Knowledge Diffusion, Trade and Innovation across Countries and Sectors

Jie Cai, Nan Li and Ana Maria Santacreu*

October 30, 2017

Abstract

We develop a quantifiable multi-country, multi-sector endogenous growth model in which comparative advantage is endogenously determined by innovation and knowledge diffusion. We quantify the effect of a trade liberalization on innovation, comparative advantage and welfare. Changes in trade frictions reallocate innovation and comparative advantage across sectors: innovation reallocates towards sectors with larger increases in comparative advantage, and comparative advantage reallocates towards sectors with stronger knowledge spillovers. Knowledge spillovers amplify the effect as countries and sectors benefit from technology developed elsewhere. In contrast to one-sector models without knowledge spillovers, we find significant dynamic gains from trade, driven by innovation and diffusion.

Keywords: Knowledge spillovers; welfare gains from trade; endogenous growth; R&D

JEL Classification: F12, O33, O41, O47

*Cai, Shanghai University of Finance and Economics, cai.jie@mail.shufe.edu.cn; Li, International Monetary Fund, address: 700 19th Street NW, Washington DC 20431, email: nanli1@gmail.com; Santacreu, Federal Reserve Bank of St. Louis, email: am.santacreu@gmail.com. We are grateful to Lorenzo Caliendo, Jonathan Eaton, Sam Kortum, Andrei Levchenko, Fernando Parro, and Dan Trefler for useful comments, as well as and participants of the seminars at Yale University, 2017 and 2016 Society for Economic Dynamics, the 2017 Econometric Society Summer Meeting, IMF, Federal Reserve Bank of St. Louis, Penn State University, RIDGE, Fudan University, and Tsinghua University for helpful comments. The views in this paper are those of the authors and do not necessarily reflect the views of the Federal Reserve Banks of St. Louis or the Federal Reserve System. The views expressed herein are those of the authors and should not be attributed to the IMF, its Executive Board, or its management.

1 Introduction

The world has increasingly become a highly interconnected network of countries and sectors that not only trade goods and services but also exchange ideas. Recently, a growing strand of the trade literature has examined how the benefits of trade liberalization may spread across sectors through production input-output linkages (Caliendo and Parro (2015)). However, countries and sectors are also linked along a different dimension—innovation and knowledge diffusion. Indeed, technological advances never happen in isolation (David (1990), Rosenberg (1982)). Knowledge in one country and sector (country-sector) can be used to enhance innovation in another, and much like production input-output linkages, knowledge diffusion across countries and sectors is far from uniform. Therefore, in a world with multiple sectors, when changes in trade costs alter the knowledge composition of the economy, the latter also conditions trade patterns and aggregate growth (as shown in the empirical research by Hausmann, Hwang, and Rodrik (2007), Hidalgo, Klinger, Barabási, and Hausmann (2007)). Furthermore, although trade flows often serve as a vehicle for knowledge diffusion (Alvarez, Buera, and Lucas Jr (2013)), other channels may also diffuse ideas across countries and sectors. The literature so far has either treated these two channels as separate issues or has modeled them together as one channel (e.g., more trade necessarily implies more knowledge spillovers).

We develop a multi-sector and multi-country endogenous growth model in which comparative advantage and the composition of knowledge are endogenously determined by innovation and knowledge diffusion. We study the effect of changes in trade costs on innovation, comparative advantage and welfare. The production side of the model is a multi-sector version of Eaton and Kortum (2002), as in Caliendo and Parro (2015). Different from those papers, which assume that technology is exogenous, we introduce dynamics through innovation and knowledge diffusion.

Knowledge diffusion, which occurs when a firm that operates in a country-sector learns about ideas developed in other countries and sectors, is exogenous. That is, innovations developed in one country-sector diffuse to other countries and sectors at an exogenous rate. In contrast to models that do not feature knowledge spillovers, our model can deliver convergence in relative productivity at the sector level, which has been documented by several empirical studies (see Levchenko and Zhang (2016), Hausmann and Klinger (2007), Cameron, Proudman, and Redding (2005), Proudman and Redding (2000), and Bernard and Jones (1996a,b)). In a Ricardian framework, this evolution of sectoral productivity changes comparative advantage and welfare. Moreover, knowledge spillovers amplify the effect of trade liberalizations on innovation and welfare, as countries and industries have access to a larger pool of ideas (Keller (2004)).

Our model has implications for welfare gains from trade that differ from both static models of trade and dynamic one-sector models of trade and innovation without knowledge spillovers. Relative to static models, the endogenous evolution of comparative advantage provides an additional source of welfare gains. Changes in trade costs cause changes in innovation which, through knowledge flows, spread across countries and generate changes in revealed comparative advantage, hence welfare. The effect of trade on innovation is driven by both the multi-sector dimension of our model and the presence of knowledge spillovers. Moreover, long-run growth rates are endogenous in our model and reductions in trade frictions increase the growth rate along the balanced-growth path (BGP). Standard one-sector models of trade and innovation find a negligible impact of trade on innovation, growth and welfare (Atkeson and Burstein (2010) and Buera and Oberfield (2016)). In contrast, our model generates dynamic gains from trade.

Despite its complexity, the model comes with the benefit of tractability, as we build upon the Ricardian trade model of Eaton and Kortum (2002) with Bertrand competition (Bernard, Eaton, Jensen, and Kortum (2003)). The innovation and international technology diffusion processes are modeled in a similar fashion as those in Eaton and Kortum (1996, 1999). Technology evolves over time through two channels: (i) Innovators in each sector invest final output to introduce a new idea, which if successful can be used to produce an intermediate good. The efficiency of innovation in a country-sector increases with the stock of knowledge available in that particular country-sector. (ii) Ideas diffuse across both sectors and countries according to an exogenous process of diffusion.

The model is solved in two stages. Given the distribution of firm productivity together with trade barriers, we solve for a static competitive equilibrium for the world economy. The equilibrium is static in that we take as given the technology levels that determine the patterns of trade. We then allow for the technology profile to evolve endogenously owing to a process of innovation and diffusion. The second stage allows us to characterize the innovation and knowledge diffusion processes that drive the endogenous evolution of comparative advantage and dynamic welfare gains from trade. A similar approach has been used in Alvarez, Buera, and Lucas Jr (2008). Different from their paper, our diffusion channel produces a Frechet distribution of productivity, as in Eaton and Kortum (1999). Furthermore, Alvarez, Buera, and Lucas Jr (2008) abstract from innovation, which is a key channel in our model. Buera and Oberfield (2016) study a model with technology diffusion and innovation that delivers the same Frechet distribution. However, different from the predictions of our model, trade has no impact on innovation in their framework, as they consider a one-sector model.

We calibrate the model to cross-country and cross-sector data on patent citations, sectoral research and development (R&D) intensity, production and international trade. The production structure of the model delivers a gravity equation at the sector level that can be estimated to obtain both trade barriers and the level of technology for every sector-country pair, as in Levchenko and Zhang (2016). Cross-country and cross-sector patent citations allow us to discipline the direction and speed at which knowledge in a particular sector-country is utilized in the innovation of other sectors in other countries.[1] Finally, data on R&D intensity at the sector-country level allow us to calibrate the parameters that govern the evolution of technology through innovation. Our model is able to capture the rich heterogeneity of the data along multiple dimensions.

We conduct a counterfactual exercise to study the effect of a uniform trade liberalization on innovation, comparative advantage and welfare. In contrast to one-sector models of trade and innovation without knowledge spillovers, changes in trade frictions have a non-negligible effect on innovation in our model, as there is a reallocation of R&D toward sectors in which the country has comparative advantage.[2] Knowledge diffusion is an additional source of technological progress in our paper, and it is important to explain convergence in relative productivity found in the data or to explain growth miracles as in Buera and Oberfield (2016). We find that comparative advantage reallocates towards sectors with stronger knowledge spillovers.

Our quantitative framework has implications for welfare gains from trade, which we decompose into static and dynamic gains. A trade liberalization strengthens a country's comparative advantage and hence the static gains from trade. In addition, there are dynamic gains from trade due to a reallocation of R&D across sectors and knowledge diffusion. Knowledge diffusion has two opposite effects on welfare. On the one hand, it enables faster productivity convergence and makes countries more similar to each other, dampening the gains from trade. On the other hand, it provides strong dynamic gains, because countries can innovate with access to a larger foreign knowledge pool. In our quantitative exercise, the second channel dominates and knowledge spillovers introduce additional gains from trade.

