The Global Productivity Distribution 
and Ricardian Comparative Advantage∗ 

Thomas Sampson† 
London School of Economics 

May 2017 

Abstract 

This paper studies the origins of Ricardian comparative advantage. I develop an endogenous growth model where productivity differences across firms and countries emerge from R&D investment by incumbent firms with heterogeneous R&D capabilities. Less productive firms have an advantage of backwardness and spillovers depend upon the productivity frontier in each country. The theory characterizes how knowledge spillovers and learning affect the global productivity distribution, international income inequality and the growth rate. Countries with an absolute advantage in R&D, or with fewer barriers that restrict knowledge spillovers, develop a comparative advantage in industries where the advantage of backwardness is lower and where spillovers are more localized.

∗I am grateful to seminar participants in Barcelona, Bologna and the ISGEP workshop in Pescara for helpful discussions. 
†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

The idea productivity differences across countries and industries shape the pattern of international trade dates from Ricardo (1817). More recently, important theoretical advances have shown how Ricardo’s insight can be formalized when there are more than two industries (Dornbusch, Fischer and Samuelson 1977) and in an economy with many countries (Eaton and Kortum 2002). Contemporary formulations of the theory of Ricardian comparative advantage predict countries have higher relative exports in industries where they have higher relative productivity. Empirical work finds strong support for this prediction and implies productivity differences are an important determinant of the pattern of trade (Golub and Hsieh 2000; Costinot, Donaldson and Komunjer 2012; Kerr 2013).

Less is known about why international productivity gaps differ by industry. Most studies of comparative advantage take productivity levels as given and analyze the consequences of productivity differences, not the causes. In contrast, this paper studies the origins of productivity differences. The productivity of an industry depends upon the productivity of its firms. In turn, the productivity of each firm depends upon both the firm’s past innovations and what the firm has learnt from other domestic and foreign firms. Thus, innovation and learning shape comparative advantage as highlighted by Posner (1961).

Existing dynamic models of comparative advantage analyze how the innovation technology determines comparative advantage in high-tech production (Grossman 1990; Grossman and Helpman 1991). 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 and imitation drive productivity growth. To understand the effects of learning on the pattern of trade, this paper develops a theory of how innovation and learning jointly determine the global productivity distribution and Ricardian comparative advantage.

The theory is based upon a new endogenous growth model with many countries and industries. The model has several important features. (i) Productivity growth results from R&D investment by incumbent firms with heterogeneous R&D capabilities. (ii) Firms further from the productivity frontier have an advantage of backwardness that raises their productivity growth conditional on R&D investment (Gerschenkron 1962). For firms close to the productivity frontier R&D is primarily about creating new knowledge, while at less advanced firms R&D takes the form of imitation and learning (Cohen and Levinthal 1989). (iii) Knowledge spillovers within and across countries raise the effectiveness of R&D. Spillovers depend upon the frontier productivity in each country meaning that shifts in the productivity distribution generate spillovers. (iv) The R&D technology and knowledge spillovers differ across both countries and industries. In equilibrium, it is these differences that give rise to comparative advantage.

Some countries are better at innovation and learning than others. These countries have well-developed institutions and infrastructure to support R&D (Nelson 1993) and a high capacity to absorb and learn from existing knowledge and technologies (Caselli and Coleman 2001). I capture these differences through cross-country variation in R&D efficiency, which measures a country’s absolute advantage in R&D. I also allow for international differences in the extent to which the legal system and innovation policy impose barriers to knowledge spillovers that increase the excludability of technical knowledge and reduce the spillovers a

---

1Costinot and Donaldson (2012) take an alternative approach and infer productivity using crop yield predictions from agronomic models that take into account water availability, soil characteristics and climatic conditions.
country generates.

The advantage of backwardness is assumed to be industry-specific consistent with evidence of inter-
industry variation in how current productivity affects future productivity growth (Doraszelski and Jauman-
dreu 2013). Knowledge spillovers enable firms to learn from innovations that occur anywhere in the world,
but I assume domestic spillovers are stronger than international spillovers and restrict spillovers to occur
within industries. Empirical studies of R&D and patent citations also find the strength of spillovers across
firms and the geographic scope of knowledge flows differ by industry (Bernstein and Nadiri 1989; Peri
2005). Consequently, I allow the parameters that control the strength of domestic knowledge spillovers
and the strength of global knowledge spillovers to vary across industries.

To understand how these country and industry characteristics affect comparative advantage I embed
firms’ R&D choices in a many country, many industry Armington trade model. I keep the general equilib-
rium structure simple. There are no trade costs. Within each country-industry pair there is perfect competi-
tion between firms that produce an homogeneous output under decreasing returns to scale. Labor is the only
factor of production and a free entry condition determines the number of producers.

Firms make forward-looking R&D investment decisions taking into account the current and future levels
of prices, wages and knowledge spillovers together with their own R&D capability and their current produc-
tivity level. All else equal, firms that have higher R&D capabilities and/or produce in countries that have
higher R&D efficiency or receive greater knowledge spillovers will invest more in R&D and grow faster.
This results in the emergence of productivity gaps between firms both within and across countries and leads
to less productive firms benefitting from the advantage of backwardness. The productivity gap between any
pair of firms in the same industry is stable if and only if the advantage of backwardness exactly offsets all
other sources of heterogeneity and both firms grow at the same rate. Consequently, any steady state of the
global economy features endogenous productivity differences across firms and countries.

Long-run growth occurs in steady state when the combined strength of domestic and global knowledge
spillovers is sufficiently large relative to the advantage of backwardness. In particular, the global economy
exhibits sustained growth if and only if increasing the productivity of all firms by the same proportion
does not affect productivity growth holding all else constant. To build understanding I start by analyzing a
model with weak knowledge spillovers in which growth eventually runs out of steam and the steady state
equilibrium is stationary. I then allow for growth and show the global economy has a unique balanced growth
path on which the growth rate of each industry is the same in all countries. The equilibrium growth rate is
increasing in the R&D efficiency of each country and decreasing in each country’s barriers to knowledge
spillovers.

On the balanced growth path, the productivity distribution in each country-industry pair shifts outwards
over time, while productivity gaps between firms remain constant. Within countries firms with higher R&D
capability are more productive. Across countries, relative productivity is greater when R&D efficiency is
higher and barriers to knowledge spillovers are lower. Lower barriers to knowledge spillovers increase the
spillovers received by all countries, but since domestic spillovers are stronger than international spillovers
they disproportionately benefit domestic firms. Crucially, the size of the equilibrium productivity gaps re-
sulting from firm and country heterogeneities are industry-specific because they depend upon the advantage
of backwardness and the strength of knowledge spillovers. And these relative productivity levels are what determines both the pattern of Ricardian comparative advantage and the degree of inequality within industries and across countries.

A higher advantage of backwardness reduces the productivity gaps that support balanced growth. Consequently, in a single sector economy where countries only differ in R&D efficiency, international inequality in wages, income and consumption is decreasing in the advantage of backwardness. And when there are many industries countries with higher R&D efficiency have a comparative advantage in industries where the advantage of backwardness is lower.

Global knowledge spillovers do not affect comparative advantage because they have a symmetric impact on all countries. But stronger domestic knowledge spillovers are relatively more beneficial in higher productivity countries that generate more spillovers. Thus, the extent to which knowledge spillovers are local, rather than global, matters for international inequality and comparative advantage. In a single sector economy, countries with higher R&D efficiency or lower barriers to knowledge spillovers have higher relative wage, income and consumption levels when spillovers are more localized. Similarly, with many industries, countries that have higher R&D efficiency or lower barriers to knowledge spillovers have a comparative advantage in industries where the localization of knowledge spillovers is greater.

These results identify novel sources of comparative advantage driven by the impact of knowledge spillovers and learning on firm-level R&D. The combination of an advantage of backwardness and global knowledge spillovers is necessary for the existence of a balanced growth path with stationary relative productivity differences, while the endogenous size of these differences determines international inequality and Ricardian comparative advantage. The theory highlights how absolute advantage in R&D is a source of comparative advantage in industries where leaders can pull further ahead of followers and how it is not the strength, but the geography of knowledge spillovers that matters for comparative advantage.

The most important antecedent to this paper is Grossman and Helpman (1991, chs.7-8) who analyze the determinants of comparative advantage in expanding variety and quality ladder growth models. They study two country, two sector models in which R&D only occurs in the high-tech sector and analyze how factor endowments, country size and initial conditions determine comparative advantage in high-tech production. When knowledge spillovers are global, the relatively skill abundant country has a comparative advantage in high-tech because R&D is relatively skill intensive, while if spillovers are purely domestic initial conditions and country size shape comparative advantage. In contrast to Grossman and Helpman’s focus on innovation, I consider how knowledge spillovers and learning affect comparative advantage. Moreover, while Grossman and Helpman consider economies where average firm productivity within each industry does not vary by country, this paper analyzes Ricardian productivity differences across countries.

Analysis of the pattern of trade and the location of innovation in endogenous growth models can also be found in Grossman and Helpman (1990), Taylor (1993) and Durkin Jr. (1997). But in these papers exogenous variation in the productivity of R&D relative to output production determines comparative advantage. Similarly, Somale (2016) develops a many industry version of Eaton and Kortum (2001) in which

---

2 Initial conditions also matter for comparative advantage in learning-by-doing models (Krugman 1987; Redding 1999). However, in this paper the existence of global knowledge spillovers implies steady state comparative advantage is not path dependent.
comparative advantage in production results from exogenous comparative advantage in innovation and a home market effect in demand. Unlike these papers I do not assume any exogenous variation in comparative advantage across countries.

Learning and imitation do shape trade flows in product cycle models, but the product cycle literature focuses on intra-industry trade (Vernon 1966; Krugman 1979; Grossman and Helpman 1991, chs.11-12). In addition, learning plays an important role in recent work on idea flows in single sector economies (Lucas and Moll 2014; Perla and Tonetti 2014; Sampson 2016a; Buera and Oberfield 2016) 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 consider comparative advantage.

This paper is also related to Romalis (2004), Nunn (2007), Chor (2010), Cuñat and Melitz (2012) and Manova (2013) who study empirically how cross-sectional variation in steady state comparative advantage depends upon the interaction of country and industry characteristics, but do not consider productivity dynamics. An alternative approach is taken by Hanson, Lind and Muendler (2013) who analyze how the pattern of comparative advantage changes over time, but remain agnostic about the causes of comparative advantage.

In addition to offering insight into the origins of comparative advantage, this paper makes a methodological contribution to the growth literature by developing a new endogenous growth model. The theory has several attractive properties. First, it gives a tractable model of how the growth rate and distribution of productivity within and across countries depend upon R&D investment by incumbent firms. Foster, Haltiwanger and Krizan (2001) estimate between around one-half and three-quarters of productivity growth in US manufacturing from 1977-87 can be attributed to incumbent firms. Likewise, using 1992-2002 data, Garcia-Macia, Hsieh and Klenow (2015) conclude most growth in US manufacturing comes from incumbent firms and that improvements in the production of existing varieties are a more important source of growth than creative destruction or the introduction of new varieties. The model developed in this paper provides an alternative framework for studying incumbent firm R&D to recent work on incumbent innovation and imitation in closed economy, quality ladder growth models (Klette and Kortum 2004; Akcigit and Kerr 2016; König, Lorenz and Zilibotti 2016).

Second, the theory shows how asymmetric countries and industries can be incorporated into the recent literature on idea flows with heterogeneous firms to study the global productivity distribution. As in Parente and Prescott (1994), Barro and Sala-i-Martín (1997), Eaton and Kortum (1999), Howitt (2000) and

---

3 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.

4 For important previous work on incumbent firm R&D investment with symmetric countries see Atkeson and Burstein (2010) and Perla, Tonetti and Waugh (2015). Unlike Atkeson and Burstein (2010) I model knowledge spillovers as depending upon the location of the productivity distribution. In contrast to Perla, Tonetti and Waugh (2015) who study technology upgrading by low productivity firms, I analyze an economy where all firms invest in R&D and R&D expenditure is higher at more productive firms.

5 Perla, Tonetti and Waugh (2015), Sampson (2016a) and Impullitti and Licandro (2016) study the effects of trade on growth in symmetric country, single sector, heterogeneous firm models with productivity spillovers. An alternative approach to modeling international productivity gaps is adopted by Basu and Weil (1998) and Acemoglu and Zilibotti (2001) who develop models in which all countries have access to the same technologies, but productivity differences result from whether the technologies are appropriate to a country’s factor supplies.
Buera and Oberfield (2016) international knowledge spillovers ensure countries’ relative productivity levels are stationary in steady state. However, in contrast to the existing literature, productivity growth results from R&D investment by incumbent firms and international inequality in productivity, wages and incomes depends upon the advantage of backwardness and the strength of knowledge spillovers.\footnote{Parente and Prescott (1994) calibrate a model with firm-level investment in technology upgrading, but assume exogenous growth of the technology frontier.}

Third, the paper models the advantage of backwardness and knowledge spillovers as functions of observable characteristics of the firm productivity distribution. Consequently, the industry parameters that shape comparative advantage could be estimated using firm-level data. The paper shows how the solution to the firm’s R&D problem can be used to obtain an estimating equation for the advantage of backwardness. Estimating the R&D and knowledge spillovers parameters would be one way to test the paper’s theory of Ricardian comparative advantage. More generally, the link between the R&D technology and observable microeconomic variables should facilitate using data to discipline the model.

The remainder of the paper is organized as follows. To motivate the theory, Section 2 briefly summarizes evidence documenting sources of heterogeneity across countries and industries in the R&D technology and knowledge spillovers. Section 3 introduces the global productivity model. Section 4 solves a stationary version of the model, while Section 5 solves for the balanced growth path and analyzes its properties. Finally, Section 6 offers some concluding remarks.

