expression for country $\tilde{s}$ gives (39).

Using (4) and (6) 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 (63) into this expression and taking logs we obtain equation (41) where:

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

and substituting (66) into this expression gives equation (40) where:

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

Next, differentiating the definition of Ψ_{js} and using that the R&D threshold ψ_{js}^* is given by (27) 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 (66) 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:

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

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.

Derivation of balanced growth path consumption prices from Section 3.5

From (5) and (7) 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 (18), (57) and (65) 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:

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

B Model Extensions

This appendix shows that the theoretical and quantitative results derived in Sections 3 and 4 are robust to allowing for productivity differences not caused by variation in R&D efficiency and to allowing the adoption technology to depend upon the quality of a country's national innovation system. I introduce two extensions of the baseline model. First, suppose there are exogenous productivity differences at the country-industry level. Instead of (8), 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, assume technology adoption is more efficient in countries with higher B_s. This assumption is consistent with evidence that adoption and innovation draw upon similar capabailities (Rosenberg 1990). Instead of (15), suppose the adoption technology is given by:

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

where $\nu \in [0, 1)$. Imposing $\nu < 1$ ensures adoption efficiency is less sensitive to the quality of a country's national innovation system than R&D efficiency.

With these two generalizations the model can be solved using the same series of steps detailed in Section 3. The main differences from the baseline model are as follows. The R&D threshold (27) is now given by:

$$\psi_{js}^* = \frac{\eta^{\gamma_j}}{B_s^{1-\nu}}.$$

Steady state relative productivity and technology investment employment are still given by (30) and (31), respectively, except that p_{js} is multiplied by A_{js} in both equations. The general equilibrium conditions (35)-(37) are unchanged other than that the definition of Z_{js} becomes:

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

Crucially, relative average steady state firm productivity levels are still given by (39), 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 (40) is replaced by:

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

Together these observations imply that all the main theoretical results from the baseline model continue to hold, including Propositions 1, 2 and 3. However, allocative efficiency differences generate additional variation in trade and income levels, while the elasticity of the R&D threshold ψ^*_{js} to B_s is lower than in the baseline model.

Taking a first order approximation to the generalized model for large ψ^*_{js} yields equations (42) and (43), but with B^A replaced by B_s^ν. It follows that using (46) to obtain R&D efficiency gives $k(1-\nu)\log B_s$ up to an additive constant.

Calculating innovation-dependence using the approximation to Ψ_{js} yields:

$$ID_j = \frac{\kappa_j(1-\beta)}{\gamma_j(1-\beta)-\alpha} + \frac{\nu}{\gamma_j}\frac{\gamma_j(1-\beta)-\alpha(1+\kappa_j)}{\gamma_j(1-\beta)-\alpha}.$$

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. Moreover, innovation-dependence is also increasing in ν because a higher ν implies that R&D efficiency matters more for adoption. Using this new expression for innovation-dependence, bilateral exports can still be written as (44), but with $(\sigma-1)\log A_{js}$ added to the right hand side. Consequently, by estimating (47) we can recover $\frac{ID_j}{k(1-\nu)}$. This implies the estimates in Table 1 should still be interpreted as estimates of innovation-dependence in the generalized model, only the scaling constant has changed.

Finally, instead of (45), Z_{js} satisfies:

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

As in the baseline model, Z_{js} is a function of $B_s^{(\sigma-1)ID_j}$. It follows that wage and income differences due to variation in R&D efficiency can be quantified without calibrating either A_{js} or ν and that the calibration results from Section 4.5 have the same interpretation in the generalized model as in the baseline case.

C Data

R&D: R&D intensity is the ratio of business R&D expenditure in the OECD's ANBERD database to industry value-added in the OECD's STAN database for 2 digit ISIC Revision 4 manufacturing industries. 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.

Trade: Bilateral trade for 2 digit ISIC Revision 4 goods industries is from the OECD's STAN Bilateral Trade by Industry and End-use database. Sales of domestic production to the domestic market 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 and converted to US dollars using exchange rates from the IMF's International Financial Statistics.

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 can be calculated using the Penn World Tables 9.0. 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. 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.