We find that after a trade liberalization the following occurs: (i) Larger countries experience lower gains from trade. (ii) R&D reallocates towards sectors that experience larger increases in revealed comparative advantage; comparative advantage reallocates towards sectors with larger knowledge spillovers. This is especially the case in those countries that have

[1] Starting from Griliches (1981), patenting statistics have been used to proxy innovation in the large literature of firm innovation. In addition, since the work by Jaffe, Trajtenberg, and Henderson (1993), patent citations have come to be considered as the most informative tool for the purpose of tracing knowledge flows. Research using cross-country patent or citation data include Eaton and Kortum (1996, 1999), Mancusi (2008) and Bottazzi and Peri (2003).

[2] These result has been found in Somale (2014) in a semi-endogenous model of growth with multiple sectors and no knowledge diffusion.

larger gains from trade. (iii) An increase in innovation translates into an increase in the growth rate and income per capita on the BGP. As a result, we find significant dynamic gains from trade, which are heterogeneous across countries.

Finally, we study the role of different channels of our model and find the following: (i) In a world in which we do not allow for knowledge spillovers, welfare gains from trade are smaller, especially for smaller countries that spend less on R&D, as they have more to gain from the diffusion of foreign ideas. In the extreme case of instantaneous diffusion, welfare gains from trade are larger. (ii) A one-sector version of our model without knowledge spillovers generates negligible dynamic gains. R&D intensity and the growth rate stay the same, as the effect of larger market access associated with a trade liberalization is canceled out by the effect of larger import competition.

Our paper merges and extends several strands of existing literature. The first is the literature on innovation, diffusion and international trade. Eaton and Kortum (1996) and Eaton and Kortum (1999) posit technological innovations and their international diffusion through trade as potential channels of embodied technological progress. Santacreu (2015) develops a model in which trade allows countries to adopt innovations that have been developed abroad, and thus diffusion does not take place without trade. Our main departure from these previous papers is that we consider a multi-sector environment in which sectors interact both in the product space and in the technology space. In addition, we allow knowledge diffusion and trade to operate separately, even though common economic forces may contribute to the development of both and diffusion and trade may benefit and reinforce each other.

The second is the strand of literature that extends Eaton and Kortum (2002) to a multi-sector model of trade (Chor (2010), Costinot, Donaldson, and Komunjer (2012)). A recent growing body of research in this area also explores the trade and growth implications of interdependence across different sectors through intermediate input-output relationships (Eaton, Kortum, Neiman, and Romalis (2016), Caliendo and Parro (2015)). Our paper differs in several dimensions. First, our focus is on innovation and knowledge diffusion. Second, besides the factor-demand linkages, this paper also simultaneously considers the intrinsic interconnections of technologies embodied in different sectors, which turns out to be significant and relevant when studying innovation and diffusion (Cai and Li (2017)). Related to the current work, Cai and Li (2016) study knowledge spillovers across sectors within a country and how trade costs affect the distribution of endogenous knowledge accumulation across sectors. Different from our paper, however, cross-sector knowledge diffusion is not considered across countries and intermediate input-demand linkages across sectors are absent. Perla, Tonetti, and Waugh (2015) study the effect of trade on growth in a symmetric country model in which firms learn from existing knowledge from other firms.

Our paper is related to two recent papers that feature multi-sector trade models of innovation with endogenous comparative advantage: Somale (2014) and Sampson (2016). Somale (2014) studies a multi-sector semi-endogenous model of trade and innovation without knowledge spillovers. He focuses on the role of innovation to generate endogenous comparative advantage. Different from his paper, we introduce knowledge spillovers in our model as an additional source of comparative advantage. Furthermore, we use data on R&D at the sector level to discipline the innovation process of the model, so that we can capture explicitly the effect of trade on innovation. With our methodology, we are able to identify what sectors in a country experience increases in R&D after a trade liberalization, and we link these results to welfare gains from trade. Sampson (2016) develops a theoretical Armington framework of innovation and learning as sources of endogenous comparative advantage. Different from his paper, our emphasis is on the quantification of the model, which allows us to do counterfactuals.

The rest of the paper proceeds as follows. Section 2 presents the model. Section 3 describes the balanced-growth path, Section 4 describes the calculation of welfare gains from trade. Section 5 explores the quantitative implications of our model. Finally, Section 6 concludes.

2 The Model

We develop a general equilibrium model of trade in intermediate goods, with sector heterogeneity and input-output linkages, in which technology evolves endogenously through innovation and knowledge diffusion. The model builds upon the Ricardian trade model of Eaton and Kortum (2002) with Bertrand competition (Bernard, Eaton, Jensen, and Kortum (2003)). The innovation and diffusion processes are modeled as in Eaton and Kortum (1996, 1999).

There are M countries and J sectors. Countries are denoted by i and n and sectors are denoted by j and k. Labor is the only factor of production, and we assume it to be mobile across sectors within a country but immobile across countries. In each country, there is a representative consumer who consumes a non-traded final good and saves. A perfectly competitive final producer in the country combines the composite output of each J sectors in the domestic economy with a Cobb-Douglas production function. In each sector there is a producer of a composite good that operates under perfect competition and sells the good to the final producer and to intermediate producers from all sectors in that country. Intermediate producers use labor and composite goods of each other sector in that country to produce varieties that are traded and are used by the composite producer of that sector, either domestic or foreign. These firms operate under Bertrand competition and are heterogeneous in their productivity. Trade is Ricardian. Finally, the technology of each sector evolves endogenously through innovation and technology diffusion. The innovation process follows the quality-ladders literature in that new innovations increase the quality of the product in a given sector. Diffusion is assumed to be exogenous. Foreign firms that decide to use a domestic innovation pay royalties to the innovator. We assume that royalty payments are perfectly enforced.

2.1 Consumers

In each country there is a representative household who chooses consumption optimally to maximize its life-time utility

$$U_{nt} = \int_{t=0}^{\infty} \rho^t u\left(C_{nt}\right) dt, \tag{1}$$

where $\rho \in (0,1)$ is the discount factor and C_{nt} represents consumption of country n at time t. The household finances R&D activities of the entrepreneurs and owns all the firms.

We assume that household's preferences are represented by a CRRA utility function

$$u\left(C_{nt}\right) = \frac{C_{nt}^{1-\gamma}}{1-\gamma}$$

with an intertemporal elasticity of substitution, $\gamma > 0$.

2.2 Final Production

Domestic final producers use the composite output from each domestic sector j in country n at time t, Y_{nt}^j, to produce a non-traded final output Y_{nt} according to the following Cobb-Douglas production function:

$$Y_{nt} = \prod_{j=1}^{J} \left(Y_{nt}^j\right)^{\alpha^j}, \tag{2}$$

with $\alpha^j \in (0,1)$ the share of sector production on total final output, and $\sum_{j=1}^{J} \alpha^j = 1$.

Final producers operate under perfect competition. Their profits are given by

$$\Pi_{nt} = P_{nt} Y_{nt} - \sum_{j=1}^{J} P_{nt}^j Y_{nt}^j,$$

where P_{nt} is the price of the final product and P_{nt}^j is the price of the composite good produced in sector j from country n.

Under perfect competition, the price charged by the final producer to the consumers is equal to the marginal cost; that is

$$P_{nt} = \prod_{j=1}^{J} \left(\frac{P_{nt}^j}{\alpha^j} \right)^{\alpha^j} .$$

The demand by final producers for the sector composite good is given by

$$Y_{nt}^j = \alpha^j \frac{P_{nt}}{P_{nt}^j} Y_{nt}.$$

2.3 Intermediate Producers

In each sector j there is a continuum of intermediate producers indexed by $\omega \in [0,1]$ that use labor, $l_{nt}^j(\omega)$, and a composite intermediate good from every other sector k in the country, $m_{nt}^{jk}(\omega)$, to produce a variety ω according to the following constant returns to scale technology[3]:

$$q_{nt}^j(\omega) = z_n^j(\omega)[l_{nt}^j(\omega)]^{\gamma^j} \prod_{k=1}^{J} [m_{nt}^{jk}(\omega)]^{\gamma^{jk}}, \tag{3}$$

with $\gamma^j + \sum_{k=1}^{J} \gamma^{jk} = 1$. Here γ^{jk} is the share of materials from sector k used in the production of intermediate ω is sector j, and γ^j is the share of value added. Firms are heterogeneous in their productivity $z_n^j(\omega)$.

The cost of producing each intermediate good ω is

$$c_{nt}^j(\omega) = \frac{c_{nt}^j}{z_{nt}^j(\omega)},$$

where c_n^j denotes the cost of the input bundle. With constant returns to scale,

$$c_{nt}^j = \Upsilon^j W_{nt}^{\gamma^j} \prod_{k=1}^{J} (P_{nt}^k)^{\gamma^{jk}}, \tag{4}$$

with $\Upsilon^j = \prod_{k=1}^{J} (\gamma^{jk})^{-\gamma^{jk}} (\gamma^j)^{-\gamma^j}$ and W_{nt} the nominal wage rate.