\section{R&D and Knowledge Spillovers: Country and Industry Heterogeneity}

The institutions and infrastructure that support R&D vary across countries. Nelson (1993) uses the term ‘national innovation system’ to summarize the complex set of factors that determine innovative performance. Important components of a national innovation system include the relationships between firms and research universities, government support for innovation, the supply of trained scientists and engineers, managerial culture, attitudes towards risk and how firms are owned and financed. Collectively these factors generate large cross-country differences in innovation inputs and outcomes or, as Ohlin (1933, p.86) puts it, “Nations vary much in inventive ability”. For example, in 2015 the US and Japan together accounted for 47\% of all applications filed under the World Intellectual Property Organization’s Patent Cooperation Treaty, while producing 30\% of world GDP.\footnote{Patent application data obtained from \url{http://www.wipo.int/pressroom/en/articles/2016/article_0002.html} on 27 September 2016. GDP data converted to US dollars at market exchange rates taken from the World Bank’s World Development Indicators.}

Since knowledge is non-rival and partially non-excludable, the invention of new ideas, products and technological processes not only advances the technology frontier, but also causes knowledge spillovers to firms and countries behind the frontier. Gerschenkron (1962) calls this the advantage of backwardness. However, countries differ in their capacity to absorb knowledge and learn from existing technologies as argued by Parente and Prescott (1994). For example, Comin, Hobijn and Rovito (2008) document that the rates at which new technologies are adopted differ greatly across countries and are highly correlated with GDP per capita, while Comin and Hobijn (2009) find evidence political institutions affect technology diffu-
sion through changing the cost of erecting barriers to adoption. Policies to support learning and knowledge absorption have also been credited with an important role in the rapid industrialization of East Asian countries such as Japan, South Korea and Taiwan in the second half of the twentieth century (Nelson and Pack 1999). Kim (1997) documents the extensive involvement of the South Korean government in facilitating technological learning during South Korea’s industrialization.

To capture international differences in national innovation systems and knowledge absorption capacities in a simple manner I allow the efficiency of R&D investment to have an exogenous country-specific component, meaning absolute advantage in R&D differs across countries. Embedded in this modeling choice is the assumption that more innovative countries also have higher knowledge absorption capacities. 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.” The model also includes a second source of cross-country heterogeneity, differences in barriers to knowledge spillovers. This parameter captures variation in institutions, such as legal protection for intellectual property, that affect the extent to which a firm’s technical knowledge spills over to other domestic and international firms.

Evidence the R&D technology varies across industries comes from production function estimates that find the elasticity of output to R&D and the rate of return to R&D differ by industry (Hall, Mairesse and Mohnen 2010). Most of this literature uses observed R&D to construct measures of firms’ knowledge capital stocks. However, in recent work Doraszelski and Jaumandreu (2013) develop a new approach to estimating firm-level productivity dynamics that treats productivity as unobservable to the econometrician and does not assume firms have a knowledge capital stock. They find that the effect of current productivity on future changes in productivity differs across industries. I model this heterogeneity by allowing for the dependence of the R&D technology on current productivity, which captures the advantage of backwardness, to be industry specific.

An extensive empirical literature documents the existence and geographic localization of knowledge spillovers (Griliches 1992; Jaffe, Trajtenberg and Henderson 1993; Eaton and Kortum 1999; Branstetter 2001; Keller 2002). Less work has been done on how knowledge spillovers vary by industry, but available studies still demonstrate substantial cross-industry variation. Bernstein and Nadiri (1989) estimate intra-industry knowledge spillovers for four US industries: Chemicals, Instruments, Machinery and Petroleum. They find the elasticity of variable and average costs to other firms’ R&D capital is twice as large in the Chemicals and Petroleum industries as in Instruments and Machinery, implying the strength of domestic spillovers varies across industries.

Peri (2005) studies knowledge flows, as measured by patent citations, within six sectors: Chemical, Computers, Drugs, Electronics, Mechanical and Others. While confirming the localized nature of knowledge, he finds knowledge flows in Electronics and, particularly, Computers are more global than in the other sectors. Consistent with Peri’s results, Malerba, Mancusi and Montobbio (2013) find the size of international, relative to domestic, R&D spillovers is greater in the Electronics industry (which includes Computers) than in Chemicals or Machinery. Additional evidence geography has a smaller effect on spillovers.
in the Computer industry than in other industries comes from Irwin and Klenow (1994) and Acharya and Keller (2009). In a study of learning-by-doing among US and Japanese firms in the Semiconductor industry, Irwin and Klenow (1994) find no evidence spillovers are stronger within countries than between the US and Japan. Acharya and Keller (2009) estimate the effect of R&D spillovers on output for 22 industries in 17 countries. They find the elasticity of output to US R&D varies more across industries than the elasticity of output to domestic R&D suggesting there exist cross-industry differences in the strength of international knowledge spillovers. Specifically, they find the elasticity with respect to US R&D is largest in the Radio, TV and Communication Equipment, and Office, Accounting and Computing Machinery industries.

Collectively these papers show both the strength and geography of knowledge spillovers vary across industries. This variation likely results from many factors including the extent to which knowledge is tacit or, alternatively, codifiable; the ease of reverse engineering products and technologies; 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). The next section develops an endogenous productivity model that includes cross-industry differences in the strength and geographic scope of knowledge spillovers together with the other sources of country and industry heterogeneity described above.

3 A Global Productivity Model

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.

3.1 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:

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

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 $\iota_s$. An individual in country $s$ with initial assets $a(t)$ chooses a consumption path to maximize utility subject to the budget constraint:

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

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:

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

$$
\lim_{\tau \to \infty} \left\{ a_s(\tau) \exp \left[ - \int_1^\tau v_s(\tilde{\tau}) d\tilde{\tau} \right] \right\} = 0.
$$

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:

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

where $X_{js}$ denotes consumption of industry $j$ in country $s$ and $\mu_j$ 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:

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

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

There is free trade, implying consumers in all countries pay the same price $p_{js}$ for industry $j$ output that is produced in country $s$. Therefore, solving consumers’ intratemporal optimization problem yields:

$$
P_j X_{js} = \mu_j z_s c_s L_s, \quad \text{(4)}
$$

$$
z_s = \prod_{j=1}^{J} p_j^{\mu_j}, \quad \text{(5)}
$$

$$
x_{j\tilde{s}s} = \left( \frac{p_{j\tilde{s}}}{p_j} \right)^{-\sigma} X_{js}, \quad \text{(6)}
$$

$$
P_j = \left( \sum_{s=1}^{S} p_{js}^{1-\sigma} \right)^{\frac{1}{1-\sigma}}, \quad \text{(7)}
$$

where $P_j$ denotes the industry $j$ price index which does not vary across countries, as shown by (7). It follows from (5) that the consumption price $z_s = z$ is also constant across countries.
3.2 Production

Within each country-industry pair, all firms produce the same homogeneous output, but firms differ along two dimensions: capability $\psi$ and productivity $\theta$. Capability is a time-invariant firm characteristic that affects a firm’s R&D efficiency and, consequently, the evolution of its productivity. Section 3.3 describes the R&D technology and productivity dynamics. I assume the distribution of capabilities in industry $j$ has support $[\psi, \bar{\psi}]$ and cumulative distribution function $G(\psi)$ that does not vary across countries.

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 $\theta$ that employs $l^P$ production workers produces output:

$$y = \theta (l^P)^\beta,$$

with $0 < \beta < 1$.

Note the production technology is independent of the firm’s capability $\psi$.

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 $\theta$ as given. 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 $l^P_{js}$, output $y_{js}$ and production profits $\pi^P_{js}$ are given by:

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

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

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

Production employment, output and production profits are all increasing in 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 3.5, while $H_{js}$ depends upon firms’ R&D investment choices. Summing up across firms we have that aggregate production employment $L^P_{js}$ in industry $j$ and country $s$ is:

$$L^P_{js} = M_{js} \left( \frac{\beta p_{js}}{w_s} \right)^{\frac{1}{1-\beta}} \int_\theta^1 \theta^{\frac{1}{1-\beta}} dH_{js}(\theta),$$

---

8 R&D investment affects future productivity growth (see Section 3.3), but not the current value of $\theta$. Consequently, the firm’s static production decision is separable from its dynamic R&D investment decision.

9 Decreasing returns to scale in production can result from production becoming less efficient as the firm’s span of control increases. An alternative way to introduce concavity is to assume each firm faces a downward sloping demand curve. For example, suppose each firm produces a differentiated variety and there is monopolistic competition between firms and a constant elasticity of substitution between varieties. This alternative approach would make little difference to the firm’s optimization problem, but would make the model analytically intractable in general equilibrium given the existence of many asymmetric countries and industries.
and aggregate output is:

\[ Y_{js} = M_{js} \left( \frac{\beta P_{js}}{w_s} \right)^{\frac{\gamma}{1-\gamma}} \int_\theta \theta^{\frac{1-\gamma}{\gamma}} dH_{js}(\theta). \]  (12)

### 3.3 R&D and Productivity

In addition to producing output, firms invest in R&D to increase their future productivity. Knowledge spillovers allow firms that perform R&D to utilize the knowledge created by past innovations both at home and abroad. Following Cohen and Levinthal (1989) I assume there are two faces of R&D: innovation and learning. Firms with high productivity levels are close to the technology frontier and increase productivity primarily through innovation. By contrast, productivity growth at firms far from the frontier results mainly from learning existing knowledge and adopting technologies already used by more productive firms. Using industry level data for OECD countries, Griffith, Redding and Van Reenen (2004) provide evidence R&D leads to productivity growth through both innovation and catch-up learning.

The R&D technology is such that a firm with capability \( \psi \) and productivity \( \theta \) that employs \( l^R \) workers in R&D has productivity growth:

\[ \frac{\dot{\theta}}{\theta} = B_s \chi_{js} \psi^{\gamma_j} \left( l^R \right)^{\alpha - \delta}, \]  (13)

where \( \gamma_j > 0, \delta > 0 \) and \( \alpha \in (0,1) \). The variable \( \chi_{js} \) represents the knowledge level in industry \( j \) in country \( s \). \( \chi_{js} \) is defined in Section 3.4 and depends upon the frontier technical efficiency in all \( S \) countries and the strengths of both domestic and global knowledge spillovers. A higher knowledge level makes R&D more efficient. The efficiency of R&D also varies across firms and countries. Firms with a higher capability \( \psi \) are better at R&D. In equilibrium, this will lead to cross-firm heterogeneity in R&D investment and productivity. Country-wide R&D efficiency is given by \( B_s \), which captures cross-country differences in both the quality of a country’s national innovation system and a country’s capacity to absorb knowledge spillovers. Countries with a higher \( B_s \) have an absolute advantage in R&D.

The industry-specific parameter \( \gamma_j \) gives the elasticity of productivity growth to a firm’s current productivity holding R&D investment constant. Since \( \gamma_j > 0 \) productivity growth is decreasing in firm productivity. This means R&D is harder for more productive firms that have less to learn from their competitors and must rely more on innovation to increase their productivity. Variation in \( \gamma_j \) across industries results from differences in the extent to which growth becomes harder as technology advances. In industries with high \( \gamma_j \) an increase in \( \theta \) has a large negative effect on the efficiency of R&D implying raising productivity is much more costly at higher productivity firms. By contrast, a low \( \gamma_j \) implies current productivity is a less important determinant of productivity growth. I will refer to \( \gamma_j \) as the advantage of backwardness in industry \( j \). The returns to scale in R&D are given by \( \alpha \), while \( \delta \) is the rate at which a firm’s technical knowledge depreciates causing its productivity to decline.

Firms face a constant instantaneous probability \( \zeta > 0 \) of suffering a shock that leads to the death of the firm. Taking this risk into account, each firm chooses a path for R&D employment to maximize its value subject to the R&D technology (13). Let \( V_{js}(\psi, \theta) \) be the value of a firm with capability \( \psi \) and productivity
\( V_{js}(\psi, \theta) \) equals the expected present discounted value of the firm’s production profits minus its R&D costs:

\[
V_{js}(\psi, \theta) = \int_{t}^{\infty} \exp \left[ - \int_{t}^{\tau} (\ell_s + \zeta) \, d\tilde{\tau} \right] \left[ \pi_{js}^P(\theta) - w_s l_R \right] \, d\tau,
\]

(14)

where \( \pi_{js}^P(\theta) \) is given by (10). All endogenous variables in this expression, including the firm’s value function, may depend on \( t \), but I have suppressed this dependence to simplify notation.

### 3.4 Knowledge Spillovers

Knowledge is non-rival and at least partially non-excludable. Consequently, firms learn from the R&D successes of their domestic and foreign competitors. The knowledge level \( \chi_{js} \) captures these spillovers. I assume \( \chi_{js} \) is increasing in the frontier productivity in industry \( j \) in all countries. I also allow for domestic spillovers to be stronger than international spillovers and for countries to differ in their capacity to create spillovers. I do not allow for interindustry spillovers in this paper.

To be specific, let \( \omega \) index firms and let \( \Omega_{js} \) denote the set of firms operating in industry \( j \) in country \( s \). Define:

\[
\theta_{js}(\omega) = \sup_{\tilde{\omega} \in \Omega_{js}, \tilde{\omega} \neq \omega} \{ \theta(\tilde{\omega}) \}.
\]

\( \theta_{js}(\omega) \) is the supremum of the productivity of all firms in industry \( j \) in country \( s \) excluding firm \( \omega \). Within each country-industry pair there will always be either zero or a continuum of firms with any given productivity level.\(^\text{10}\) Therefore, \( \theta_{js}(\omega) = \theta_{ja} \) and does not vary with \( \omega \). I now define the knowledge level \( \chi_{js} \) of industry \( j \) in country \( s \) to be:

\[
\chi_{js} = \left( \frac{\theta_{js}}{N_s} \right)^{\eta_j} \left( \sum_{s=1}^{S} \frac{\theta_{js}}{N_s} \right)^{\nu_j}.
\]

(15)

Since \( \chi_{js} \) is independent of \( \theta(\omega) \) each firm takes the current and future values of \( \chi_{js} \) as given when choosing its R&D investment.