Country-level variables: GDP per capita, population, nominal wages, physical capital per employee and human capital are from the Penn World Tables 9.0. GDP per capita is defined as output-side real GDP at chained purchasing power parties (PPPs) divided by population. The nominal wage is calculated as labor's share of GDP times output-side GDP at current PPPs times the price level of current GDP divided by persons engaged. Physical capital per employee is given by the capital stock at current PPPs divided by persons engaged.

The Worldwide Governance Indicators are from the World Bank. 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. 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. All these variables are time-varying.

Firm-level R&D: The share of firms that perform R&D is computed from the UK's Annual Business Survey, which is a representative sample of production, construction, distribution and service industries. Firms are asked whether they have "plans to carry out in-house Research and Development during the next two years". For each 2 digit goods industry I compute the weighted share (using sampling weights) of respondents that answer yes to this question and report the average share for 2008-09. The data is reported for UK SIC 2007 industries, which corresponds to ISIC Revision 4. The data does not cover Northern Ireland.

To measure R&D intensity, I match the Annual Business Survey with the Business Expenditure on Research and Development data set and compute R&D intensity as the ratio of total R&D expenditure to approximate gross value-added at basic prices. Industry R&D intensity is then calculated as the median of all firm-level observations pooled for 2008-09. 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.

Industry growth rates are estimated using OECD STAN data on value-added volumes per person engaged from 1995-2014. 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.

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. 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 calculated using equation (46), but with data from 2008-15 including 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. 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 nominal wages can be calculated using the Penn World Tables 9.0.

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 and converted to US dollars using exchange rates from the IMF's International Financial Statistics. Imports and exports by industry are from the OECD's STAN Bilateral Trade by Industry and End-use database. 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.

Caliendo and Parro (2015) estimate trade elasticities for ISIC Revision 3 goods sectors at approximately the 2 digit level of aggregation. I take their benchmark estimates from the 99% sample in 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. Caliendo and Parro estimate the trade elasticity for Petroleum is 64.85, around four times larger than for the industry with the second highest elasticity. Using this elasticity delivers very high trade cost estimates in the Petroleum industry for some country-pairs. Consequently, to facilitate solving the model numerically, I censor the calibrated trade costs in Petroleum at the maximum estimated trade cost across all other industries.