2.4 Composite Intermediate Goods (Materials)

Each sector j produces a composite good combining domestic and foreign varieties from that sector. Composite producers operate under perfect competition and buy intermediate

[3]The notation in the paper is such that every time there are two subscripts or two superscripts, the one on the right corresponds to the source country and the one on the left corresponds to the destination country.

products ω from the minimum cost supplier.

The production for a composite good in sector j and country n is given by the Ethier (1982) CES function,

$$Q^j_{nt} = \left(\int r^j_{nt}(\omega)^{1-1/\sigma} d\omega \right)^{\sigma/(\sigma-1)}, \tag{5}$$

where $\sigma > 0$ is the elasticity of substitution across intermediate goods and $r^j_{nt}(\omega)$ is the demand of intermediate goods from the lowest cost supplier in sector j.

The demand for each intermediate good ω is given by

$$r^j_{nt}(\omega) = \left(\frac{p^j_{nt}(\omega)}{P^j_{nt}} \right)^{-\sigma} Q^j_{nt},$$

where

$$P^j_{nt} = \left(\int p^j_{nt}(\omega)^{1-\sigma} d\omega \right)^{\frac{1}{1-\sigma}}. \tag{6}$$

Composite intermediate goods are used as final goods in the final production and as materials for the production of the intermediate goods:

$$Q^j_{nt} = Y^j_{nt} + \sum_{k=1}^{J} \int m^{kj}_{nt}(\omega) d\omega.$$

2.5 International Trade

Trade in goods is costly. In particular, there are iceberg transport costs from shipping a good that is produced in sector j from country i to country n, $d^j_{ni} > 1$. We follow Bernard, Eaton, Jensen, and Kortum (2003) and assume Bertrand competition.

The p^{th} most efficient producer of variety ω from sector j in country i has productivity z^j_{pi} and can deliver one unit of goods to country n at cost

$$c^j_{pni}(\omega) = d^j_{ni} \frac{c^j_i}{z^j_{pi}(\omega)}.$$

With Bertrand competition, as with perfect competition, composite producers in each sector buy from the lowest cost supplier. The lowest cost of good ω in country n is given by

$$c^j_{1n}(\omega) = min_i \left\{ c^j_{1ni}(\omega) \right\}.$$

In addition, Bertrand competition implies that the price charged by the producer will be the production cost of the second-lowest producer:

$$c^j_{2n}(\omega) = min\left\{c^j_{2ni^*}(\omega), min_{i\neq i^*}\{c^j_{1ni}(\omega)\}\right\},$$

where i^* satisfies $c^j_{1ni^*}(\omega) = c^j_{1n}(\omega)$. The low-cost supplier will not want to charge a markup above $\bar{m} = \sigma/(\sigma-1)$. Hence,

$$p^j_n(\omega) = min\left\{c^j_{2n}(\omega), \bar{m}c^j_{1n}(\omega)\right\}.$$

Ricardian motives for trade are introduced as in Eaton and Kortum (2002), since productivity is allowed to vary by country-sector. The productivity of producing intermediate good ω in country i and sector j is drawn from a Frechet distribution with parameter T^j_i and shape parameter θ. A higher T^j_i implies a higher average productivity of that country-sector, while a lower θ implies more dispersion of productivity across varieties:

$$F(z^j_i) = Pr\left[Z \leq z^j_i\right] = e^{-T^j_{it}z^{-\theta}},$$

and

$$Pr\left[p^j_{ni,t} < p\right] = 1 - e^{-T^j_{it}\left(d^j_{ni}c^j_{it}/p\right)^{-\theta}}.$$

Because each sector j in country n buys goods from the second cheapest supplier, the cost of producing good ω in sector j and country n is $p^j_{nt}(\omega) = min_i\left\{p^j_{nit}(\omega)\right\}$. Then, $c^j_{nt}(\omega)$ are realizations from

$$G^j_n(p) = 1 - \prod_{i=1}^{M}\left(Pr\left[p^j_{nit} > p\right]\right) = 1 - \prod_{i=1}^{M} e^{-T^j_{it}\left(d^j_{ni}c^j_{it}/p\right)^{-\theta}} = 1 - e^{-\Phi^j_{nt}p},$$

with Φ^j_{nt} each country n and sector j accumulated technology expressed as

$$\Phi^j_{nt} = \sum_{i=1}^{M} T^j_{it}(d^j_{ni}c^j_{it})^{-\theta}. \tag{7}$$

From here, we can obtain the distribution of prices of goods in sector j in country n as

$$P^j_{nt} = B\left(\Phi^j_{nt}\right)^{-1/\theta}, \tag{8}$$

with $B = \left[\frac{1+\theta-\sigma+(\sigma-1)(\bar{m})^{-\theta}}{1+\theta-\sigma}\Gamma\left(\frac{2\theta+1-\sigma}{\theta}\right)\right]^{1/(1-\sigma)}$. For prices to be well defined, we assume

$\sigma < (1+\theta)$.[4]

2.6 Expenditure shares

The probability that country i is the lowest cost supplier of a good in sector j to be exported to country n is

$$\pi_{nit}^j = \frac{T_{it}^j \left(c_{it}^j d_{ni}^j\right)^{-\theta}}{\Phi_{nt}^j}, \tag{9}$$

where π_{nit}^j is also the fraction of goods that sector j in country i sells to any sector in country n. In particular, the share country n spends on sector j products from country i is

$$\pi_{nit}^j = \frac{X_{nit}^j}{X_{nt}^j}. \tag{10}$$

2.7 Endogenous Growth: Innovation and International Technology Diffusion

We model the innovation process within each industry in a country as in Kortum (1997). Innovation follows the quality-ladders literature, in that a blueprint (i.e. an idea) is needed to produce an intermediate good. Ideas are developed with effort, and they increase the efficiency of production of an intermediate good. In each sector j and country n, there are entrepreneurs that invest final output to come up with an idea. Within each sector, research efforts are targeted at any of the continuum of intermediate goods. In each country n and sector j, ideas are drawn at a Poissson rate: If a fraction of final output, s_{nt}^j, is invested into R&D by the entrepreneur, then ideas are created at the rate

$$\lambda_{nt}^j \left(s_{nt}^j\right)^{\beta_r}, \tag{11}$$

with $\lambda_{nt}^j = \lambda_n^j A_{nt}^j$, where λ_n^j is a scaling parameter that captures the efficiency of innovation in sector j in country n, A_{nt}^j is the stock of knowledge in sector j and country n, and $\beta_r \in (0,1)$ is a parameter of diminishing returns to investing into R&D. This process has been microfounded in Eaton and Kortum (1996, 1999) and it ensures that there is a balanced-growth path without scale effects. Note that λ_{nt}^j depends positively on the stock of knowledge, A_{nt}^j; that is, countries that have accumulated more knowledge have higher efficiency of innovation.

Once an idea has arrived in sector j and country n, there is no forgetting. New ideas

[4] Details of these derivations can be found in Bernard, Eaton, Jensen, and Kortum (2003).

created in each sector j and country n increase its average productivity, A^j_{nt}.

Ideas may also diffuse exogenously to other sectors and countries. Through diffusion, the stock of knowledge in each sector j and country n is composed of knowledge that has been developed by each sector k in country i. An idea discovered at time t in country i and sector k diffuses to country n and sector j at time $t + \tau^{jk}_{ni}$. We assume that the diffusion lag, τ^{jk}_{ni}, has an exponential distribution with parameter ε^{jk}_{ni} as the speed of diffusion, so that $Pr[\tau^{jk}_{ni} \leq x] = 1 - e^{-\varepsilon^{jk}_{ni}x}$. Therefore, the flow of ideas diffusing to country n and sector j is given by

$$\dot{A}^j_{nt} = \sum_{i=1}^{M}\sum_{k=1}^{J} \varepsilon^{jk}_{ni} \int_{-\infty}^{t} e^{-\varepsilon^{jk}_{ni}(t-s)} \lambda^k_{is} \left(s^k_{is}\right)^{\beta_r} ds. \tag{12}$$

Therefore, the growth of the stock of knowledge in a particular sector j and country n at time t depends on the past research effort by each other sector k in each other country i up to time t and diffused at rate ε^{jk}_{ni}. If $\varepsilon^{jk}_{ni} \to \infty$, then there is instantaneous diffusion. If $\varepsilon^{jk}_{ni} \to 0$, then there is no diffusion.