The knowledge level is the product of two terms: domestic spillovers given by \( (\theta_{js}/N_s)^{\eta_j} \) and global spillovers. The parameter \( \eta_j \) gives the elasticity of the knowledge level to the domestic productivity frontier, which I will call the strength of domestic knowledge spillovers. I assume \( \eta_j \geq 0 \). The term that captures global spillovers is an increasing function of the productivity frontiers in all \( S \) countries including country \( s \). The elasticity \( \nu_j > 0 \) gives the strength of global knowledge spillovers. The sum \( \eta_j + \nu_j \) is the total strength of knowledge spillovers in industry \( j \). When \( \eta_j = 0 \) the knowledge level \( \chi_{js} \) does not vary by country implying knowledge is global and available to all firms regardless of their location. Whenever \( \eta_j > 0 \) an increase in \( \theta_{js} \) generates both global spillovers, which increase the knowledge level in all countries including

---

\(^{10}\)In steady state this follows from the assumption there are a continuum of firms with each capability \( \psi \), as shown in Section 4.1. Outside steady state it also requires assuming an initial condition in which there are either zero or a continuum of firms with each \( (\psi, \theta) \) pair.
country $s$, and also domestic spillovers which are limited to country $s$. In this case within-country spillovers are stronger than cross-country spillovers and the knowledge level differs by country.

The knowledge level also depends upon the set of country-specific parameters $N_{\bar{s}}$. I assume $N_{\bar{s}} > 0$ for all $\bar{s}$. $N_{\bar{s}}$ is an inverse measure of the extent to which firms learn from knowledge created in country $\bar{s}$. Countries with policies and institutions that limit knowledge spillovers have a high $N_{\bar{s}}$. I will refer to $N_{\bar{s}}$ as the barriers to knowledge spillovers.

### 3.5 Entry

Entrants must pay a fixed cost to establish a firm. To generate 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 $\psi$ and initial productivity $\theta$ of each entrant are determined by a random draw from the joint distribution of $\psi$ and $\theta$ in the entrants’ country and industry at the moment the new firm is created. This implies the existence of spillovers from incumbents to entrants within a country-industry pair.\(^{11}\) There is free entry and the free entry condition requires the cost of entry equals the expected value of entry meaning:

$$f^E w_s = \int_{(\psi,\theta)} V_{js}(\psi,\theta) d\tilde{H}_{js}(\psi,\theta),$$  
(16)

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

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

$$\dot{M}_{js} = -\zeta M_{js} + \frac{L^E_{js}}{f^E}.$$  
(17)

### 3.6 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:

$$L_s = \sum_{j=1}^J \left( L^P_{js} + L^R_{js} + L^E_{js} \right),$$  
(18)

where production employment $L^P_{js}$ in industry $j$ is given by (11), $L^R_{js}$ denotes aggregate R&D employment in industry $j$ and $L^E_{js}$ is aggregate employment in entry in industry $j$.

Industry output markets clear country-by-country implying domestic output $Y_{js}$ equals the sum over all countries of consumption of goods produced in country $s$. This means:

\(^{11}\)In Sampson (2016a) spillovers from incumbents to entrants lead to endogenous growth through a dynamic selection mechanism. By contrast, in this paper the absence of selection implies incumbent firm R&D is the source of all productivity growth.
\[ Y_{js} = \sum_{\tilde{s}=1}^{S} x_{js\tilde{s}}. \quad (19) \]

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:

\[ a_s L_s = \sum_{j=1}^{J} M_{js} \int_{(\psi, \theta)} V_{js}(\psi, \theta) d\tilde{H}_{js}(\psi, \theta). \quad (20) \]

I also let global consumption expenditure be the numeraire implying:

\[ \sum_{s=1}^{S} z c_s L_s = 1. \quad (21) \]

Finally, I assume the parameters that govern the returns to scale in production and R&D, the advantage of backwardness and the strength of domestic spillovers satisfy the following restriction.

**Assumption 1.** For all industries \( j \), the parameters of the global economy satisfy: \[ \frac{1}{1-\beta} > \gamma_j > \frac{\alpha}{1-\beta} + \eta_j. \]

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

### 3.7 Equilibrium

An equilibrium of the global economy is defined by time paths for consumption per capita \( c_s \), assets per capita \( a_s \), wages \( w_s \), the interest rate \( \epsilon_s \), the consumption price \( z \), consumption levels \( X_{js} \) and \( x_{js\tilde{s}} \), prices \( P_j \) and \( p_{js} \), production employment \( L^P_{js} \), industry output \( Y_{js} \), the mass of firms \( M_{js} \), knowledge levels \( \chi_{js} \), R&D employment \( L^R_{js} \), employment in entry \( L^E_{js} \) and the joint distribution of firms’ capabilities and productivity levels \( \tilde{H}_{js}(\psi, \theta) \) for all countries \( s, s = 1, \ldots, S \) 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 (11) and (12), respectively; (iv) firms’ productivity levels evolve according to the R&D technology (13) and firms choose R&D employment to maximize their value (14); (v) knowledge levels are given by (15); (vi) 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 (16) holds and the mass of firms evolves according to (17), and; (viii) labor, output and asset market clearing imply (18)-(20) 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 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 \( B_s \), barriers to knowledge spillovers \( N_s \) and population \( L_s \). Meanwhile,
there is variation across industries in the advantage of backwardness $\gamma_j$, the strengths of domestic and global knowledge spillovers $\eta_j$ and $\nu_j$, respectively, and the expenditure share $\mu_j$.

4 Stationary Equilibrium

The properties of the equilibrium depend qualitatively on the relative magnitudes of the advantage of backwardness $\gamma_j$ and the total strength of knowledge spillovers $\eta_j + \nu_j$. Consider the following assumption.

Assumption 2. For all industries $j$, the advantage of backwardness is strictly greater than the sum of the strengths of domestic and global knowledge spillovers: $\gamma_j > \eta_j + \nu_j$.

Assumption 2 imposes an upper bound on the total strength of knowledge spillovers. The sum $\eta_j + \nu_j$ gives the elasticity of the knowledge level to productivity growth when the productivity frontier increases by the same proportion in all countries. Under Assumption 2 the R&D technology (13) implies that if the productivity of all firms in the world increases by the same proportion, then the efficiency of R&D declines for all firms. This means knowledge spillovers are too weak to generate ongoing productivity growth, implying there cannot be a steady state equilibrium in which the productivity frontier grows without bound. Since productivity increases are the only possible source of long-run growth in this economy it follows that, in equilibrium, sustained growth is not possible.

In Section 5.1 I analyze the consequences of relaxing Assumption 2 to allow for endogenous growth. However, in this section I impose Assumption 2 and look for a stationary equilibrium of the global economy. I define a stationary equilibrium as an equilibrium in which all aggregate country and industry level variables are constant. In what follows I outline how to solve for a stationary equilibrium. Full details of the solution together with proofs of the propositions can be found in Appendix A.

4.1 Firm-level R&D and Productivity Dynamics

For the economy to be in a stationary equilibrium the productivity distribution $H_{js}(\theta)$ must remain constant over time. Is this requirement consistent with optimal firm behavior? The first step in solving the model is to characterize firms’ R&D choices and productivity growth rates.

Firms in industry $j$ in country $s$ choose R&D employment taking the time paths of $\iota_s$, $w_s$, $p_{js}$ and $\chi_{js}$ as given. Suppose the economy is in a stationary equilibrium implying $\iota_s$, $w_s$, $p_{js}$ and $\chi_{js}$ are time invariant. Then the Euler equation (2) implies the interest rate is constant across countries and equal to the discount rate $\rho$. Since $\rho > 0$ the transversality condition (3) is satisfied.

To solve the firm’s R&D problem it is useful to let $\phi \equiv \theta^{\frac{1}{1-\beta}}$ be a transformation of the firm’s productivity level. Taking the time derivative of $\phi$ and using the R&D technology (13) implies:

$$\frac{\dot{\phi}}{\phi} = \frac{1}{1-\beta} \left[ B_s \chi_{js} \psi \phi^{-\gamma_j (1-\beta)} (I^R)^{\alpha} - \delta \right].$$

(22)

Now, by substituting the production profits function (10) into the value function (14), using $\iota_s = \rho$ and changing variables from $\theta$ to $\phi$, the optimization problem of a firm with capability $\psi$ can be written as:
subject to the growth of $\phi$ being given by (22) and an initial value for $\phi$ at time $t$. Since $w_s$ and $p_{js}$ are time invariant, the payoff function depends upon time only through exponential discounting meaning the firm faces a discounted infinite-horizon optimal control problem with state variable $\phi$ and control variable $l^R$. I prove in Appendix A that any solution to the firm’s problem must satisfy:

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

Equations (22) and (23) are a system of differential equations for $\phi$ and $l^R$. Setting $\dot{\phi} = 0$ and $\dot{l}^R = 0$ shows the system has a unique steady state $\phi^*_j, l^R_j$ given by:

$$\phi^*_j = \left[ \alpha \beta \psi \left( B_s \chi_j \psi \right) \frac{1}{\beta} \left( \frac{p_{js}}{w_s} \right)^{\frac{1}{\beta}} \frac{\delta^{\alpha - 1}}{\rho + \zeta + \gamma_j \delta} \gamma_j^{(1 - \beta) - 1} \right],$$

$$l^R_j = \left[ \alpha \beta \psi \left( B_s \chi_j \psi \right) \frac{1}{\beta} \left( \frac{p_{js}}{w_s} \right)^{\frac{1}{\beta}} \frac{\delta^{\alpha - 1}}{\rho + \zeta + \gamma_j \delta} \gamma_j^{(1 - \beta) - 1} \right].$$

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 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 $\phi$ below $\phi^*_j$, while the opposite is true for firms with initial $\phi$ above $\phi^*_j$.

The steady state has several important properties. First, in steady state the productivity of each surviving firm is time invariant. In a stationary equilibrium the productivity distribution $H_{js}(\theta)$ must remain constant over time. 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. Recall entrants draw their capability and productivity from the joint distribution of $\psi$ and $\theta$ among incumbent firms, while all firms face the same instantaneous exit probability $\zeta$. 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 choices generate a time invariant productivity distribution if and only if all incumbent firms are in steady state. Consequently, in any stationary equilibrium all firms are in steady state and each surviving firm’s productivity and R&D employment are time invariant and given by (24) and (25), respectively.\(^\text{12}\)

Second, $\phi^*_j$ is increasing in $\psi$ implying within any country-industry pair firms with greater capabilities

\(^{12}\)Formally, a stationary equilibrium only requires a mass $M_{js}$ of firms to be in steady state, which allows for individual firms with zero mass to deviate from steady state. I overlook this distinction since it does not matter for industry or aggregate outcomes.
have higher steady state productivity levels. This explains why, even though capability differs across firms, all firms have constant productivity in steady state. The advantage of backwardness raises the R&D efficiency of less productive firms which offsets the disadvantage of having a lower capability in such a way that in steady state R&D investment cancels out knowledge depreciation and productivity remains constant at all firms. Similarly, level differences in productivity across countries exactly offset all other sources of cross-country heterogeneity to ensure the productivity distribution is time invariant in all countries. The existence of an advantage of backwardness is also necessary for the stability of the steady state because it implies a negative relationship between a firm’s productivity level and its productivity growth holding all else constant.

Third, steady state firm behavior 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 $\phi^*_j$ and $I^{Rs}_j$ are both increasing in $\psi$, and; (ii) R&D intensity is independent of firm size since using (9) and (25) implies the steady state ratio of R&D investment to sales satisfies:

$$\frac{w_{s}I^{Rs}_{js}}{p_{js}y_{js}(\phi^*_{js})} = \frac{\alpha\delta}{\rho + \zeta + \gamma_j \delta},$$

which is constant across firms within an industry. R&D intensity is increasing in the returns to scale in R&D $\alpha$ and the knowledge depreciation rate $\delta$ and decreasing in the advantage of backwardness $\gamma_j$, the interest rate $\rho$ and the firm exit rate $\zeta$.

---

13 Although empirical studies usually find R&D intensity is independent of firm size among firms that report positive R&D, they also find smaller firms are more likely to report zero R&D investment. However, as Klette and Kortum (2004) point out this may be an artifact of the R&D data which probably under measures informal investment in productivity growth, particularly at smaller firms that do not have formal R&D divisions. This bias is likely to be exacerbated by the definition of R&D used in this paper, which covers investment in both innovation and learning, whereas R&D data primarily measures investment in innovation. The under measurement will be most severe for firms at which learning is the main source of productivity growth, e.g. less productive firms and firms in developing countries.
To obtain the distribution of productivity in a stationary equilibrium recall $G(\psi)$ is the cumulative distribution function of firms’ capabilities. Using (24) and the definition of $\phi$ it follows that in a stationary equilibrium the productivity distribution of both incumbent firms and entrants is given by:

$$H_{js}(\theta) = G\left[\alpha^{-\alpha} \beta^{\alpha} \frac{1}{B_s \chi_{js}} \frac{w_s}{p_{js}} \left(\rho + \zeta + \gamma_j \delta\right)^{\alpha} \theta^{\frac{1-\alpha}{1-\beta}}\right].$$

In this economy both the location and the shape of the productivity distribution are endogenously determined by firms’ R&D decisions.\textsuperscript{14} The location varies by country and industry and depends upon $w_s$, $\chi_{js}$ and $p_{js}$ which are determined in general equilibrium as described in Section 4.2 below.