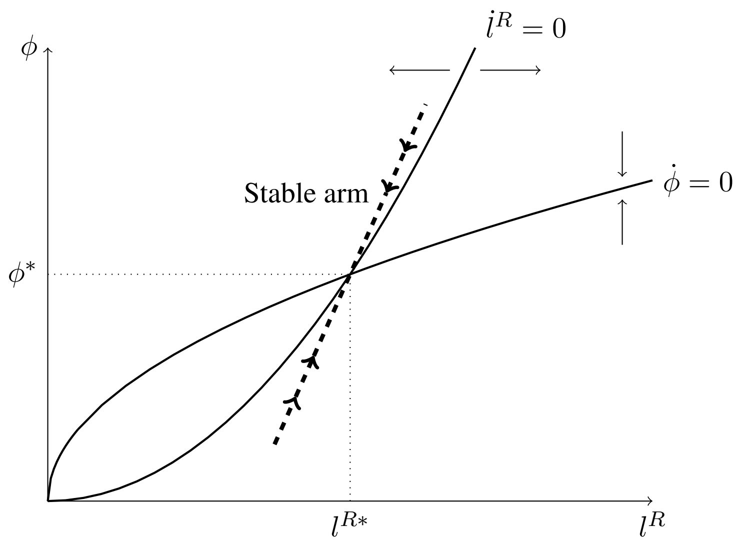

Figure 1: Firm steady state and transition dynamics

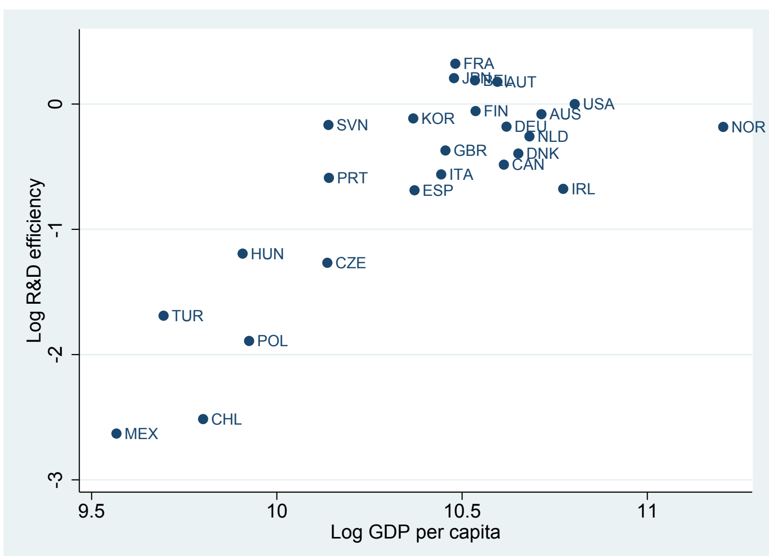

Figure 2: R&D efficiency

Notes: R&D efficiency for 2010-14 calculated using OECD's ANBERD and STAN databases. GDP per capita in 2010 from Penn World Tables 9.0.

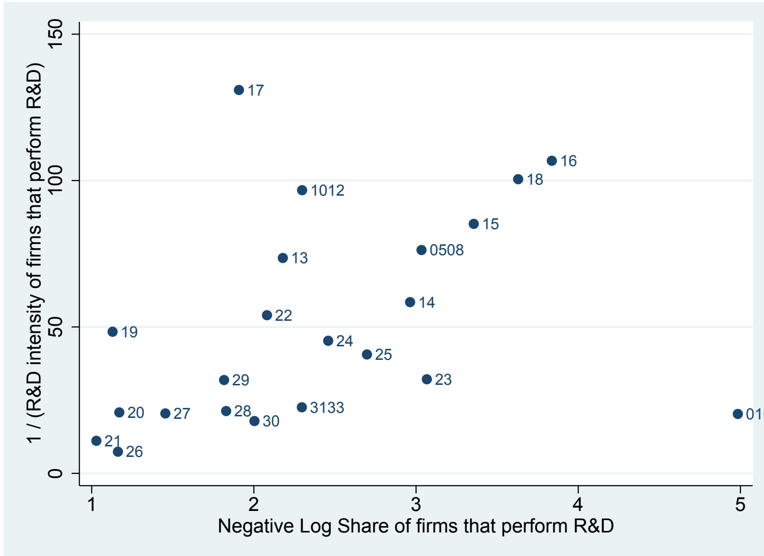

Figure 3: Firm-level R&D investment

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.

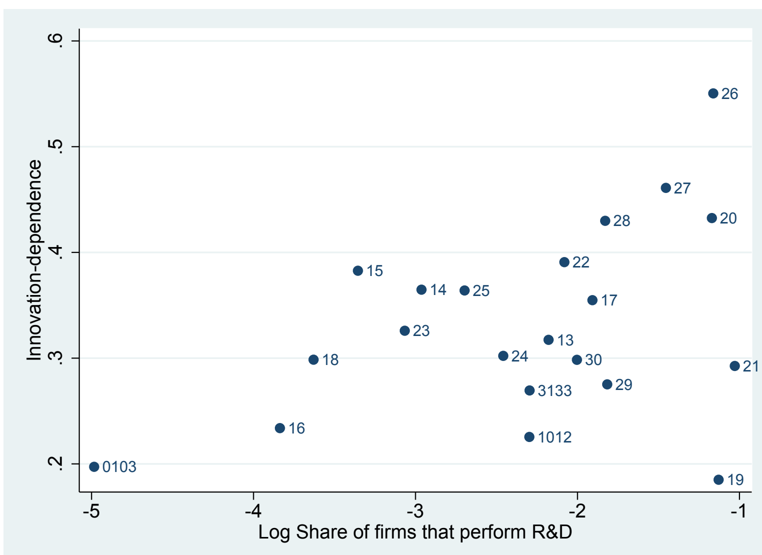

Figure 4: Innovation-dependence and selection into R&D

Notes: Innovation-dependence estimates from Table 1, column (c). Share of firms that invest in R&D computed for UK industries from Office for National Statistics' Annual Business Survey 2008-09. Industries are 2 digit ISIC Revision 4 goods industries.