2.7.1 The Incentives to Innovate

Following the quality-ladders literature, an idea is the realization of two random variables. One is the good ω to which the idea applies. An idea applies to only one good in the continuum. The good ω to which it is associated is drawn from the uniform distribution $[0, 1]$. The other is the quality of the idea, $q^j(\omega)$, which is drawn from the Pareto distribution $H(q) = 1 - q^{-\theta}$. In equilibrium, only the best idea for each input is actually used to produce an intermediate good in any sector and country. In that case, the idea can be used to produce an intermediate product ω in sector j and country n with efficiency $z^j_n(\omega)$. Therefore, the efficient technology $z^j_n(\omega)$ for producing good ω in country n is the best idea for producing it yet discovered. This modeling choice follows Eaton and Kortum (2006), with some modifications.

The stock of ideas at time t in each sector j and country n is A^j_{nt}. Because there is a unit interval of intermediate goods, the number of ideas for producing a specific good is Poisson with parameter A^j_{nt}. This Poisson arrival implies that the probability of k ideas for producing a good by date t in sector j and country n is $\left(A^j_{nt}\right)^k e^{-A^j_{nt}}/k!$. If there are k ideas, the probability that the best one is below the best quality q is $[H(q)]^k$. Summing over all possible k, $F(q) = e^{-A^j_{nt}q^{-\theta}}$. Therefore, the quality distribution of successful ideas inherits the distribution of productivity of the intermediate goods produced in a country. Our probabilistic distribution assumptions for the quality of an idea imply that the probability of an idea being successful is $1/A^j_{nt}$. This introduces a competitive effect, by which the larger the stock of knowledge in a sector-country, the lower the probability that the new idea lowers the cost there.

Entrepreneurs finance R&D by issuing equity claims to the households. These claims pay nothing if the entrepreneur is not successful in introducing a new technology in the market, and it pays the stream of future profits if the innovation succeeds. The value of a successful innovation in a particular sector is the expected flow of profits that will last until a new producer is able to produce the good at a lower cost. Because of the probabilistic distribution of productivity, entrepreneurs will be indifferent to what product ω to devote their efforts all products within a sector deliver the same expected profit.

The profits of an innovator in sector j in country n have two components. First, the expected future profit from selling the product in that sector and country,

$$\int_{t}^{\infty} \left(\frac{P_{nt}^{j}}{P_{ns}^{j}} \right) e^{-\rho(s-t)} \frac{\Pi_{ns}^{j}}{A_{ns}^{j}} ds, \tag{13}$$

where Π_{ns}^{j} is the profit of sector j in country n at time t and is expressed as follows

$$\Pi_{nt}^{j} = \frac{\sum_{i=1}^{M} \pi_{int}^{j} X_{it}^{j}}{1+\theta}. \tag{14}$$

Second, the innovator gets royalties from those technologies that have diffused to other countries and sectors and that have been used to produce an intermediate product there. We assume that royalty payments are perfectly enforced and are proportional to the profits that successful intermediate good producers in other sectors and countries obtain from using that technology. The expected royalty payment to an entrepreneur in country n and sector j from a technology that has been diffused and adopted by a producer in sector k of country i is

$$\sum_{i=1 i\neq n}^{M} \sum_{k=1 k\neq j}^{J} \int_{t}^{\infty} \left(\frac{P_{it}^{k}}{P_{is}^{k}} \right) e^{-\rho(s-t)} (1 - e^{-\varepsilon_{in}^{kj}(s-t)}) \frac{\Pi_{is}^{k}}{A_{is}^{k}} ds, \tag{15}$$

where $(1 - e^{-\varepsilon_{in}^{kj}(s-t)})$ is the probability that the idea created in country n has been diffused to country i.

Hence, the value of an idea that has been developed in country n and sector j is

$$V_{nt}^{j} = \int_{t}^{\infty} \left(\frac{P_{nt}^{j}}{P_{ns}^{j}} \right) e^{-\rho(s-t)} \frac{\Pi_{ns}^{j}}{A_{ns}^{j}} ds + \sum_{i=1 i\neq n}^{M} \sum_{k=1 k\neq j}^{J} \int_{t}^{\infty} \left(\frac{P_{it}^{k}}{P_{is}^{k}} \right) e^{-\rho(s-t)} (1 - e^{-\varepsilon_{in}^{kj}(s-t)}) \frac{\Pi_{is}^{k}}{A_{is}^{k}} ds. \tag{16}$$

Even if there were no royalties in the model, the innovator in a country-sector still makes profits from those ideas used in that country-sector that are sold domestically and abroad through exports, as in equation (14).

The first-order condition for optimal R&D is

$$\beta_r \lambda_{nt}^j V_{nt}^j \left(s_{nt}^j\right)^{\beta_r - 1} = P_{nt} Y_{nt}. \tag{17}$$

Therefore, the optimal R&D investment is a positive function of the value of an innovation, V_n^j, and the efficiency of innovation, λ_n^j, and is[5]

$$s_{nt}^j = \left(\beta_r \lambda_{nt}^j \frac{V_{nt}^j}{P_{nt} Y_{nt}}\right)^{\frac{1}{1-\beta_r}}. \tag{18}$$

2.8 Productivity and Comparative Advantage

We assume that the average productivity of each sector j in country i, T_{it}^j, is driven by two components: (i) The first is a time-varying component, A_{it}^j, which reflects the stock of knowledge of country i in sector j at time t. We refer to this component as "knowledge-related productivity," and it reflects the part of productivity that is driven by innovation and diffusion of foreign innovations through knowledge spillovers, as we describe in more detail in Section 2.7; (ii) The second is a time-invariant component, $T_{i,p}^j$, which captures the part of productivity that is not explained by innovation or diffusion. Factor endowments, institutions, geography, multinational production, or human capital could be factors embodied in this component.

Without loss of generality, we make the following assumption[6]:

$$T_{it}^j = A_{it}^j T_{p,i}^j. \tag{19}$$

Therefore, the dynamics of the level of technology T_{it}^j are driven by the dynamics of the knowledge-related productivity, A_{it}^j. As it will be clear in our quantitative exercise, $T_{p,i}^j$ is computed as a residual following development accounting. On the one hand, we will be able

[5]The optimization problem of the innovator is as follows. Innovators choose the amount of final output to be allocated into R&D. In our model, s_n^j is the fraction of final output that is spent on R&D activity. Therefore, innovators choose $S_{nt}^j = s_{nt}^j Y_{nt}$ to maximize

$$\dot{A}_n^j V_n^j - P_n S_n^j$$

subject to equation (12).

[6]This formulation is similar to the one introduced in Arkolakis, Ramondo, Rodríguez-Clare, and Yeaple (2013).

to identify A_{it}^{j} from innovation and diffusion data; on the other hand, we identify T_{it}^{j} from trade data. The part of T_{it}^{j} that cannot be explained by innovation and diffusion is $T_{i,p}^{j}$.

2.9 Balance of Payments

The current account balance equals the trade balance plus the net foreign income derived from net royalty payments. Total imports in country n are given by

$$IM_{nt} = \sum_{i=1 i \neq n}^{M} \sum_{k=1 k \neq j}^{J} X_{nit}^{k} = \sum_{k=1 k \neq j}^{J} X_{nt}^{k} \sum_{i=1 i \neq n}^{M} \pi_{nit}^{k}. \tag{20}$$

Total exports in country n are given by

$$EX_{nt} = \sum_{i=1 i \neq n}^{M} \sum_{k=1 k \neq j}^{J} X_{int}^{k} = \sum_{i=1 i \neq n}^{M} \sum_{k=1 k \neq j}^{J} \pi_{int}^{k} X_{it}^{k}.$$

Net royalty payments are given by

$$RP_{nt} = \sum_{j=1}^{J} RP_{nt}^{j}$$

and

$$RP_{nt}^{j} = \sum_{i=1 i \neq n}^{M} \sum_{k=1 k \neq j}^{J} \left(\chi_{in,t}^{kj} \Pi_{i}^{kt} - \chi_{ni,t}^{jk} \Pi_{nt}^{j} \right),$$

where $\chi_{in,t}^{kj}$ is the fraction of technologies developed by entrepreneurs from sector j in country n that are used by sector k in country i.

The balance of payments implies

$$EX_{nt} = IM_{nt} - RP_{nt}.$$

3 Endogenous Growth along the Balanced-Growth Path (BGP)

We now derive an expression for the growth rate of the economy on the BGP that can be used to understand the main mechanisms of the model. International and cross-sector diffusion guarantee that the knowledge-related productivity A_{nt}^{j} (and from equation (19) also average productivity, T_{nt}^{j}), grows at a common rate, g_A, across all countries and sectors. We normalize all the endogenous variables so that they are constant on the BGP and denote the normalized variables with a hat; therefore, we remove time subscripts in our derivation.