Within each industry the shape of the productivity distribution does not vary by country, but depends upon the exogenous firm capability distribution $G(\psi)$ and upon the parameters $\alpha$, $\beta$ and $\gamma_j$. Consider two firms in the same country and industry with capabilities $\psi$ and $\psi'$, respectively. The ratio of these two firms’ steady state productivity levels is:

$$\frac{\theta^*_j(\psi')}{\theta^*_j(\psi)} = \left(\frac{\psi'}{\psi}\right)^{\frac{1-\beta}{1-\beta - \alpha}}. \tag{27}$$

Thus, the productivity gap between firms is strictly increasing in $\alpha$ and $\beta$ and strictly decreasing in $\gamma_j$. This implies that, in a stationary equilibrium, inequality in productivity across firms within a country-industry pair is strictly increasing in $\alpha$ and $\beta$, but strictly decreasing in $\gamma_j$.\textsuperscript{15} An increase in the advantage of backwardness $\gamma_j$ reduces the steady state productivity gaps between firms leading to lower productivity inequality. An increase in $\beta$ raises the returns to scale in production giving higher capability, larger firms a greater incentive to raise productivity by increasing R&D investment. Consequently, productivity inequality grows. Similarly, an increase in $\alpha$ raises the returns to scale in R&D which disproportionately benefits higher capability firms that employ more R&D workers and leads to greater productivity inequality. Substituting the solution to the firm’s R&D problem into equations (8)-(10) it also follows that in a stationary equilibrium inequality across firms in production employment, revenue and profits is strictly increasing in $\alpha$ and $\beta$ and strictly decreasing in $\gamma_j$.

Proposition 1 summarizes the results obtained in this section regarding the distribution of productivity in a stationary equilibrium.

**Proposition 1.** Suppose Assumptions 1 and 2 hold. In a stationary equilibrium there is no firm-level productivity growth and within any country-industry pair:

(i) Productivity and R&D employment are strictly increasing in firm capability;

(ii) Productivity inequality across firms is strictly decreasing in the advantage of backwardness, but strictly increasing in the returns to scale in production and R&D.

\textsuperscript{14}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 Boniﬁglioli, Crinò and Gancia (2015) who allow firms to choose between receiving productivity draws from distributions with different shapes.

\textsuperscript{15}Throughout the paper, 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.
4.2 General Equilibrium

To complete the solution for a stationary equilibrium we can now impose the remaining equilibrium conditions of the global economy. Using the definition of the knowledge level, the free entry condition, the goods and labor market clearing conditions and firms’ steady state productivity and employment levels I show in Appendix A that in a stationary equilibrium:

\[
L_s = \sum_{j=1}^{J} \frac{\mu_j}{\rho + \zeta} \left( \zeta + \rho \delta \right) \frac{w_s^{-\sigma} \left( B_s N_s^{-\eta_j} \right)^{\sigma-1}}{w_s^{1-\sigma} \left( B_s N_s^{-\eta_j} \right)^{\sigma-1}}. \tag{28}
\]

Since equation (28) holds for all countries \( s = 1, \ldots, S \) we have a system of \( S \) equations in the \( S \) wage variables. In Appendix A I prove this system has a unique solution, which implies the following proposition.

**Proposition 2.** Suppose Assumptions 1 and 2 hold. The global economy has a unique stationary equilibrium.

In the stationary equilibrium heterogeneity across countries in R&D efficiency \( B_s \) and barriers to knowledge spillovers \( N_s \) leads to international gaps in productivity levels, wages and incomes. Moreover, the productivity gaps necessary to support the stationary global productivity distribution are industry-specific, which generates comparative advantage. The next two sections study how innovation and learning endogenously determine the global productivity distribution, income differences and trade flows. I start by considering how wages and income vary across countries and then proceed to analyze the pattern of comparative advantage.

4.3 International Inequality

To understand the determinants of international wage and income inequality in the stationary equilibrium suppose the economy has a single industry. Since the consumption price \( z \) does not vary across countries, real wage variation depends only upon differences in nominal wages. Setting \( J = 1 \) in equation (28) gives:

\[
\frac{w_s}{w_{\tilde{s}}} = \left( \frac{L_s}{L_{\tilde{s}}} \right)^{-\frac{1}{\sigma}} \left[ \frac{B_s}{B_{\tilde{s}}} \left( \frac{N_s}{N_{\tilde{s}}} \right)^{-\eta_j} \right]^{\frac{\sigma-1}{\sigma (\gamma - \eta)}}. \tag{29}
\]

The relative wage of country \( s \) is increasing in its R&D efficiency \( B_s \), but decreasing in its barriers to knowledge spillovers \( N_s \).\(^{16}\) Moreover, the elasticities of the relative wage to increases in \( B_s \) or reductions in \( N_s \) are increasing in the strength of domestic spillovers \( \eta \) and decreasing in the advantage of backwardness \( \gamma \). Relative wages are independent of the strength of global spillovers \( \nu \). The implication of these observations for international wage inequality depends upon the cross-country correlations between \( B_s, N_s \) and \( L_s \). Consider the case where the ranking of countries by wage levels is the same as the ranking by R&D efficiency or by the inverse of barriers to knowledge spillovers.\(^{17}\) Then international wage inequality is decreasing in

\(^{16}\)The relative wage is also decreasing in population \( L_s \) due to the assumption of Armington demand.

\(^{17}\)This case is guaranteed to hold if all countries have the same population and countries with higher R&D efficiency also have lower barriers to knowledge spillovers.
the advantage of backwardness and increasing in the strength of domestic spillovers.

The explanation for these results lies in how R&D efficiency and barriers to knowledge spillovers impact the location of the productivity distribution. Consider two firms with the same capability $\psi$, but in different countries. In the stationary equilibrium the ratio of these firms’ steady state productivity levels is:

$$\frac{\theta_s^*}{\theta_{\tilde{s}}^*} = \left( \frac{B_s}{B_{\tilde{s}}} \right)^{\frac{1}{\eta - \gamma}} \left( \frac{N_s}{N_{\tilde{s}}} \right)^{\frac{-\eta}{\gamma - \eta}}.\quad (30)$$

This expression is the cross-country analogue of equation (27) which gives the ratio of steady state productivity levels for two firms in the same country and industry, but with different capabilities. Countries with a higher R&D efficiency are better at R&D. Likewise, countries with lower barriers to knowledge spillovers also have an advantage in R&D because they learn more from domestic spillovers for a given domestic productivity frontier, which raises their knowledge level. Consequently, an increase in $B_s$ or a reduction in $N_s$ shifts the productivity distribution in country $s$ outwards relative to other countries as shown by (30). This leads to an increase in the relative wage of country $s$, as becomes apparent from substituting (30) into (29) to obtain:

$$\frac{w_s}{w_{\tilde{s}}} = \left( \frac{L_s}{L_{\tilde{s}}} \right)^{-\frac{1}{\sigma}} \left( \frac{\theta_s^*}{\theta_{\tilde{s}}^*} \right)^{\frac{\sigma - 1}{\sigma}},$$

showing that the relative wage of country $s$ is increasing in the relative productivity of its firms.

An increase in the advantage of backwardness $\gamma$ disproportionately benefits lower productivity countries because it reduces the cross-country productivity gaps necessary to support the stationary equilibrium. Thus, a higher $\gamma$ reduces productivity gaps between countries just as it reduces productivity gaps between firms within the same country and industry. By contrast, the strength of domestic spillovers $\eta$ does not affect productivity gaps within country-industry pairs, but does affect international productivity gaps. Stronger domestic spillovers disproportionately benefit those countries that have higher frontier productivity levels and, consequently, generate more spillovers. This allows countries with higher $B_s$ and lower $N_s$ to pull further ahead of other countries, raising international productivity and wage gaps. The strength of global spillovers $\nu$ does not affect relative productivity levels or relative wages because global spillovers have a symmetric effect on the knowledge level, and, therefore, the location of the productivity distribution, in all countries.

In a single industry economy, assets per capita $a_s$ and consumption per capita $c_s$ are proportional to $w_s$. Therefore, relative income per capita and consumption per capita levels across countries are equal to relative wages and given by (29). It follows that changes in $\gamma$, $\eta$ and $\nu$ have the same effects on international inequality in income and consumption as on inequality in wages. Proposition 3 summarizes how wages, income and consumption vary across countries.

**Proposition 3.** Suppose Assumptions 1 and 2 hold and the economy has a single industry. In the stationary equilibrium:

(i) Each country’s wage relative to all other countries is strictly increasing in its R&D efficiency and strictly decreasing in its barriers to knowledge spillovers;
(ii) The elasticities of each country’s relative wage to increases in its R&D efficiency or reductions in its barriers to knowledge spillovers are strictly decreasing in the advantage of backwardness and strictly increasing in the strength of domestic knowledge spillovers;

(iii) The strength of global knowledge spillovers does not affect relative wage levels.
All these results continue to hold if, instead of wages, income per capita or consumption per capita levels are compared across countries.

Proposition 3 demonstrates an important distinction between the roles played by the advantage of backwardness and knowledge spillovers. Although a greater advantage of backwardness reduces the productivity gap between countries, stronger knowledge spillovers disproportionately benefit not poorer countries, but countries that are better able to take advantage of the spillovers. For domestic spillovers this means higher productivity countries that use more advanced technologies and generate greater spillovers, whereas global spillovers have no impact on relative productivity levels or wages because they affect all countries symmetrically.

4.4 Comparative Advantage

To characterize the pattern of comparative advantage it is sufficient to analyze how exports vary across countries and industries in the stationary equilibrium. Let $C$ denote a country characteristic and $I$ an industry characteristic. Countries with higher $C$ have a comparative advantage in industries with higher $I$ if and only if stationary equilibrium exports are log supermodular in $C$ and $I$. For example, countries with high R&D efficiency have a comparative advantage in industries with strong domestic knowledge spillovers if and only if, comparing across countries and industries within the same stationary equilibrium, exports are log-supermodular in $B_s$ and $\eta_j$. Let $EX_{js\tilde{s}} = p_{js}x_{js\tilde{s}}$ be the value of exports from country $s$ to country $\tilde{s}$ in industry $j$. In the stationary equilibrium log exports can be written as:

$$\log EX_{js\tilde{s}} = \nu_j^1 + \xi_{\tilde{s}}^1 + (\sigma - 1) \log \theta_{js}^\gamma - (\sigma - 1) \log w_s,$$

where $\nu_j^1$ is an industry-specific term and $\xi_{\tilde{s}}^1$ depends only upon the destination country $\tilde{s}$. The relative productivity of firms with different capabilities within a country-industry pair is the same in all countries by (27). Thus, we can interpret the frontier productivity $\theta_{js}^\gamma$ as a measure of the location of the productivity distribution in country $s$. A higher frontier productivity implies an upwards shift in the productivity distribution.

Equation (32) shows that, holding all else constant, exports are increasing in the frontier productivity and decreasing in the wage level. An increase in the frontier productivity means all firms in the industry are more productive, which leads to a lower price $p_{js}$ and higher exports. By contrast, higher wages reduce exports because the increase in labor costs raises the output price and reduces firm-level production and exports. Equation (32) also implies variation in $\theta_{js}^\gamma$ is the only possible source of comparative advantage.\textsuperscript{18}

Thus, any comparative advantage that exists in this economy is Ricardian in nature.

\textsuperscript{18}To see this, note $\log \theta_{js}^\gamma$ is the only term on the right hand side of (32) that depends upon both country characteristics and either $\gamma_j$, $\eta_j$ or $\nu_j$. 

21
The location of the productivity distribution in industry \( j \) and country \( s \) is given by:

\[
\bar{\theta}_{js}^* = \left( \frac{B_s}{B_j} \right)^{\gamma_j - \eta_j} \left( \frac{N_s}{N_j} \right)^{-\eta_j} \tag{33}
\]

or, equivalently:

\[
\log \bar{\theta}_{js}^* = v_j^2 + \frac{1}{\gamma_j - \eta_j} \log B_s - \frac{\eta_j}{\gamma_j - \eta_j} \log N_s, \tag{34}
\]

where \( v_j^2 \) is an industry-specific term. These expressions show that within any industry countries with higher R&D efficiency \( B_s \) or lower barriers to knowledge spillovers \( N_s \) are more productive. Comparative advantage results from the interaction of these country-level differences with industry characteristics. Differentiating (34) yields:

\[
\frac{\partial^2 \log \bar{\theta}_{js}^*}{\partial \gamma_j \partial \log B_s} < 0, \quad \frac{\partial^2 \log \bar{\theta}_{js}^*}{\partial \eta_j \partial \log B_s} > 0, \quad \frac{\partial^2 \log \bar{\theta}_{js}^*}{\partial \nu_j \partial \log B_s} = 0,
\]

\[
\frac{\partial^2 \log \bar{\theta}_{js}^*}{\partial \gamma_j \partial \log N_s} > 0, \quad \frac{\partial^2 \log \bar{\theta}_{js}^*}{\partial \eta_j \partial \log N_s} < 0, \quad \frac{\partial^2 \log \bar{\theta}_{js}^*}{\partial \nu_j \partial \log N_s} = 0.
\]

It follows that higher R&D efficiency raises productivity relatively more in industries where domestic spillovers are stronger or the advantage of backwardness is smaller. Similarly, a reduction in barriers to knowledge spillovers increases productivity relatively more in the same set of industries. Cross-industry variation in the strength of global knowledge spillovers does not affect relative productivity levels. Proposition 4 summarizes the implications of these results for comparative advantage.