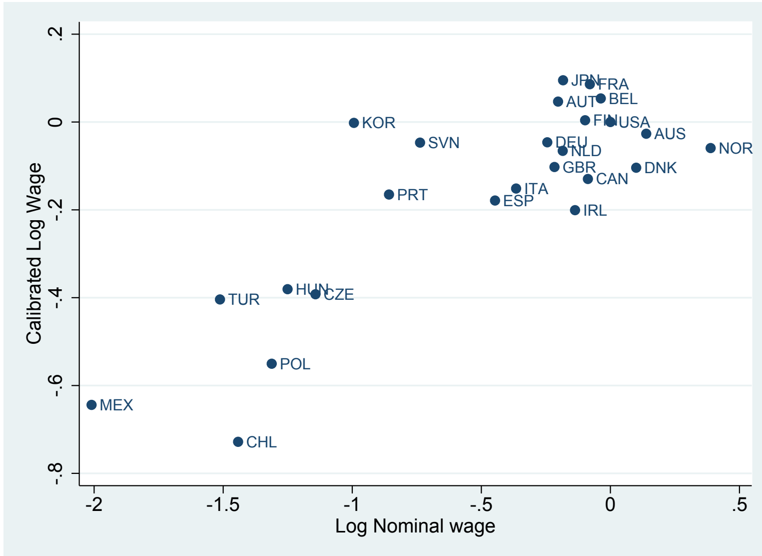

Figure 5: R&D efficiency and nominal wages

Notes: Calibrated log wage from author's calculations. Nominal wages in 2012 from Penn World Tables 9.0. Variables normalized to zero for the US.

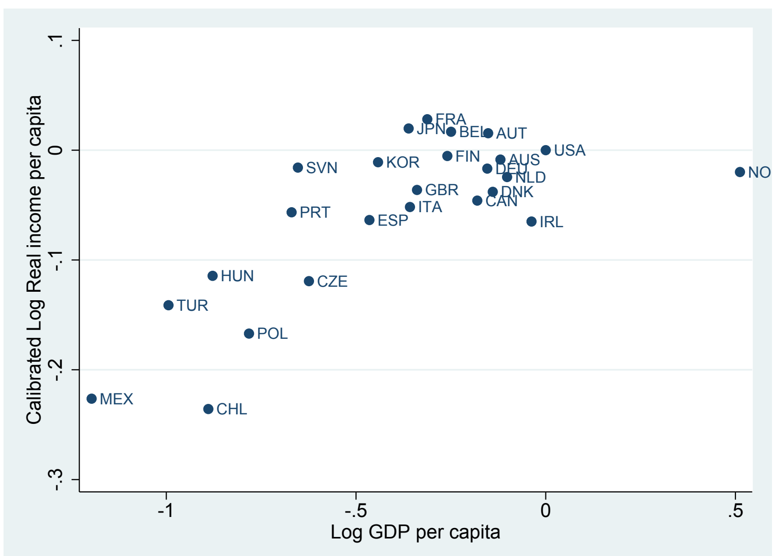

Figure 6: R&D efficiency and real incomes

Notes: Calibrated log real income per capita from author's calculations. GDP per capita in 2012 from Penn World Tables 9.0. Variables normalized to zero for the US.