From the resource constraint in equation (67), the fraction of final output that is invested into R&D, s_n^j, is constant on the BGP. This result, together with the expression for the value of an innovation, implies that

$$\hat{V}_n^j = \frac{1}{\rho - g_y + g_A}\frac{\Pi_n^j}{\hat{A}_n^j} + \sum_{i=1}^{M}\sum_{k=1}^{J}\left(\frac{1}{\rho - g_y + g_A} - \frac{1}{\rho - g_y + \varepsilon_{in}^{kj} + g_A}\right)\frac{\hat{\Pi}_i^k}{\hat{A}_i^k},$$

with $\hat{V}_n^j = \frac{V_n^j A_M^J}{W_M}$, $\hat{A}_i^k = \frac{A_i^k}{A_M^J}$, and $\hat{\Pi}_i^k = \frac{\Pi_i^k}{W_M}$, and where χ_{in}^{kj} is the fraction of profits that a firm in sector k, country i pays to the innovator in sector j, country n as royalties. We impose $\rho - g_y + g_A > 0$ and derive an expression for g_y in Appendix E.

Profits are given by

$$\hat{\Pi}_n^j = \frac{\sum_{i=1}^{M}\pi_{in}^j\hat{X}_i^j}{(1+\theta)}.$$

On the BGP, the fraction of profits paid by producers from sector k in country i that use technologies from sector j in country n is

$$\chi_{in}^{kj} = \frac{\varepsilon_{in}^{kj}}{g_A + \varepsilon_{in}^{kj}}$$

with $\chi_{nn}^{jj} = 1$.

Royalty payments are given by the fraction of technologies diffused from (n, j) to (i, k) and used to produce intermediate goods in (i, k). The fraction of technologies from (n, j) that have diffused to (i, k) is $\varepsilon_{in}^{kj}\int_{-\infty}^{t} e^{-\varepsilon_{in}^{kj}(t-s)}\dot{A}_{ns}^j ds/A_{nt}^j$. On the BGP, this is equal to $\frac{\varepsilon_{in}^{kj} g_A}{\varepsilon_{in}^{kj}+g_A}$.

To gain some intuition on why trade has an effect on R&D, let's assume that there are no royalties. Then,

$$s_n^j = \left(\beta_r\lambda_n^j\frac{1}{(1+\theta)}\frac{1}{\rho - g_y + g_A}\frac{\sum_{i=1}^{M}\pi_{in}^j\hat{X}_i^j}{\hat{Y}_n}\right)^{\frac{1}{1-\beta_r}}, \tag{21}$$

with $\hat{X}_i^j = \frac{X_i^j}{W_M}$ and $\hat{Y}_n = \frac{P_n Y_n}{W_M}$. Trade affects optimal investment into R&D at the sector level to the extent that it affects the reallocation of production into particular sectors. This result differs from previous papers in the literature that find that trade has no impact on R&D intensity. In our model, R&D reallocates towards sectors in which the country has a comparative advantage, through $\frac{\sum_{i=1}^{M}\pi_{in}^j\hat{X}_i^j}{\hat{Y}_n}$. In Appendix F, we show how in the one-sector version of our model without royalties, changes in trade costs have no effect on innovation, or on the growth rate, even when we allow for knowledge spillovers.

Substituting equation (21) into the growth rate of the stock of knowledge as in

$$g_A = \sum_{i=1}^{M} \sum_{k=1}^{J} \frac{\varepsilon_{ni}^{jk}}{g_A + \varepsilon_{ni}^{jk}} \lambda_i^k \frac{\hat{A}_i^k}{\hat{A}_n^j} \left(s_i^k\right)^{\beta_r},$$

we obtain

$$g_A = \sum_{i=1}^{M} \sum_{k=1}^{J} \frac{\varepsilon_{ni}^{jk}}{g_A + \varepsilon_{ni}^{jk}} \lambda_i^k \frac{\hat{A}_i^k}{\hat{A}_n^j} \left(\frac{1}{\rho - g_y + g_A} \beta_r \lambda_i^k \frac{1}{(1+\theta)} \frac{\sum_{n=1}^{M} \pi_{ni}^k \hat{X}_n^k}{\hat{Y}_n} \right)^{\frac{\beta_r}{1-\beta_r}}.$$

Rearranging the above equation, we obtain an expression for the growth rate of the stock of knowledge on the BGP,

$$g_A \hat{A}_n^j = \sum_{i=1}^{M} \sum_{k=1}^{J} \frac{\varepsilon_{ni}^{jk}}{g_A + \varepsilon_{ni}^{jk}} (\lambda_i^k)^{\frac{1}{1-\beta_r}} \hat{A}_i^k \left(\frac{1}{\rho - g_y + g_A} \beta_r \frac{1}{(1+\theta)} \frac{\sum_{n=1}^{M} \pi_{ni}^k \hat{X}_n^k}{\hat{Y}_n} \right)^{\frac{\beta_r}{1-\beta_r}}. \tag{22}$$

The growth rate of the stock of knowledge on the BGP depends positively on the speed of diffusion, the expected profits, and the efficiency of innovation, and it depends negatively on the dispersion parameter. Following Eaton and Kortum (1999), the Frobenius theorem guarantees that there is a unique balanced-growth path in which all countries and sectors grow at the same rate g_A. The expression for the growth rate can be expressed in matrix form as

$$g_A A = \Delta(g_A) A.$$

If the matrix $\Delta(g_A)$ is definite positive, then there exists a unique positive balanced-growth rate of technology $g_A > 0$, given research intensities and diffusion parameters. Associated with that growth rate is a vector A (defined up to a scalar multiple), with every element positive, which reflects each country-sector's relative level of knowledge along that balanced-growth path.

In Appendix B, we report the equations of the model after normalizing the endogenous variables.

4 Welfare Gains from Trade

We compute welfare gains from trade after a trade liberalization between the baseline and the counterfactual BGP. Welfare in our model is defined in equivalent units of consumption. We can use equation (1) to obtain the lifetime utility in the initial BGP as

$$\bar{U}_i^* = \int_{t=0}^{\infty} e^{-\rho t} \frac{\left(\hat{C}_i^*\right)^{1-\gamma}}{1-\gamma} e^{g^*(1-\gamma)t} dt = \frac{\left(\hat{C}_i^*\right)^{1-\gamma}}{\rho - g^*(1-\gamma)},$$

and in the counterfactual BGP as

$$\bar{U}_i^{**} = \int_{t=0}^{\infty} e^{-\rho t} \frac{\left(\hat{C}_i^{**}\right)^{1-\gamma}}{1-\gamma} e^{g^{**}(1-\gamma)t} dt = \frac{\left(\hat{C}_i^{**}\right)^{1-\gamma}}{\rho - g^{**}(1-\gamma)}$$

with $*$ denoting the baseline BGP and $**$ denoting the counterfactual BGP.

Welfare gains are defined as the amount of consumption that the consumer is willing to give up in the counterfactual BGP to remain at the same level as in the initial BGP. We call this, λ_i, which is obtained as

$$\bar{U}_i^*(\lambda_i) = \bar{U}_i^{**}$$

$$\frac{\left(\hat{C}_i^* \lambda_i\right)^{1-\gamma}}{\rho - g^*(1-\gamma)} = \frac{\left(\hat{C}_i^{**}\right)^{1-\gamma}}{\rho - g^{**}(1-\gamma)}.$$

From here,

$$\lambda_i = \frac{\hat{C}_i^{**}}{\hat{C}_i^*} \left(\frac{\rho - g^*(1-\gamma)}{\rho - g^{**}(1-\gamma)}\right)^{\frac{1}{1-\gamma}}. \tag{23}$$

Welfare gains depend on changes in normalized consumption between the BGPs and the change in growth rates. From equation (67), normalized consumption in the BGP is equal to income per capita net of R&D expenditures. That is,

$$\hat{C}_i = \hat{Y}_i - \sum_{k=1}^{J} s_i^k \hat{Y}_i = \left(1 - \sum_{k=1}^{J} s_i^k\right) \hat{Y}_i. \tag{24}$$

In static models or one-sector models of trade and innovation in which changes in trade costs do not have an effect on innovation, $g^* = g^{**}$ and $s_i^k = 0$. In that case, welfare gains from trade are computed as changes in the real wage. As in Caliendo and Parro (2015), we can obtain an expression for the real wage in country i as

$$\frac{W_i}{P_i} \propto \prod_{j=1}^{M} \left(\frac{W_i}{P_i^j}\right)^{\alpha^j}.$$

Using the first-order conditions for prices and import shares, it can be shown that

$$\frac{W_i}{P_i^j} = \left(\frac{T_i^j}{\pi_{ii}^j}\right)^{1/\theta} \frac{W_i}{c_i^j} \propto \left(\frac{T_i^j}{\pi_{ii}^j}\right)^{1/\theta} \prod_{k=1}^{J} \left(\frac{W_i}{P_i^k}\right)^{\gamma^{jk}}.$$

Therefore,

$$\frac{W_i}{P_i} \propto \prod_{j=1}^{J} \left(\left(\frac{T_i^j}{\pi_{ii}^j}\right)^{\alpha^j/\theta} \prod_{k=1}^{J} \left(\frac{W_i}{P_i^k}\right)^{\alpha_i^j \gamma^{jk}} \right). \tag{25}$$

Note that this formula resembles the standard welfare formula in Arkolakis, Costinot, and Rodríguez-Clare (2012). In a one-sector version of our model, in which $j = 1$, γ^{jk}=0, and $\alpha^j = 1$, equation (25) becomes

$$\frac{W_i}{P_i} \propto \left(\frac{T_i}{\pi_{ii}}\right)^{1/\theta}. \tag{26}$$

This is the standard formula for welfare gains from trade that has been used in the literature and depends on aggregate productivity, the home trade shares and the trade elasticity.