**Proposition 4.** Suppose Assumptions 1 and 2 hold. In the stationary equilibrium:

(i) Countries with higher R&D efficiency or lower barriers to knowledge spillovers have a comparative advantage in industries where the advantage of backwardness is lower or domestic knowledge spillovers are stronger;

(ii) Cross-industry differences in the strength of global knowledge spillovers do not affect comparative advantage.

The mechanisms behind Proposition 4 are closely related to those that determine international inequality in Proposition 3. The advantage of backwardness imposes a cost on more productive firms by reducing their R&D efficiency. In industries where the advantage of backwardness is lower, this cost is smaller. Therefore, the cross-country productivity gaps that support the stationary equilibrium are larger in industries where the advantage of backwardness is lower, which gives a comparative advantage to countries with higher \( B_s \) or lower \( N_s \).

More productive countries with higher R&D efficiency or lower barriers to knowledge spillovers also have a comparative advantage when domestic spillovers are stronger because such countries generate more spillovers. Consequently, in industries where domestic spillovers are stronger, the knowledge and productivity gaps between countries are larger. The strength of global spillovers does not affect comparative advantage.
because stronger global spillovers do not affect relative knowledge levels across countries and, therefore, have a symmetric effect on productivity in all countries.

Before introducing growth into the model, it is worth considering how other sources of heterogeneity across industries affect comparative advantage. Suppose the Armington elasticity $\sigma_j$, the capability distribution $G_j(\psi)$, the returns to scale in production $\beta_j$, the returns to scale in R&D $\alpha_j$, the knowledge depreciation rate $\delta_j$, the exit rate $\zeta_j$ 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. In particular, subject to this modification, the system of wage equations (28), the export equation (32) and the frontier productivity equation (34) are unaffected. This has three immediate implications.

First, Propositions 3 and 4 continue to hold meaning the model’s implications for how R&D and knowledge spillovers affect international inequality and comparative advantage are robust to allowing for these additional sources of inter-industry heterogeneity. Second, cross-industry variation in $G_j(\psi)$, $\beta_j$, $\alpha_j$, $\delta_j$, $\zeta_j$ or $f_j^E$ does not generate comparative advantage. Third, the effect of variation in the Armington elasticity $\sigma_j$ on comparative advantage is in general ambiguous because $\sigma_j$ affects exports through interactions with both the frontier productivity $\bar{\theta}_{j, s}$ and the wage level $w_s$ in (32) and, without imposing further parameter assumptions, neither effect necessarily dominates. However, if $\gamma_j$ and $\eta_j$ do not vary by industry then countries with a higher R&D efficiency or lower barriers to knowledge spillovers must have a comparative advantage in industries with higher $\sigma_j$. The intuition for this result comes from the relationship between the Armington elasticity and terms of trade effects. A higher Armington elasticity implies greater substitutability between the exports of different countries meaning terms of trade effects are weaker. Consequently, in industries where the Armington elasticity is higher the market share and exports of more productive countries are relatively greater, which gives them a comparative advantage.

5 Growth

Under Assumption 2 knowledge spillovers are too weak to sustain long-run growth. This simplifies the model and enables us to study the consequences of independent variation in $\gamma_j$, $\eta_j$ and $\nu_j$. However, observed productivity distributions are not stationary. Therefore, this section introduces long-run growth into the global productivity model.

Ongoing growth is possible if the total strength of knowledge spillovers $\eta_j + \nu_j$ equals the advantage of backwardness $\gamma_j$. Therefore, I replace Assumption 2 with the following assumption.

Assumption 3. For all industries $j$, the advantage of backwardness equals the sum of the strengths of domestic and global knowledge spillovers: $\gamma_j = \eta_j + \nu_j$.

Assumption 3 implies there are only two degrees of freedom among the three parameters $\gamma_j$, $\eta_j$ and $\nu_j$. Consequently, I let $\eta_j = \kappa_j \gamma_j$ and $\nu_j = (1 - \kappa_j) \gamma_j$ and treat $\gamma_j$ and $\kappa_j$ as independent parameters. In the growth model $\gamma_j$ plays a dual role. An increase in $\gamma_j$ raises both the advantage of backwardness and the total strength of spillovers. Meanwhile, $\kappa_j \in [0, 1)$ measures the localization of knowledge spillovers in industry.
An increase in $\kappa_j$ raises domestic spillovers relative to global spillovers implying the geographic scope of spillovers is reduced.

Using Assumption 3 and the definition of the knowledge level (15) we can rewrite the R&D technology (13) as:

$$\dot{\theta} = B_s \psi \left( \frac{1}{N_s} \theta \right)^{\kappa_j} \left( \sum_{\tilde{s}=1}^{S} \frac{1}{N_{\tilde{s}}} \theta_{j\tilde{s}} \right)^{(1-\kappa_j)\gamma_j} (IR)^\alpha - \delta. \quad (35)$$

All else constant, equation (35) shows productivity growth is decreasing in a firm’s productivity relative to both the domestic frontier and the frontier productivity levels in all other countries.\(^{19}\) Consequently, if the productivity of all firms in the world increases by the same proportion, then productivity growth is unaffected. This dependence of productivity growth on relative, rather than absolute, productivity levels is the key to obtaining sustained growth.

Bartelsman, Haskel and Martin (2008) and Griffith, Redding and Simpson (2009) estimate the impact of distance to the industry frontier on productivity growth using firm-level data for the UK. Consistent with the specification of the R&D technology in (35) both papers find lower productivity relative to the domestic frontier increases productivity growth. Bartelsman, Haskel and Martin also find lower productivity relative to the global frontier (defined as the maximum across countries of average productivity among the top quartile of firms) raises growth and that the global frontier effect is weaker than the domestic frontier effect.

Except for Assumption 3 replacing Assumption 2, the model is unchanged from Section 3 and equations (1)-(21) continue to hold. Equilibrium is defined as before and 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. The next section outlines how to solve for a balanced growth path and characterizes the properties of a balanced growth path equilibrium. Full details of the derivations and proofs of the propositions can be found in Appendix A.

### 5.1 Balanced Growth Path

On a balanced growth path the knowledge level $\chi_{js}$ must grow at a constant rate. Differentiating (15) implies:

$$\frac{\dot{\chi}_{js}}{\chi_{js}} = \gamma_j \left[ \kappa_j \frac{\dot{\theta}_{j\tilde{s}}}{\theta_{j\tilde{s}}} + (1 - \kappa_j) \frac{\sum_{\tilde{s}=1}^{S} \frac{1}{N_{\tilde{s}}} \theta_{j\tilde{s}}}{\sum_{\tilde{s}=1}^{S} \frac{1}{N_{\tilde{s}}} \theta_{j\tilde{s}}} \right],$$

which is time invariant if and only if the frontier productivity $\bar{\theta}_{j\tilde{s}}$ grows at the same rate $g_j$ in all countries. It follows that on a balanced growth path $\chi_{js}$ grows at rate $\gamma_j g_j$ and the productivity distribution $H_{js}(\theta)$ shifts outwards at rate $g_j$ for all $s$. This means $H_{js}(\theta, t) = H_{js}(e^{g_j(t-\tau)} \bar{\theta}, \tau)$ for all times $t$, $\tau$ and productivity levels $\theta$. The existence of global knowledge spillovers is necessary to ensure the productivity growth rate of

\(^{19}\)Nelson and Phelps (1966) initially proposed that, because of knowledge spillovers, productivity growth is a decreasing function of productivity relative to the frontier. For recent applications of this idea see Damsgaard and Krusell (2010) and Benhabib, Perla and Tonetti (2014).
each industry is constant across countries, since if $\kappa_j = 1$ then $\dot{\chi}_{js}/\chi_{js}$ is proportional to $\dot{\theta}_{js}/\theta_{js}$ and there is no mechanism that prevents productivity growth being country-specific on a balanced growth path.

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:

$$q = \sum_{j=1}^{J} \mu_j g_j.$$  \hspace{1cm} (36)

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 the consumption price $z$ declines 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 and entry in each country-industry pair is time invariant, as is the mass of firms $M_{js}$. Industry output $Y_{js}$ and the quantity sold in each market $x_{j\tilde{s}s}$ grow at rate $g_j$, while prices $p_{js}$ and $P_j$ 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 $\tau_s = \rho$. Since the discount rate $\rho > 0$ and nominal assets per capita remain constant over time, the transversality condition (3) is satisfied.

Is a balanced growth path consistent with optimal firm behavior? For the economy to be on a balanced growth path the productivity distribution $H_{js}(\theta)$ must shift outwards at a constant rate $g_j$. As in Section 4.1 firms choose R&D employment taking the time paths of $w_s, p_{js}$ and $\chi_{js}$ as given. Let relative productivity $\phi$ be defined by:

$$\phi \equiv \left( \frac{1}{\chi_{js}^{\gamma_j}} \right)^{\frac{1}{1-\beta}} = \left( \frac{1}{N_s} \frac{\theta_{js}}{\theta} \right)^{\frac{-\kappa_j}{1-\beta}} \left( \sum_{\tilde{s}=1}^{S} \frac{1}{N_{\tilde{s}}} \frac{\theta_{j\tilde{s}}}{\theta} \right)^{\frac{-(1-\kappa_j)}{1-\beta}}.$$  

Each firm’s relative productivity is a function of its productivity relative to both the domestic frontier and the frontiers in other countries. After changing variables from $\theta$ to $\phi$, the intertemporal optimization problem is isomorphic to the problem firms faced in the stationary model and each firm has a unique, locally saddle-path stable steady state given by:
In steady state all surviving firms in an industry have the same productivity growth rate \( g_j \). Entrants draw their capability and productivity from the joint distribution of \( \psi \) and \( \theta \) among incumbents and all incumbents face instantaneous exit probability \( \zeta \). Thus, entry, exit and R&D cause the productivity distribution \( H_{js}(\theta) \) to shift outwards at rate \( g_j \) if and only if all incumbent firms are in steady state. It follows that on a balanced growth path all firms are in steady state and each incumbent firm’s relative productivity and R&D employment are time invariant and given by (37) and (38), respectively.

Within each country-industry pair, less capable firms have lower steady state relative productivity levels and this provides an advantage of backwardness that exactly offsets the disadvantage from low \( \psi \) ensuring all firms grow at the same rate. Similarly, the existence of global knowledge spillovers ensures that level differences across countries in the location of the productivity distribution exactly 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.

On a balanced growth path, within each industry firm growth is independent of firm size. In addition, R&D intensity is independent of firm size and given by:

\[
\frac{w_s l^{R*}_{js}}{p_{js} \phi^*_{js}} = \frac{\alpha (\delta + g_j)}{\rho + \zeta + \gamma_j (\delta + g_j)}.
\]

Note that faster growing industries have higher R&D intensity. From (8) and (25) firm employment remains constant over time on a balanced growth path. This means the firm employment distribution is stationary. Luttmer (2010) observes that the US firm employment distribution appears to be stationary. Equation (37) also implies that within any country-industry pair the ratio of firms’ steady state productivity levels is given by (27) as in the stationary model. It follows that productivity inequality across firms within any country-industry pair is strictly decreasing in the advantage of backwardness and strictly increasing in the returns to scale in production and R&D.

Before moving on from firm behavior, we can use the solution to the firm’s R&D problem to develop a strategy for estimating the advantage of backwardness \( \gamma_j \) from firm-level data. Suppose for a panel of firms \( \omega \) and years \( t \) we have data on productivity \( \theta_{\omega t} \), R&D employment \( l^{R*}_{\omega t} \) and the industry productivity growth rate \( g_j \). Define \( \hat{\phi}_{\omega t} = e^{-g_j t} \theta_{\omega t} \) to be the detrended productivity of firm \( \omega \) which is constant in steady state. Let \( \hat{\phi}^*_{\omega} \) be the firm-specific steady state value of \( \hat{\phi}_{\omega t} \) and \( l^{R*} \) be steady state R&D employment at firm \( \omega \).
given by (38). Taking the time derivative of $\hat{\varphi}_{\omega t}$, using the R&D technology (35) and log-linearizing around the firm’s steady state gives:

$$
\frac{d}{dt} \log \hat{\varphi}_{\omega t} = -\gamma_j (\delta + g_j) \log \left( \frac{\hat{\varphi}_{\omega t}}{\hat{\varphi}_{\omega t}^*} \right) + \alpha (\delta + g_j) \log \left( \frac{l_{R_t}^\omega}{l_{R_t}^{\omega*}} \right),
$$

$$
= v_\omega - \gamma_j (\delta + g_j) \log \hat{\varphi}_{\omega t} + \alpha (\delta + g_j) \log l_{R_t}^\omega,
$$

(39)

where $v_\omega = \gamma_j (\delta + g_j) \log \hat{\varphi}_{\omega t}^* - \alpha (\delta + g_j) \log l_{R_t}^{\omega*}$ is a firm fixed effect. Given a shock that causes some firms’ detrended productivity levels to deviate from their steady state values, but does not affect the aggregate industry equilibrium, equation (39) could be used to estimate $\gamma_j$. Alternatively, the same data could be used to estimate $\gamma_j$ directly from (13) using industry-time fixed effects to control for the knowledge level and a firm fixed effect to control for the firm’s capability $\psi$. Estimating $\kappa_j$ imposes more stringent data requirements because of the need to separately identify variation in domestic and global spillovers. However, it would be possible given comparable productivity data from more than one country such as that used by Bartelsman, Haskel and Martin (2008).