Table 1: Innovation-dependence by industry

	Innovation-dependence estimates		
Industry	(a)	(b)	(c)
Agriculture, forestry and fishing (0103)	0.517	0.365	0.197
	(0.0578)	(0.0402)	(0.0832)
Mining and quarrying (0508)	0.437	0.283	-0.101
	(0.0893)	(0.0624)	(0.134)
Food products, beverages and tobacco (1012)	0.535	0.383	0.225
	(0.0481)	(0.0383)	(0.0733)
Textiles (13)	0.578	0.449	0.317
	(0.0470)	(0.0439)	(0.0587)
Wearing apparel (14)	0.564	0.402	0.365
	(0.0521)	(0.0460)	(0.0496)
Leather and related products (15)	0.536	0.407	0.383
	(0.0493)	(0.0551)	(0.0733)
Wood and of products of wood and cork, except furniture (16)	0.582	0.428	0.234
	(0.0683)	(0.0362)	(0.0609)
Paper and paper products (17)	0.633	0.472	0.355
	(0.0590)	(0.0338)	(0.0593)
Printing and reproduction of recorded media (18)	0.632	0.483	0.299
	(0.0602)	(0.0335)	(0.0619)
Coke and refined petroleum products (19)	0.534	0.388	0.185
	(0.0531)	(0.0427)	(0.0816)
Chemicals and chemical products (20)	0.651	0.500	0.433
	(0.0521)	(0.0452)	(0.0839)
Basic pharmaceutical products and pharmaceutical preparations (21)	0.698	0.527	0.293
	(0.0783)	(0.0604)	(0.134)
Rubber and plastics products (22)	0.658	0.501	0.391
	(0.0576)	(0.0334)	(0.0451)
Other non-metallic mineral products (23)	0.626	0.473	0.326
	(0.0540)	(0.0324)	(0.0518)
Basic metals (24)	0.648	0.456	0.302
	(0.0430)	(0.0424)	(0.0808)
Fabricated metal products, except machinery and equipment (25)	0.658	0.503	0.364
	(0.0596)	(0.0319)	(0.0505)
Computer, electronic and optical products (26)	0.758	0.520	0.550
	(0.0432)	(0.0400)	(0.155)
Electrical equipment (27)	0.701	0.565	0.461
	(0.0799)	(0.0569)	(0.102)
Machinery and equipment n.e.c. (28)	0.785	0.629	0.430
	(0.0800)	(0.0451)	(0.0999)
Motor vehicles, trailers and semi-trailers (29)	0.622	0.420	0.275
	(0.0538)	(0.0370)	(0.0956)
Other transport equipment (30)	0.660	0.419	0.298
	(0.0918)	(0.0570)	(0.122)
Furniture, other manufacturing (3133)	0.601	0.445	0.269
	(0.0694)	(0.0352)	(0.0599)
Trade cost controls	Yes	Yes	Yes
Productivity level controls	No	Yes	Yes
Comparative advantage controls	No	No	Yes
Observations	181,923	181,923	181,923
R-squared	0.580	0.701	0.726
F test innovation-dependence equal across industries (p-value)	0.0218	0.000	0.0639
Average innovation-dependence	0.619	0.455	0.311

Estimates give innovation-dependence relative to the shape parameter of the R&D capability distribution. Standard errors clustered by importer-industry in parentheses. 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 agreeent. 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: Calibration results

		Baseline (a)	GDP per capita (b)	Homogeneous innovation-dependence (c)	Low trade elasticity (d)	Moderate trade elasticity (e)	High trade elasticity (f)	Industry trade elasticities (g)
(i) Nominal wage	Elasticity	0.300	0.279	0.300	0.305	0.310	0.283	0.238
	Standard deviation ratio	0.363	0.331	0.353	0.366	0.369	0.345	0.304
(ii) Real income per capita	Elasticity	0.144	0.125	0.156	0.123	0.135	0.141	0.141
	Standard deviation ratio	0.195	0.167	0.208	0.167	0.182	0.192	0.197
(iii) Innovation-dependence	Average	0.316	0.277	0.295	0.340	0.324	0.306	0.379
	Standard deviation	0.113	0.123		0.348	0.184	0.091	0.258
	Correlation with baseline		0.924		0.848	0.945	0.974	0.548
(iv) Trade elasticity		6.53	6.53	6.53	2.5	4.5	8.5	Industry

Row (i) gives the elasticity of calibrated log wages to observed log nominal wages, and the ratio of the standard deviation of calibrated log wages to the standard deviation of observed log nominal wages. Row (ii) gives the same statistics for real income per capita. Row (iii) reports summary statistics on the innovation-dependence estimates for goods industries used in the calibration. Correlation with baseline is the correlation with the baseline estimates used in column (a). Row (iv) shows the trade elasticity used in the calibration. Column (a) uses estimates from Table 1, column (c) to calibrate the model. For column (b) innovation-dependence is estimated including the interaction of industry dummy variables with the importer's log GDP per capita as an additional control. In column (c) innovation-dependence is restricted to be the same in all goods industries. Industry-specific trade elasticities used for column (g) from Caliendo and Parro (2015).