5 Quantitative Analysis

We quantify our model to evaluate the role that sector heterogeneity in production and knowledge flows has on innovation and welfare. We study the effect of a trade liberalization that consists of a uniform reduction of trade barriers of 40%. We compare the economy in the baseline and counterfactual BGPs. We consider four versions of our model: (i) our baseline model with heterogeneity in innovation, production and knowledge linkages; (ii) a model with sector heterogeneity but where diffusion is almost negligible; and (iii) a one-sector model with knowledge flows across countries. In all cases, we recalibrate the parameters of the model to match the same moments of the data.

5.1 Calibration

We use data on bilateral trade flows, R&D intensity, production, and patent citations to calibrate the main parameters of the model. We assume that the world is on a BGP in 2005. We calibrate the model in two stages. In the first stage, we calibrate the production and knowledge diffusion parameters, as well as the average productivity T_i^j and trade barriers d_{in}^j, and solve for the competitive equilibrium of the model. In the second stage, we take as given the results from the competitive equilibrium and solve for the innovation parameters and the stock of knowledge. Here we explain in more detail the calibration of the average productivity parameters T_i^j, the diffusion parameters ε_{in}^{jk}, and the parameters governing the innovation process—the elasticity of innovation, β_r, and the efficiency of innovation, λ_i^j. Details on the data used in the calibration are relegated to Appendix B, and the description of the calibration procedure to recover other parameters of interest is provided in Appendix C.

5.1.1 Estimation of T_i^j: Gravity Equation at the Sector Level

To estimate the technology parameters for tradable sectors, $j \leq J-1$, we follow the procedure in Levchenko and Zhang (2016) by estimating standard gravity equations for each sector in 2005. We start from the trade shares in equation (10):

$$\pi_{ni}^j = \frac{X_{ni}^j}{X_n^j} = \frac{T_i^j \left(c_i^j d_{ni}^j\right)^{-\theta}}{\Phi_n^j}. \tag{27}$$

Dividing the trade shares by their domestic counterpart as in Eaton and Kortum (2002) and assuming $d_{nn}^j = 1$, we have

$$\frac{\pi_{ni}^j}{\pi_{nn}^j} = \frac{X_{ni}^j}{X_{nn}^j} = \frac{T_i^j \left(c_i^j d_{ni}^j\right)^{-\theta}}{T_n^j \left(c_n^j\right)^{-\theta}}. \tag{28}$$

Taking logs of both sides, we have

$$\log\left(\frac{X_{ni}^j}{X_{nn}^j}\right) = \log\left(T_i^j \left(c_i^j\right)^{-\theta}\right) - \log\left(T_n^j \left(c_n^j\right)^{-\theta}\right) - \theta \log(d_{ni}^j). \tag{29}$$

The log of the trade costs can be expressed as

$$\log(d_{ni}^j) = D_{ni,k}^j + B_{ni}^j + CU_{ni}^j + RTA_{ni}^j + ex_i^j + \nu_{ni}^j. \tag{30}$$

Following Eaton and Kortum (2002), $D_{ni,k}^j$ is the contribution to trade costs of the distance between country n and i falling into the k^{th} interval (in miles), defined as [0,350], [350, 750], [750, 1500], [1500, 3000], [3000, 6000], [6000, maximum). The other control variables include common border effect, B_{ni}, common currency effect CU_{ni}, and regional trade agreement RTA_{ni}, between country n and country i. We include an exporter fixed effect, ex_i^j, to fit the patterns in both country incomes and observed price levels as shown in Waugh (2010). ν_{ni}^j

is the error term.

Substituting (30) back into (29) results in the following gravity equation at the sector level:

$$\log\left(\frac{X_{ni}^j}{X_{nn}^j}\right) = \log\left(T_i^j\left(c_i^j\right)^{-\theta}\right) - \theta ex_i^j - \log\left(T_n^j\left(c_n^j\right)^{-\theta}\right) - \theta(D_{ni,k}^j + B_{ni}^j + CU_{ni}^j + RTA_{ni}^j + \nu_{ni}^j). \tag{31}$$

Define $\hat{F}_i^j = \log\left(T_i^j\left(c_i^j\right)^{-\theta}\right) - \theta ex_i^j$ and $F_n^j = \log\left(T_n^j\left(c_n^j\right)^{-\theta}\right)$. We then estimate the following equation using fixed effects and observables related to trade barriers, taking θ as known:

$$\log\left(\frac{X_{ni}^j}{X_{nn}^j}\right) = \hat{F}_i^j - F_n^j - \theta(D_{ni,k}^j + B_{ni}^j + CU_{ni}^j + RTA_{ni}^j + \nu_{ni}^j). \tag{32}$$

Using the estimates of equation (32), we can back out $\log(d_{ni}^j)$ based on equation (30). To obtain the exporter fixed effect in trade cost, ex_i^j, we use the importer and exporter fixed effects from the Gravity equation (32). That is, $ex_i^j = (F_i^j - \hat{F}_i^j)/\theta$. Figure 1 plots the distance parameters that we obtain from the sectoral gravity equations, d_{in}^j, against the trade share from the data that we use to estimate the gravity equations at the sector level, assuming $\theta = 8.28$.

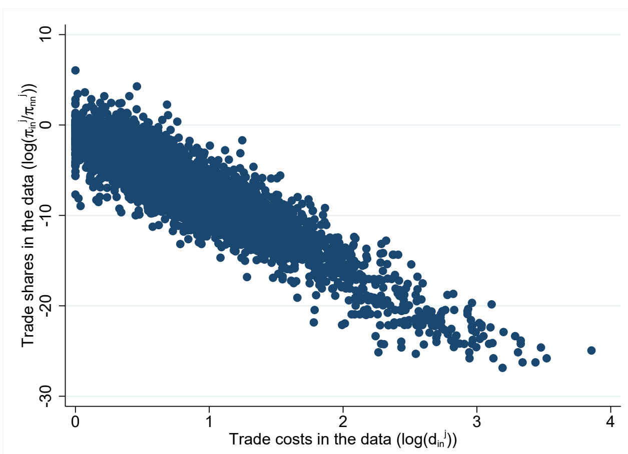


Figure 1: Trade shares and distance

The productivity of the tradable sector in country n relative to that in the United States, T_n^j/T_{US}^j, is then recovered from the estimated importer fixed effects as in

$$S_n^j = \frac{\exp(F_n^j)}{\exp(F_{US}^j)} = \frac{T_n^j}{T_{US}^j}\left(\frac{c_n^j}{c_{US}^j}\right)^{-\theta}, \tag{33}$$

in which the relative cost component can be computed by expressing (4) as

$$\frac{c_n^j}{c_{US}^j} = \left(\frac{W_n}{W_{US}}\right)^{\gamma^j} \prod_{k=1}^{J-1} \left(\frac{P_n^k}{P_{US}^k}\right)^{\gamma^{jk}} \left(\frac{P_n^J}{P_{US}^J}\right)^{\gamma^{jJ}}, \tag{34}$$

where J indicates the nontradable sector. Using data on wages (in USD), estimates of price levels in the tradable sector and the nontradable sector relative to the United States, we can back out the relative cost. The nontradable relative price is obtained using the detailed consumer price data collected by the International Comparison Program (ICP). To compute the relative price of the tradable sector, we follow the approach of Shikher (2012) by combining (8), (9), and (10) and get the following expression for relative prices of tradable goods:

$$\frac{P_n^j}{P_{US}^j} = \left(\frac{X_{nn}^j/X_n^j}{X_{US,US}^j/X_{US}^j}\frac{1}{S_n^j}\right)^{\frac{1}{\theta}}. \tag{35}$$

The right-hand side of this expression can be estimated using the observed expenditure shares of domestic product in country n and in the United States and the estimated importer fixed effects. Substituting the estimates for relative prices and wages in each country-sector and using the estimated S_n^j, we can construct the relative productivity T_n^j/T_U^jS based on equation (33).