The final step in solving for a balanced growth path is to use the economy’s general equilibrium conditions to determine the productivity growth rate $g_j$ in each industry and wages $w_s$ in each country. General equilibrium requires that for all countries $s = 1, \ldots, S$:

$$
L_s = \sum_{j=1}^J \frac{\mu_j}{\rho + \zeta} \left( \zeta + \rho \beta + \frac{\rho \alpha (\delta + g_j)}{\rho + \zeta + \gamma_j (\delta + g_j)} \right) \frac{w_s^{-\sigma} \left( \frac{1}{B_s^\gamma_j} N_s^{-\kappa_j} \right)^{\frac{\sigma-1}{1-\kappa_j}}}{\sum_{s=1}^S w_s^{1-\sigma} \left( \frac{1}{B_s^\gamma_j} N_s^{-\kappa_j} \right)^{\frac{\sigma-1}{1-\kappa_j}}},
$$

(40)

and for all industries $j = 1, \ldots, J$:

$$
(\delta + g_j)^{1-\alpha} \left( 1 - \beta + [\gamma_j (1 - \beta) - \alpha] \frac{\delta + g_j}{\rho + \zeta} \right) = \alpha f^{-\frac{1}{\psi}} \left[ \frac{\sum_{s=1}^S \left( \frac{1}{B_s^\gamma_j} N_s^{-1} \right)^{\frac{1-\kappa_j}{1-\kappa_j}}}{\int_{\psi}^\psi \left( \frac{\psi}{w} \right)^{\gamma_j (1-\beta)-\alpha} dG(\psi)} \right].
$$

(41)

Equations (40) and (41) form a system of $S + J$ equations in the $S + J$ unknowns $w_s$ and $g_j$. Since wages do not appear in (41) this system can be solved recursively. For each industry $j$ the left hand side of (41) is a strictly increasing and unbounded function of $g_j$ implying there exists a unique solution for $g_j$.\(^{20}\)

Conditional on $g_j$ for $j = 1, \ldots, J$ equation (40) is isomorphic to the general equilibrium wage equation in the stationary model (28). Therefore, (40) gives a unique solution for the wage levels $w_s$ by the same reasoning used to prove Proposition 2. Proposition 5 summarizes the balanced growth path equilibrium.

\(^{20}\)To ensure existence I assume the parameters of the global economy are such that when $g_j = 0$ the left hand side of (41) is strictly less than the right hand side. This is guaranteed if, for example, $\delta = 0$.  

27
Proposition 5. Suppose Assumptions 1 and 3 hold. There exists a unique balanced growth path equilibrium on which:

(i) The productivity growth rate within each industry is the same for all firms in all countries;
(ii) Wages, income and consumption per capita grow at the same rate in all countries.

On the 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, as in the stationary model, cross-country heterogeneity does generate international gaps in productivity levels, wages and incomes. Section 5.2 studies the global productivity distribution that supports the balanced growth path equilibrium and analyzes the origins of comparative advantage.

Using (36) and (41) we can characterize the determinants of growth. Consumption growth \( q \) is strictly increasing in each of the industry growth rates \( g_j \). Because of global knowledge spillovers an increase in R&D efficiency in any country raises growth in all countries meaning \( g_j \) is increasing in the R&D efficiency \( B_s \) of each country \( s \). Likewise, \( g_j \) is decreasing in \( N_s \) for all \( s \) because higher barriers to knowledge spillovers in any country lower the knowledge level and, consequently, the productivity growth rate, in all countries.

The industry growth rate \( g_j \) is increasing in the entry cost \( f_E \) and firms’ effective discount rate \( \rho + \zeta \). An increase in either the cost of entry or the rate at which firm’s discount future profits leads to an increase in \( p_{js}/w_s \) to ensure the free entry condition is satisfied. Higher output prices relative to wages raise static profits relative to R&D costs which tends to increase R&D investment and growth. There is also a direct negative effect of higher \( \rho + \zeta \) on R&D investment due to firms’ discounting future profits more heavily, but this is more than offset by the indirect effect operating through \( p_{js}/w_s \). These relationships are the opposite of those found in expanding variety (Romer 1990) and quality ladders (Aghion and Howitt 1992) endogenous growth models. In these models and in the idea flows model of Sampson (2016a) entry drives growth and higher entry costs or greater discounting reduce entry and lower growth. By contrast, in this paper incumbent R&D investment drives growth and lower entry reduces the competition faced by incumbent firms which leads to faster growth. This illustrates how, when growth results from incumbent firm R&D, the determinants of growth differ from those in economies where entrants are the source of growth.

Growth \( g_j \) is also decreasing in the knowledge depreciation rate \( \delta \), but is independent of population \( L_s \). As in Sampson (2016a) there are no scale effects on growth because an increase in population does not affect the location of the productivity distribution in any country and, therefore, does not generate knowledge spillovers. Growth is increasing in the returns to scale in production \( \beta \) because a higher \( \beta \) reduces the concavity in firms’ intertemporal optimization problems. This raises the returns to productivity growth leading firms to invest more in R&D.

A higher maximum R&D capability \( \psi \) increases growth by making frontier firms more efficient at R&D, which raises their productivity and generates spillovers that increase the knowledge level everywhere. However, conditional on \( \psi \), growth is decreasing in the capability of non-frontier firms. This occurs because firms that are behind the frontier do not generate knowledge spillovers, but do compete with frontier firms. When these firms have lower capabilities and, consequently, lower relative productivity levels they employ
less labor and produce less output, which raises \( p_{js}/w_s \) leading to an increase in R&D investment at frontier firms causing higher knowledge spillovers and growth. In general, the effects of the advantage of backwardness \( \gamma_j \) and the localization of knowledge spillovers \( \kappa_j \) on \( g_j \) are ambiguous. But if there is a single country with \( N_s \geq 1 \) then a higher advantage of backwardness reduces growth because it compresses the productivity distribution and increases the competition faced by frontier firms. Proposition 6 summarizes the determinants of growth.

**Proposition 6.** Suppose Assumptions 1 and 3 hold. On the balanced growth path, the growth rate \( g_j \) in any industry \( j \) is strictly increasing in the R&D efficiency \( B_s \) of all countries, the entry cost \( f^E \), the discount rate \( \rho \), the exit rate \( \zeta \), the returns to scale in production \( \beta \) and the maximum R&D capability \( \psi \); it is strictly decreasing in the barriers to knowledge spillovers \( N_s \) in all countries, the knowledge depreciation rate \( \delta \) and the relative capabilities \( \psi/\bar{\psi} \) of non-frontier firms.

### 5.2 The Global Productivity Distribution

On the balanced growth path all firms have the same productivity growth rate, implying the relative productivity of any pair of firms remains constant over time. The relative productivity of two firms with the same capability that belong to the same industry, but different countries is given by:

\[
\frac{\theta^*_{js}}{\theta^*_{j\bar{s}}} = \left( \frac{B_s}{B_{\bar{s}}} \right)^{\frac{1}{1-\kappa_j}} \left( \frac{N_s}{N_{\bar{s}}} \right)^{-\frac{\kappa_j}{1-\kappa_j}}.
\]

Note that substituting \( \kappa_j = \eta_j/\gamma_j \) into this expression yields equation (33), which gives relative productivity levels in the stationary equilibrium. Thus, the determinants of productivity differences on the balanced growth path are the same as in the stationary equilibrium except that under Assumption 3 there are only two degrees of freedom among the parameters that govern the advantage of backwardness and the strength of knowledge spillovers.

Each country’s relative productivity is increasing in its relative R&D efficiency and decreasing in its relative barriers to knowledge spillovers. Intuitively, a country where R&D investment is less effective or barriers restrict knowledge flows falls behind the global frontier until its advantage of backwardness is sufficient to offset poor fundamentals ensuring it grows at the same rate as other countries. Equation (42) gives the relative productivity levels that support this balanced growth path.

The elasticity of relative productivity to R&D efficiency is decreasing in \( \gamma_j \) and increasing in the localization of knowledge spillovers \( \kappa_j \). Recall that \( \gamma_j \) has a dual role when Assumption 3 holds. An increase in \( \gamma_j \) raises both the advantage of backwardness and the total strength of spillovers. From the stationary model we know that an increase in the advantage of backwardness tends to reduce the elasticity of relative productivity to R&D efficiency, while an increase in the total strength of spillovers holding \( \kappa_j \) constant has the opposite effect because it raises the strength of domestic spillovers. Equation (42) shows that the former effect dominates on the balanced growth path. An increase in \( \kappa_j \) raises the elasticity of relative productivity to R&D efficiency because when spillovers are more localized they disproportionately benefit more productive countries that generate more spillovers. An increase in \( \kappa_j \) also increases the cost of barriers to
knowledge spillovers by making the elasticity of relative productivity to $N_s$ more negative. Again, this results from more localized spillovers increasing the knowledge gap between countries. However, in contrast to the stationary model, the elasticity of relative productivity to $N_s$ is independent of $\gamma_j$ because, in this case, the dual effects of variation in $\gamma_j$ exactly offset each other.

Suppose the ranking of countries by productivity levels is the same as the ranking by R&D efficiency and by the inverse of barriers to knowledge spillovers meaning that more productive countries have both higher $B_s$ and lower $N_s$. Then international productivity inequality within each industry is decreasing in the advantage of backwardness and increasing in the localization of knowledge spillovers. Keller (2002) finds knowledge spillovers have become less sensitive to distance and more localized over time. Interpreting this as evidence of a decline in $\kappa_j$ the global productivity model predicts convergence in relative productivity levels within industries, which is consistent with the results of Levchenko and Zhang (2016).

Using (42) we can also analyze how wages, consumption and incomes differ across countries and the pattern of comparative advantage. As in the stationary model, when $J = 1$ assets per capita $a_s$ and consumption per capita $c_s$ are proportional to wages $w_s$ and relative wages are given by (31). Thus, relative wages are increasing in relative productivity levels and the discussion of the determinants of relative productivity above also applies to relative wages, consumption per capita and income per capita. To conserve space I omit any further details.

Turning to comparative advantage, balanced growth path exports can be written as:

$$
\log \text{EX}_{js} = v_j^3 + \xi^2_s + (\sigma - 1) \log \overline{\theta}_{js} - (\sigma - 1) \log w_s,
$$

which is analogous to equation (32) in the stationary model.\footnote{In this expression both the industry-specific term $v_j^3$ and the frontier productivity $\log \overline{\theta}_{js}$ are non-stationary, while the value of exports, the destination country-specific term $\xi^2_s$ and wages $w_s$ are time invariant due to the choice of global consumption expenditure as the numeraire. The analysis in this section compares exports and productivity levels across countries and industries at a point in time.} Thus, the pattern of comparative advantage depends upon the location of the productivity distribution in each country-industry pair as captured by $\log \overline{\theta}_{js}$. From (42) it follows:

$$
\log \overline{\theta}_{js} = v_j^4 + \frac{1}{(1 - \kappa_j)\gamma_j} \log B_s - \frac{\kappa_j}{1 - \kappa_j} \log N_s,
$$

and differentiating this expression yields:

$$
\frac{\partial^2 \log \overline{\theta}_{js}}{\partial \gamma_j \partial \log B_s} < 0, \quad \frac{\partial^2 \log \overline{\theta}_{js}}{\partial \kappa_j \partial \log B_s} > 0,
$$

$$
\frac{\partial^2 \log \overline{\theta}_{js}}{\partial \gamma_j \partial \log N_s} = 0, \quad \frac{\partial^2 \log \overline{\theta}_{js}}{\partial \eta_j \partial \log N_s} < 0.
$$

This means countries with higher R&D efficiency have a comparative advantage in industries where the advantage of backwardness is lower or knowledge spillovers are more localized. Countries with lower barriers to knowledge spillovers also have a comparative advantage in industries with more local knowledge
spillovers, but their comparative advantage is unaffected by the advantage of backwardness. Proposition 7 summarizes these results.

**Proposition 7.** Suppose Assumptions 1 and 3 hold. On the balanced growth path:

(i) Countries with higher R&D efficiency have a comparative advantage in industries where the advantage of backwardness is smaller;

(ii) Countries with higher R&D efficiency or lower barriers to knowledge spillovers have a comparative advantage in industries where the localization of knowledge spillovers is greater.

Proposition 7 characterizes how comparative advantage depends upon cross-industry differences in knowledge spillovers and learning. On the balanced growth path, a stable pattern of Ricardian comparative advantage exists because international productivity gaps support an equilibrium in which productivity growth does not vary by country. It is the advantage of backwardness which makes the existence of such an equilibrium possible. In industries where innovation is harder relative to learning, the advantage of backwardness is stronger meaning productivity dispersion due to differences in R&D efficiency is smaller and countries that are less efficient at R&D have a comparative advantage. Comparative advantage also depends upon the geographic scope of knowledge spillovers. When domestic spillovers are stronger relative to global spillovers the advantage of more productive countries is magnified implying countries with a higher R&D efficiency or lower barriers to knowledge spillovers have a comparative advantage.

### 6 Conclusions

There are substantial and persistent productivity differences across firms, industries and countries. These differences give rise to both Ricardian comparative advantage and international income inequality. Any explanation of these differences must account for how firm-level innovation and learning determine productivity. Some firms and countries are better at R&D and knowledge absorption than others and the ease with which firms can discover new ideas and improve existing technologies differs across industries, as does the strength and geographic scope of knowledge spillovers. Yet developing a general equilibrium model of the global economy that incorporates these heterogeneities is far from straightforward. Building on the ideas that there exists an advantage of backwardness and knowledge spillovers depend upon the location of the productivity distribution, this paper presents a new theory of global productivity differences when productivity growth is driven by incumbent firm R&D.

The theory generates a rich set of predictions concerning comparative advantage. In particular, it highlights how the advantage of backwardness and the localization of knowledge spillovers shape comparative advantage. Countries with an absolute advantage in R&D have a comparative advantage in industries where the advantage of backwardness is lower and knowledge spillovers are more local. Countries where there are high barriers to the generation of knowledge spillovers have a comparative advantage in industries where spillovers are more global. The paper also analyzes how firms’ R&D choices affect firm size inequality within industries, income inequality across countries and the global growth rate.