To compute the relative productivity in nontradable sectors, we combine (7), (8), and set the trade cost in nontradable sector d_{ni}^J to infinity for all i and n. This implies $\Phi_n^J = T_n^J \left(c_n^J\right)^{-\theta}$ based on equation (7). Substituting this expression into (8), we express the nontradable good price as

$$p_n^J = \frac{c_n^J}{(T_n^J)^{1/\theta}}. \tag{36}$$

The relative technology in nontradable sector can then be constructed based on

$$\frac{T_n^J}{T_{US}^J} = \left(\frac{c_n^J}{c_{US}^J}\frac{P_{US}^J}{P_n^J}\right)^{\theta}. \tag{37}$$

Again, the cost ratios are calculated following (34) and the price ratios for the non-tradable sectors are from the ICP database.

We now have estimated the relative productivity for all countries relative to the United States in every sector. To estimate the level of productivity, we need the U.S. productivity level. First, using OECD industry account data, we estimate the empirical sectoral productivity for each U.S. sector by the Solow residual (without capital in the production function):

$$\ln Z^j_{US} = \ln Y^j_{US} - \gamma^j \ln L^j_{US} - \sum_{k=1}^{J} \gamma^{jk} \ln M^{jk}_{US}, j = 1, 2, ..., J, \tag{38}$$

where Z^j_{US} is measured U.S. productivity in sector j, Y^j_{US} is the output, L^j_{US} is the labor input and M^{jk}_{US} is the intermediate input from sector k. Finicelli et al. (2013) show that trade and competition introduce selection in the productivity level, and the relationship between empirical productivity and the level of technology T^j_{US} in an open economy is given by

$$T^j_{US} = \left(Z^j_{US}\right)^{\theta} \left[1 + \sum_{i \neq US} S^j_i \left(d^j_{US,i}\right)^{-\theta}\right]^{-1}, \tag{39}$$

in which S^j_i and $d^j_{US,i}$ are estimated using (33) and (30) respectively. Lastly, we normalize the U.S. nontradable technology to 1, and express all T^j_{US} relative to T^J_{US} as

$$\hat{T}^j_{US} = \left(\frac{Z^j_{US}}{Z^J_{US}}\right)^{\theta} \left[1 + \sum_{i \neq US} S^j_i \left(d^j_{US,i}\right)^{-\theta}\right]^{-1}. \tag{40}$$

Throughout our analysis we assume that θ is common across countries and set it equal to 8.28.[7]

5.1.2 The Speed of Knowledge Diffusion

We discipline the speed of knowledge diffusion, ε^{jk}_{ni}, using citation data across countries and sectors obtained from the U.S. Patent and Trade Office (USPTO) for the period 2000-2010. In the innovation literature, citation data have been used to trace the direction and intensity of knowledge flows between economic units (such as firms or countries) and across technological classes.[8] In the dataset, each patent is assigned to one of the IPC (International patent classification) categories. We use the probability mapping between IPC and ISIC Rev. 3 provided by www.wipo.int to assign patents into our 19 sectors.[9]

[7] We have also run our gravity equation at the sector level using $\theta = 4$ and a sector specific θ from Caliendo and Parro (2015). We find that the technology parameters estimated under different θ are highly correlated, as 00 documented in Levchenko and Zhang (2016). In particular, the calibration of technology parameters for $\theta = 4$ and $\theta = 8.28$ is 0.98, whereas the correlation of the technology parameter when θ is common and when we use the θ from Caliendo and Parro (2015) is 0.8.

[8] Although patent statistics have been widely used in studies of firm innovations, not all innovations are patented, especially process innovations, which are often protected in other ways such as copyright, trademarks and secrecy (see Levin et al.,1987). Our measure implicitly assumes that for any sector, the unpatented and patented knowledge utilizes knowledge (patented or unpatented) from other sectors in the same manner with the same likelihood and intensity.

[9] Details of the concordance are available at http://www.wipo.int/meetings/en/doc_details.jsp?doc_id=117672.

First, patents applied in year t by sector j in country n are all grouped into a patent bin (nj, t). Based on our model assumption of the exponential distribution of citation lags, we can express the share of total citations from country n sector j made in year t to patents applied in year s by country i sector k as

$$\widehat{citeshare}_{ni,s}^{jk,t} \equiv \frac{citation_{ni,s}^{jk,t}}{\sum_{s=0}^{t} \sum_{ik} citation_{ni,s}^{jk,t}} = \frac{\varepsilon_{ni}^{jk} e^{-\varepsilon_{ni}^{jk}(t-s)} P_{i,s}^{k}}{S_{n}^{j,t}}, \tag{41}$$

where $P_{i,s}^{k}$ is the total number of patent applications in bin (ik, s), $S_{n}^{j,t} = \sum\limits_{s=0}^{t} \sum_{i,k} \varepsilon_{ni}^{jk} e^{-\varepsilon_{ni}^{jk}(t-s)} P_{i,s}^{k}$ is the knowledge stock available to country n sector j at time t that was ever invented at $s \leq t$, and $citation_{ni,s}^{jk,t}$ is the number of citations from patents in nj, t to patents ik, s. $P_{i,s}^{k}$, similar to the term $\lambda_{is}^{k} \left(s_{is}^{k}\right)^{\beta_r}$ in equation (12), measures the new knowledge generated in ik at time s, and $\varepsilon_{ni}^{jk} e^{-\varepsilon_{ni}^{jk}(t-s)}$ is the share of $P_{i,s}^{k}$ that arrives in nj at time t.

In addition, we assume that $S_{n}^{j,t}$ is a stock variable that grows at a constant rate g_{n}^{j}, i.e. $S_{n}^{j,t} = S_{n}^{j,0} e^{t g_{n}^{j}}$. Suppose (t_0, s_0) is the base year-pair. Dividing both sides of equation (41) by the base year-pair observation leads to

$$\frac{\widehat{citeshare}_{ni,s}^{jk,t}}{\widehat{citeshare}_{ni,s_0}^{jk,t_0}} = \frac{\varepsilon_{ni}^{jk} e^{-\varepsilon_{ni}^{jk}(t-s)} P_{i,s}^{k}}{S_{n}^{j,0} e^{t g_{n}^{j}}} \Big/ \frac{\varepsilon_{ni}^{jk} e^{-\varepsilon_{ni}^{jk}(t_0-s_0)} P_{i,s_0}^{k}}{S_{n}^{j,0} e^{t_0 g_{n}^{j}}} = \frac{P_{i,s}^{k}}{P_{i,s_0}^{k}} e^{-\varepsilon_{ni}^{jk}[(t-s)-(t_0-s_0)]} e^{(t_0-t) g_{n}^{j}}. \tag{42}$$

Parameters $\{\varepsilon_{ni}^{jk}, g_{n}^{j}\}$ are then estimated using general methods of moments (GMM) by the quadratic distance between the empirical counterpart of the citation share and equation (42). For each (nj, ik) country-sector pair over T periods, we have $T(T-1)/2$ observations of (t, s) year-pairs and 2 unknowns.

Figure 2 shows that the distribution of diffusion speed across countries and sectors is highly heterogeneous and skewed.

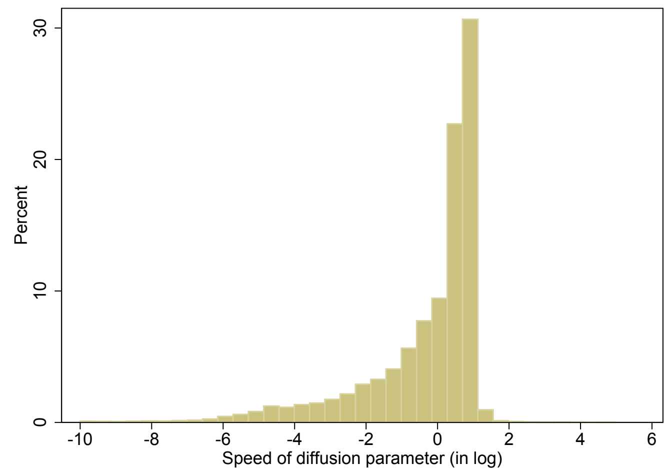


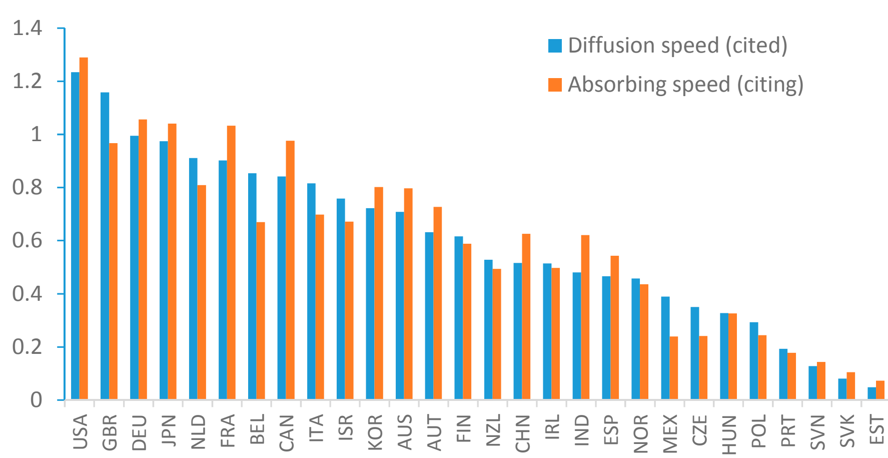
Figure 2: Distribution graph of ϵ_{ni}^{jk}