Beyond the observation there exists a positive correlation between countries’ measured R&D expenditure as a share of GDP and their high-tech exports (Fagerberg 1995), there has been surprisingly little
empirical work on the dynamic origins of Ricardian comparative advantage. This may, in part, be due to a lack of testable predictions derived from economic theory. Hopefully, the paper will stimulate further empirical work in this area.

The framework developed in this paper should also prove useful in addressing other important questions related to productivity. Governments often view industrial policy as a tool to promote export growth in particular industries. The theory could be used to analyze how optimal industrial policy design depends upon the R&D technology, knowledge spillovers and country characteristics. The model could also be applied to analyze inter-sectoral spillovers and to study the important role of foreign direct investment in international technology diffusion. Alternatively, extending the model to include products and monopolistic competition would allow for an analysis of how growth depends upon the interaction of incumbent R&D and selection in a dynamic Melitz (2003) environment.

Finally, endogenous growth models that include incumbent firm R&D map the origins of growth to observable firm behavior. Patent data can be used to estimate the innovation technology in quality ladders models with incumbent innovation (Akcigit and Kerr 2016), but is not informative about learning by firms behind the technology frontier. By linking the R&D technology and knowledge spillovers to firm-level productivity dynamics, the framework developed in this paper offers a promising way to use firm data to impose greater empirical discipline on models of learning and growth.

References


Ricardo, David. 1817. “Principles of Political Economy and Taxation.”


Appendices

A Proofs and Derivations

Solution to firm’s R&D problem in Section 4.1

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:

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

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

\[ 0 = \frac{\partial H}{\partial l^R} = -w_s + \lambda \frac{\alpha}{1-\beta} B_s \chi_{js} \psi \phi^{1-\gamma_j(1-\beta)} (l^R)^{\alpha-1}, \]

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

\[ 0 = \lim_{\tau \to \infty} \left[ e^{-(\rho+\zeta)(\tau-t)} H(\phi, l^R, \lambda) \right], \tag{44} \]

where equation (44) is the transversality condition. Differentiating the upper expression with respect to \( \tau \) gives:

\[ (1 - \alpha) \frac{\dot{l}^R}{l^R} = [1 - \gamma_j(1 - \beta)] \frac{\dot{\phi}}{\phi} + \frac{\dot{\lambda}}{\lambda}, \tag{45} \]

and using the first order conditions of the Hamiltonian to substitute for \( \lambda \) and \( \dot{\lambda} \) and (22) to substitute for \( \dot{\phi} \) we obtain equation (23).

Equations (22) and (23) are an autonomous nonlinear system of differential equations in \((\phi, l^R)\) whose unique steady state \((\phi^*_{js}, l^*_{js})\) is given by (24) and (25). 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 \( DF \) of the function \( F \) is:

\[ DF \begin{pmatrix} \phi^*_{js} \\ l^*_{js} \end{pmatrix} = \begin{pmatrix} \frac{\alpha}{1-\beta} \phi^*_{js} \delta & -\gamma_j \delta \\ \frac{1-\gamma_j(1-\beta)}{1-\alpha} \frac{l^*_{js}}{\phi^*_{js}} (\rho + \zeta + \gamma_j \delta) & \rho + \zeta + \gamma_j \delta \end{pmatrix}. \]
The trace of the Jacobian is $\rho + \zeta$ which is positive. The determinant of the Jacobian is:

$$\begin{vmatrix} DF \left( \frac{\phi^{*}_{js}}{l^{R*}_{js}} \right) \end{vmatrix} = -\delta (\rho + \zeta + \gamma j \delta) \frac{\gamma j (1 - \beta) - \alpha}{(1 - \alpha)(1 - \beta)},$$

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 $\phi$ lies within this neighborhood, the system of differential equations given by (22) and (23) 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 (45) 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 (44).

The solution to (22) and (23) 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 $\lambda$ is the current-value costate variable obtained from the solution to (22) and (23). Equation (43) implies $\lambda$ is always strictly positive. Therefore, given any path for $\phi$ on which $\phi$ is always positive we have $\lim_{\tau \to \infty} \left[ e^{-\left(\rho + \zeta\right)(\tau - t)} \lambda \phi \right] \geq 0$.

Now define:

$$\overline{H}(\phi, \lambda) = \max_{l^{R}} \mathcal{H}(\phi, l^{R}, \lambda),$$

$$= \left[ 1 - \beta \left( \frac{\beta p_{js}}{w_{s}} \right)^{\frac{1}{1 - \beta}} w_{s} - \frac{\lambda \delta}{1 - \beta} \right] \phi + \frac{1 - \alpha}{\alpha} \frac{\beta p_{js}}{w_{s}} \left( \frac{\alpha \lambda B_{s} \chi_{js} \psi}{1 - \beta} \right)^{\frac{1}{1 - \alpha}} \phi^{1 - \gamma j (1 - \beta)} \phi,$$

where the second line follows from solving the maximization problem in the first line. Assumption 1 implies $\overline{H}(\phi, \lambda)$ is strictly concave in $\phi$. Thus, the sufficiency conditions of Theorem 7.14 in Acemoglu (2009) hold, implying the solution to (22) and (23) is the unique solution to the firm’s optimal control problem.

**Derivation of equation (28)**

Suppose the global economy is in a stationary equilibrium. Using (10), (14), (24) and (25) yields that, in a stationary equilibrium, the steady state value of a firm with capability $\psi$ is:

$$V_{js}(\psi, \theta_{js}^{*}) = \left( 1 - \beta - \frac{\alpha \delta}{\rho + \zeta + \gamma j \delta} \right) \frac{w_{s}}{\rho} - \frac{\lambda \delta}{1 - \beta} \left[ \alpha^{\gamma j \beta} B_{s} \chi_{js} \psi \left( \frac{p_{js}}{w_{s}} \right)^{\gamma j} \frac{\delta^{\alpha - 1}}{(\rho + \zeta + \gamma j \delta)^{\alpha}} \right] \frac{1}{\gamma j (1 - \beta) - \alpha},$$

where $\theta_{js}^{*} = (\phi_{js}^{*})^{1 - \beta}$ is the firm’s steady state productivity. Assumption 1 implies $1 - \beta > \frac{\alpha \delta}{\rho + \zeta + \gamma j \delta}$ which ensures $V_{js}(\psi, \theta_{js}^{*})$ is positive. Section 4.1 showed that in a stationary equilibrium 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 (16) yields:

38
\[ f_E = \left(1 - \beta - \frac{\alpha \delta}{\rho + \zeta + \gamma_j \delta}\right) \frac{1}{\rho + \zeta} \int_{\psi}^{\psi_0} \psi \gamma_j (1 - \beta - \alpha) dG(\psi) \times \left[ \alpha \beta \gamma_j B_{s} \chi_{js} \left( \frac{p_{j,s}}{w_{s}} \right)^{\gamma_j} \frac{\delta^{\gamma_j}}{(\rho + \zeta + \gamma_j \delta)^{\alpha}} \right] \]  

(46)

Define \( \chi_{js} = \left( \sum_{s=1}^{S} \overline{\theta}_{js} N_{s}^{\gamma_j} \right)^{\nu_j} \). From (15) the knowledge level \( \chi_{js} \) can then be written as:

\[ \chi_{js} = \overline{\theta}_{js} N_{s}^{\gamma_j} \chi_{j} \]  

(47)

and using (24) to substitute for the steady state technical efficiency frontier \( \overline{\theta}_{js} = \left( \theta_{js}^{\ast} \right)^{1 - \beta} \) implies:

\[ \chi_{js} = \left( N_{s}^{\gamma_j} \chi_{j} \right)^{\nu_j \left( \gamma_j (1 - \beta) - \alpha \right)} \left[ \alpha \beta \gamma_j B_{s} \psi \right]^{\frac{1}{\alpha}} \left( \frac{p_{j,s}}{w_{s}} \right)^{\frac{1}{\alpha}} \frac{\delta^{\gamma_j - \eta_j}}{(\rho + \zeta + \gamma_j \delta)^{\alpha}} \right]^{\frac{\eta_j \gamma_j (1 - \beta)}{(\gamma_j - \eta_j) (1 - \beta) - \alpha}} \]  

(48)

Note that Assumption 1 ensures \( (\gamma_j - \eta_j)(1 - \beta) - \alpha > 0 \).

Next, observe that in a stationary equilibrium:

\[ \int_{\theta}^{\psi_{0}} \psi \psi_{0} dH_{js}(\theta) = \int_{\psi}^{\psi_{0}} \psi \phi_{js}^{*} dG(\psi), \]

where \( \phi_{js}^{*} \) is given by (24) and depends upon \( \psi \). Thus, by substituting (11) and (25) into the labor market clearing condition (18) and using (17) with \( \dot{M}_{js} = 0 \) to solve for \( L_{js}^{E} \), we obtain:

\[ L_{s} = \sum_{j=1}^{J} M_{js} \left\{ \int_{\psi}^{\psi_{0}} \psi \gamma_j (1 - \beta - \alpha) dG(\psi) \left(1 + \frac{\beta \rho + \zeta + \gamma_j \delta}{\alpha \delta}\right) \right. \]

\[ \times \left. \left[ \alpha \gamma_j (1 - \beta) \beta \gamma_j B_{s} \chi_{js} \left( \frac{p_{j,s}}{w_{s}} \right)^{\gamma_j} \frac{1}{\delta} \left( \frac{\delta}{\rho + \zeta + \gamma_j \delta} \right)^{\gamma_j (1 - \beta)} \gamma_j (1 - \beta - \alpha) \right]^{\gamma_j (1 - \beta - \alpha)} \right\} \]  

(49)

Similarly, substituting (4), (6), (12), (21) and (24) into the goods market clearing condition (19) we obtain:

\[ \frac{\sigma + 1 - \gamma_j}{\gamma_j (1 - \beta - \alpha)} \frac{\gamma_j}{\gamma_j (1 - \beta - \alpha)} = \mu_{j} \left( \frac{p_{j,s}^{\sigma - 1}}{M_{js} w_{s}} \right) \left[ \int_{\psi}^{\psi_{0}} \psi \gamma_j (1 - \beta - \alpha) dG(\psi) \right]^{-1} \]

\[ \times \left[ \alpha \gamma_j (1 - \beta) \beta \gamma_j B_{s} \chi_{js} \left( \frac{1}{w_{s}} \right)^{\gamma_j} \frac{\delta^{\gamma_j - 1}}{(\rho + \zeta + \gamma_j \delta)^{\alpha}} \right]^{\gamma_j (1 - \beta - \alpha)} \]  

(50)

Equations (46)-(50) together with equation (7), which gives the industry price index, form a system of equations that can be used to solve for \( w_{s}, M_{js}, p_{j,s}, \chi_{js} \) and \( P_{j} \). Equation (28) is obtained by simplifying...
this system of equations in the following manner. Start by using (50) and then (46) to substitute for \( M_{js} \) and \( \chi_{js} \) in (49) to obtain:

\[
L_s = \sum_{j=1}^{J} \frac{\mu_j}{\rho + \zeta} \left( \zeta + \rho \beta + \frac{\rho \alpha \delta}{\rho + \zeta + \gamma_j \delta} \right) \frac{1}{w_s} \left( \frac{P_j}{P_{js}} \right)^{\sigma-1}.
\]

(51)

Next use (48) to substitute for \( \chi_{js} \) in (46) giving:

\[
\frac{p_{js}}{w_s} = \left[ \alpha^{\alpha} \beta^{\beta} (\gamma_j - \eta_j) B_s N_s^{-\eta_j} \chi_j \psi \right]^{-\frac{1}{\gamma_j - \eta_j}}
\times \left\{ \int \left( \frac{\psi}{\psi} \right)^{1 - \beta - \frac{\alpha \delta}{\rho + \zeta + \gamma_j \delta}} \left[ \int \frac{1}{\psi} \frac{dG(\psi)}{\psi} \right]^{-1} \right\}^{-(\gamma_j - \eta_j)(1 - \beta) - \alpha},
\]

and from the definition of the industry price index (7) it then follows that:

\[
\left( \frac{P_j}{P_{js}} \right)^{\sigma-1} = \frac{w_s^{1-\sigma} B_s N_s^{-\nu_j} \chi_j \psi}{\sum_{s=1}^{S} w_s^{1-\sigma} B_s N_s^{-\nu_j} \chi_j \psi}.
\]

(53)

Finally, substituting this equation into the expression above for \( L_s \) gives equation (28).

**Proof of Proposition 2**

The proof of Proposition 2 uses results from Allen, Arkolakis and Li (2015). Let \( w = (w_1, \ldots, w_S) \) be the \( S \)-dimensional wage vector and by subtracting \( L_s \) from both sides write equation (28) as \( f(w) = 0 \) where \( f : \mathbb{R}^S_+ \to \mathbb{R}^S \). Let \( f_s(w) \) denote element \( s \) of the vector \( f \). For all \( s = 1, \ldots, S \) define the scaffold function \( F : \mathbb{R}^{S+1}_+ \to \mathbb{R}^S \) by:

\[
F_s(\tilde{\mathbf{w}}, w_s) = \sum_{j=1}^{J} \frac{\mu_j}{\rho + \zeta} \left( \zeta + \rho \beta + \frac{\rho \alpha \delta}{\rho + \zeta + \gamma_j \delta} \right) \frac{w_s^{-\sigma} B_s N_s^{-\eta_j} \chi_j \psi}{\sum_{s=1}^{S} w_s^{-\sigma} B_s N_s^{-\eta_j} \chi_j \psi} - L_s.
\]

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

To prove the existence of a solution to equation (28) it is now sufficient to show that conditions (i)-(iii) of Lemma 1 in Allen, Arkolakis and Li (2015) are satisfied. Condition (i) follows from observing that, for any \( \tilde{\mathbf{w}} \), \( F_s(\tilde{\mathbf{w}}, w_s) \) is strictly decreasing in \( w_s \), positive for \( w_s \) sufficiently close to zero and negative for \( w_s \) sufficiently large. To see that condition (ii) holds, note that \( 1 - \sigma < 0 \) implying \( F_s(\tilde{\mathbf{w}}, w_s) \) is strictly increasing in \( \tilde{w}_s \) for all \( s \).

Now, given \( \lambda > 0 \) and \( \tilde{\mathbf{w}} \in \mathbb{R}^S_+ \) define \( w_s(\lambda) \) by \( F_s[\lambda \tilde{\mathbf{w}}, w_s(\lambda)] = 0 \). Let \( u \in (0, 1) \) be such that \(-1 + \sigma u < 0\). Then \( F_s[\lambda \tilde{\mathbf{w}}, \lambda^{-u} w_s(1)] \) is strictly negative if \( \lambda > 1 \) and strictly positive if \( \lambda < 1 \). Since
\( F_s(\tilde{w}, w_s) \) 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 to equation (28) exists.

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

**Derivation of equation (30)**

The steady state productivity of a firm with capability \( \psi \) in industry \( j \) and country \( s \) is given by (24) and \( \theta^*_j = (\phi^*_j)^{1-\beta} \). Therefore, the steady state productivity ratio of two firms with the same capability in the same industry, but different countries is:

\[
\frac{\theta^*_j}{\theta^*_{\tilde{j}}} = \left[ \frac{B_s \chi_{j}s}{B_{\tilde{s}} \chi_{j}\tilde{s}} \right]^{1-\beta} \left[ \frac{p_{j}s w_{\tilde{s}}}{p_{j}\tilde{s} w_s} \right]^{\alpha} \left[ \frac{\gamma_j (1-\sigma) - \alpha}{\rho + \zeta} \right].
\]

Using (48) and (52) to substitute for \( \chi_{j}s/\chi_{j}\tilde{s} \) and \( (p_{j}s w_{\tilde{s}})/(p_{j}\tilde{s} w_s) \) then implies that in the stationary equilibrium:

\[
\frac{\theta^*_j}{\theta^*_{\tilde{j}}} = \left( \frac{B_s}{B_{\tilde{s}}} \right)^{1-\beta} \left( \frac{N_s}{N_{\tilde{s}}} \right)^{\eta_j/\gamma_j - \eta_j}. \]

This yields equation (30) in the single industry case.

**Proof of Proposition 3**

Start by substituting (46) into (50) to obtain:

\[
M_{js} = \frac{1}{f^E} \frac{w_j}{w_s} \left( \frac{P_j}{p_{j}s} \right)^{\sigma-1} \left( 1 - \beta - \frac{\alpha \delta}{\rho + \zeta + \gamma_j \delta} \right) \frac{1}{\rho + \zeta}. \]

Setting \( J = 1 \) and using (51) then gives:

\[
\left( \zeta + \rho \beta + \frac{\rho \alpha \delta}{\rho + \zeta + \gamma_j \delta} \right) M_s = \frac{1}{f^E} \left( 1 - \beta - \frac{\alpha \delta}{\rho + \zeta + \gamma_j \delta} \right) L_s.
\]

Next, substitute the free entry condition (16) into the asset market clearing condition (20) yields:

\[
a_s L_s = \sum_{j=1}^{J} M_{js} f^E w_s. \quad (54)
\]

Setting \( J = 1 \) and using \( M_s \propto L_s \) then implies that in the single industry case assets per capita \( a_s \) is proportional to wages \( w_s \).
To complete the proof note that setting $\dot{a}_s = 0$ in (1) and using $\iota_s = \rho$ gives:

$$zc_s = \rho a_s + w_s.$$  

It follows that consumption per capita is proportional to wages in the stationary equilibrium.

**Proof of Proposition 4**

To derive equation (32) start by substituting (4) and (6) into $EX_{js\tilde{s}} = p_j x_{js\tilde{s}}$ to show:

$$EX_{js\tilde{s}} = \left( \frac{P_j}{p_{js}} \right)^{\sigma - 1} \mu_j zc_s L_{\tilde{s}}.$$  

Then use (53) to substitute for $P_j/p_{js}$ and take logarithms to obtain:

$$\log EX_{js\tilde{s}} = v_j^5 + \xi_1^s + \frac{\sigma - 1}{\gamma_j - \eta_j} \log B_s - \frac{(\sigma - 1)\eta_j}{\gamma_j - \eta_j} \log N_s - (\sigma - 1) \log w_s.$$  

(55)

Substituting (34) into the expression above gives equation (32). When interpreting (32) and (34) it is important to remember that $v_j^j, v_j^j, v_j^5$ and $\xi_1^s$ are endogenous and, in general, will vary across stationary equilibria.

The remainder of the proposition follow from differentiating (34).

**Other sources of comparative advantage in Section 4.4**

To see that $G_j(\psi), \beta_j, \alpha_j, \delta_j, \xi_j$ and $f_j^E$ do not affect comparative advantage note that they only affect log exports through $v_j^j$ in (55).

When the Armington elasticity varies across industries we have:

$$\frac{\partial^2 \log EX_{js\tilde{s}}}{\partial \sigma_j \partial \log B_s} = \frac{1}{\gamma_j - \eta_j} - \frac{\partial \log w_s}{\partial \log B_s},$$  

$$\frac{\partial^2 \log EX_{js\tilde{s}}}{\partial \sigma_j \partial \log N_s} = -\frac{\eta_j}{\gamma_j - \eta_j} - \frac{\partial \log w_s}{\partial \log N_s}.$$  

Wages are given by (28). Since the summation in the denominator of (28) does not vary by country, we can analyze wage variation across countries on the balanced growth path by differentiating (28) holding the summation term fixed. This yields:

$$\frac{\partial \log w_s}{\partial \log B_s} = \frac{\sum_{j=1}^J \frac{\sigma_j - 1}{\sigma_j} \frac{1}{\gamma_j - \eta_j} \lambda_{js}}{\sum_{j=1}^J \lambda_{js}},$$  

$$\frac{\partial \log w_s}{\partial \log N_s} = -\frac{\sum_{j=1}^J \frac{\sigma_j - 1}{\sigma_j} \frac{\eta_j}{\gamma_j - \eta_j} \lambda_{js}}{\sum_{j=1}^J \lambda_{js}}.$$
where:
\[
\lambda_{js} = \frac{\mu_j}{\rho + \zeta_j} \left( \zeta_j + \rho \beta_j + \frac{\rho \alpha_j \delta_j}{\rho + \zeta_j + \gamma_j \delta_j} \right) \frac{w_{s}^{-\sigma_j} (B_s N_s^{-\eta_j})^{\frac{\sigma_j - 1}{\gamma_j - \eta_j}}}{{\sum_{s=1}^{S} w_{s}^{1-\sigma_j} (B_s N_s^{-\eta_j})^{\frac{\sigma_j - 1}{\gamma_j - \eta_j}}}}.
\]

meaning \(\sum_{j=1}^{J} \lambda_{js} = L_s\) by (28). Note that wages are strictly increasing in \(B_s\) and strictly decreasing in \(N_s\).

If \(\gamma_j\) and \(\eta_j\) do not vary across industries then \(\frac{\partial \log w_s}{\partial \log B_s} < 1\) implying \(\frac{\partial^2 \log EX_{js}}{\partial \sigma_j \partial \log B_s} > 0\). Similarly, \(\frac{\partial^2 \log EX_{js}}{\partial \sigma_j \partial \log N_s} < 0\). More generally, it is possible that the sign of each of these expressions varies across industries implying the absence of a clear pattern of comparative advantage, but it cannot be the case that \(\frac{\partial^2 \log EX_{js}}{\partial \sigma_j \partial \log B_s}\) is always negative or \(\frac{\partial^2 \log EX_{js}}{\partial \sigma_j \partial \log N_s}\) is always positive.

**Analysis of growth rates in Section 5.1**

On a balanced growth path the individual’s budget constraint (1) gives:

\[
\dot{w}_s = \frac{\dot{a}_s}{a_s} = q_s + \frac{\dot{z}}{z}. \tag{56}
\]

Next, the growth rate of production employment can be obtained by differentiating (11). Since the productivity distribution \(H_{js}(\theta)\) shifts outwards at rate \(g_j\) this implies:

\[
\frac{\dot{L}_{js}^P}{L_{js}^P} = \frac{\dot{M}_{js}}{M_{js}} + \frac{1}{1 - \beta} \left( \frac{\dot{p}_j}{p_j} - \frac{\dot{w}_s}{w_s} + g_j \right). \tag{57}
\]

There is no population growth, meaning that on a balanced growth path \(\dot{L}_{js}^P = \dot{L}_{js}^E = 0\). Therefore, (54) and (56) imply \(\dot{M}_{js} = 0\). Using (56) together with the expression above we then have:

\[
q_s = \frac{\dot{p}_j}{p_j} + g_j - \frac{\dot{z}}{z}. \tag{57}
\]

Now, differentiating the industry price index (7) yields:

\[
\frac{\dot{P}_j}{P_j} = \frac{\sum_{s=1}^{S} p_{js}^{1-\sigma} \dot{p}_{js}}{\sum_{s=1}^{S} p_{js}^{1-\sigma}},
\]

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

\[
\frac{\dot{p}_{js}}{p_{js}} = \frac{\dot{P}_j}{P_j}.
\]

Substituting this expression into (57) implies implies the growth rate of consumption per capita \(q_s = q\) is the same in all countries and given by:
\[ q = \frac{\dot{P}_j}{P_j} + g_j - \frac{\dot{z}}{z}. \]  

(58)

Differentiating the numeraire condition (21) we obtain:

\[ q = -\frac{\dot{z}}{z}. \]  

(59)

and substituting this result into (58) shows the industry price index \( P_j \) declines at rate \( g_j \). Therefore, differentiating (5) implies:

\[ q = -\frac{\dot{z}}{z} = -\sum_{j=1}^{J} \mu_j \frac{\dot{P}_j}{P_j} = \sum_{j=1}^{J} \mu_j g_j, \]

which shows equation (36) holds. To obtain \( \tau_s = \rho \) we can then substitute (59) into the Euler equation (2).

Note also that using (4) to substitute for \( X_{js} \) in (6) and appealing to (59) together with the fact prices decline at rate \( g_j \) implies \( x_{js} \) grows at rate \( g_j \). It then follows from the industry output market clearing condition (19) that industry output \( Y_{js} \) also grows at rate \( g_j \).

**Solution to firm’s R&D problem in Section 5.1**

On a balanced growth path, \( \tau_s = \rho \) and \( w_s \) are time invariant, \( p_{js} \) declines at rate \( g_j \) and \( \chi_{js} \) grows at rate \( \gamma_j g_j \). Taking the time derivative of \( \phi \) and using the R&D technology (13) implies:

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

(60)

R&D raises a firm’s relative productivity, while growth in the knowledge level causes relative productivity to decline.

Substituting the production profits function (10) into the value function (14) and changing variables from \( \theta \) to \( \phi \), the optimization problem of a firm with capability \( \psi \) can be written as:

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

subject to the growth of \( \phi \) being given by (60) and an initial value for the firm’s relative productivity at time \( t \). Since \( \chi_{js} \) grows at rate \( \gamma_j g_j \), \( p_{js} \) declines at rate \( g_j \) and \( w_s \) is time invariant, the term in square brackets in the integrand is independent of time \( \tau \). Therefore, the payoff function depends upon time only through exponential discounting and the firm faces a discounted infinite-horizon optimal control problem with state variable \( \phi \) and control variable \( l^R \). This problem is isomorphic to the firm’s R&D problem in Section 4.1 and has a unique, locally saddle-path stable steady state given by equations (37) and (38).
Derivation of equations (40) and (41)

The derivation of equation (40) follows the same steps used to derive equation (28) in the stationary equilibrium. To avoid repetition I omit the details.

To derive equation (41) start by substituting (47) and $\theta^*_j = \chi_j \phi^*_j$ into (37) which gives:

$$
\theta^*_j = \left[ \frac{\alpha \beta}{1-\beta} B N^{-\kappa_j \gamma_j} \psi \bar{x}_j \left( \frac{p_j}{w_s} \right)^{\frac{\alpha}{1-\beta}} \frac{(\delta + g_j)^{\alpha-1}}{[\rho + \zeta + \gamma_j (\delta + g_j)]^{\alpha}} \right]^{1-\beta}.
$$

Then use the balanced growth path equivalent of (52) to substitute for $p_j w_s$ in the above equation and set $\psi = \psi$ before substituting the resulting expression into $\bar{x}_j = \left( \sum_{s=1}^{S} \frac{\bar{\theta}_s}{N_s} \right)^{1-\kappa_j \gamma_j}$ to obtain (41).

Proof of Proposition 7

From (61) we have that the relative productivity of two firms in industry $j$ with the same capability $\psi$, but in different countries is:

$$
\frac{\theta^*_j}{\theta^*_j} = \left[ \frac{B_j}{B_i} \left( \frac{N_i}{N_j} \right)^{-\kappa_j \gamma_j} \left( \frac{p_j w_s}{p_j w_s} \right)^{\frac{\alpha}{1-\beta}} \right]^{1-\beta^{1-\kappa_j \gamma_j}}.
$$

Using the balanced growth path equivalent of (52) to substitute for $p_j w_s$ in this expression gives (42). The proof then follows from the same reasoning used to prove Proposition 4.