In addition, Table 1 reports the average speed of diffusion by cited sector (i.e., a sector that diffuses knowledge) and citing sector (i.e, a sector that acquires knowledge). It shows that patents in the chemicals, computer, electronic and medical instruments sectors have the highest diffusion speed, while patents in the wood products sector have the lowest diffusion speed. Across sectors, the citing speed (or speed of absorption) is highly correlated with the cited speed. Figure 3 shows the average speed of diffusion and absorption by country. Unsurprisingly, new knowledge created in the United States, United Kingdom, Germany and Japan diffuse the fastest. The speed of diffusion of knowledge in the United States and United Kingdom on average is less than a year (captured by $1/\epsilon$). Countries that diffuse knowledge (get cited) rapidly also tend to acquire new knowledge from other countries (citing others) fast. Canada, France, and emerging innovation powerhouse like China and India are faster at acquiring new knowledge than diffusing their own knowledge. In Appendix D we present further characteristics of our estimated cross-country cross-sector speed of knowledge diffusion ϵ_{ni}^{jk}. As demonstrated by the estimated gravity-like equation, we find that factors that affect communication and the exchange of ideas, such as linguistic distance, a common border, a common colonizer or a currency union/free trade agreement zone, do have a significant effect on the speed of diffusion. Interestingly, trade also turns out to have a significant positive impact on the citation speed, in line with the literature that argues that trade often serves as a vehicle for knowledge exchange (Alvarez, Buera, and Lucas Jr (2013)) .

Table 1: Average diffusion speed by sectors

ISIC	Industry	Cited	Citing
C24	Chemicals and chemical products	0.948	0.853
C30T33X	Computer, electronic and medical instruments	0.939	0.931
C01T05	Agriculture, hunting, forestry and fishing	0.932	0.912
C17T19	Textiles, textile products, leather and footwear	0.771	0.888
C29	Machinery and equipment, n.e.c.	0.752	0.850
C10T14	Mining and Quarrying	0.793	0.747
C28	Fabricated metal products, except machinery and equipment	0.667	0.736
C40T95	Nontradables	0.630	0.633
C21T22	Pulp, paper, paper products, printing and publishing	0.619	0.595
C15T16	Food products, beverages and tobacco	0.628	0.611
C25	Rubber and plastics products	0.570	0.537
C27	Basic metals	0.594	0.581
C23	Coke, refined petroleum products and nuclear fuel	0.478	0.506
C34	Motor vehicles, trailers and semi-trailers	0.406	0.377
C26	Other non-metallic mineral products	0.445	0.471
C31	Electrical machinery and apparatus, n.e.c.	0.473	0.467
C36T37	Manufacturing n.e.c. and recycling	0.336	0.319
C35	Other transport equipment	0.310	0.310
C20	Wood and products of wood and cork	0.175	0.144

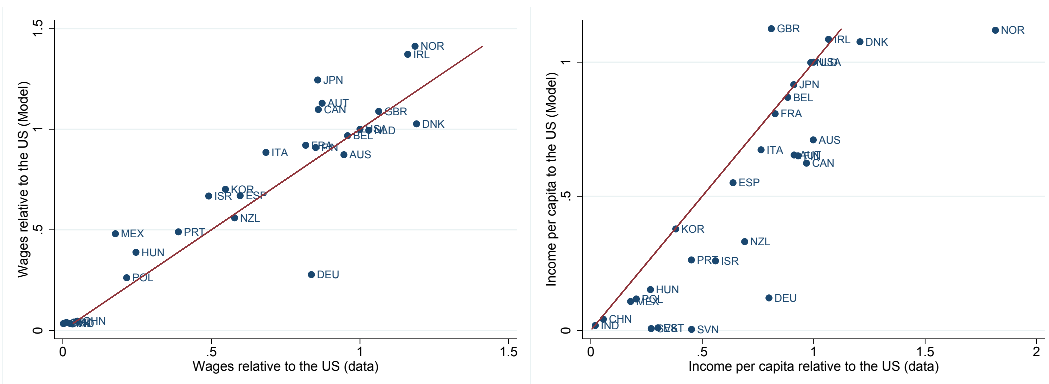
Figure 3: Average speed of diffusion by country



Note: This figure presents the average diffusion speed and absorbing speed by country. Average diffusion (absorbing) speed is calculated as the average ϵ by cited (citing) country.

5.1.3 Parameters of Innovation

We calibrate the parameters of innovation $\{\beta_r, \lambda_n^j, \hat{A}_n^j\}$ in two steps. First, we solve for the static trade equilibrium taking as given the estimated sectoral productivity T_i^j, the estimated trade barriers d_{in}^j, and production input-output linkages parameters $\{\alpha^j, \gamma^j, \gamma^{jk}\}$ estimated using the U.S. input-output table for 2005. Our calibration strategy delivers relative wages and relative income per capita that are broadly consistent with those observed in the data (see Figure 4).

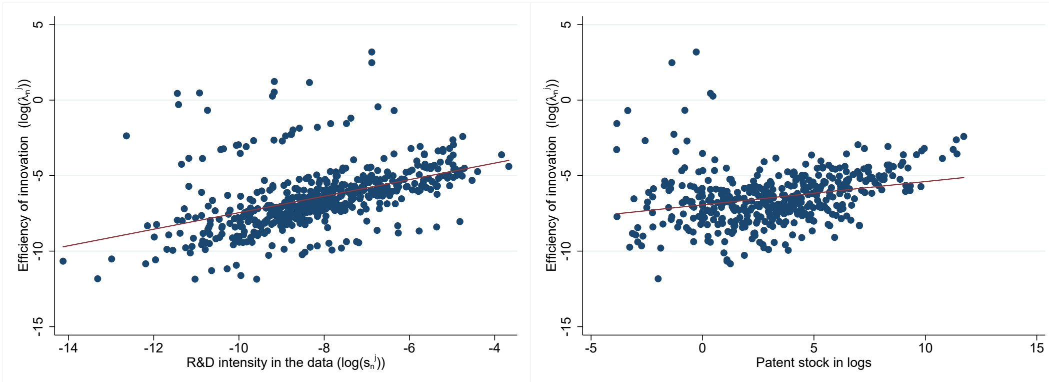


(a) Wages relative to the US

(b) Income per capita

Figure 4: Validation: Relative wages and relative income per capita

Having computed wages and trade shares, in the second step we use the estimated parameters for knowledge diffusion, ε_{in}^{kj}, data on R&D intensity at the country-sector level, s_n^j, and the expression for the growth rate of the economy on the BGP in equation (22) to calibrate the innovation parameters $\lambda_n^j, \beta_r, A_n^j$. We proceed as follows: First, we assume a growth of income per capita (productivity) on the BGP of $g_y = 3\%$. This corresponds to a growth rate for the stock of knowledge on the BGP of $g_A = \theta\left(1 + \sum_{j=1}^{J} \alpha_j \Lambda_j\right)^{-1} g_y = 0.25$. Because all countries and sectors' stock of knowledge grows at the same rate, all countries have the same productivity growth on the BGP (see Appendix E for details on the derivation). Second, we use the Frobenius theorem and equation (22) to obtain a value for the efficiency of innovation, λ_i^k, and the elasticity of innovation, β_r. Given data for s_n^j, the estimated values for ε_{ni}^{jk}, and g_A, we can use the Frobenius theorem and iterate on equation (22) to obtain β_r and λ_n^j. We obtain that $\beta_r = 0.24$ and λ_n^j ranges from $7 * 10^{-6}$ to 24, with mean 0.10 and standard deviation 1.2. Figure 5 plots the calibrated values for λ_i^k against two measures of innovation in the data: R&D intensity (as an input of innovation and a flow variable) and the stock of patents (as an output of innovation and a stock variable). As the figure shows, there is a positive relationship between the efficiency of innovation, λ_i^k, and both measures of innovation.

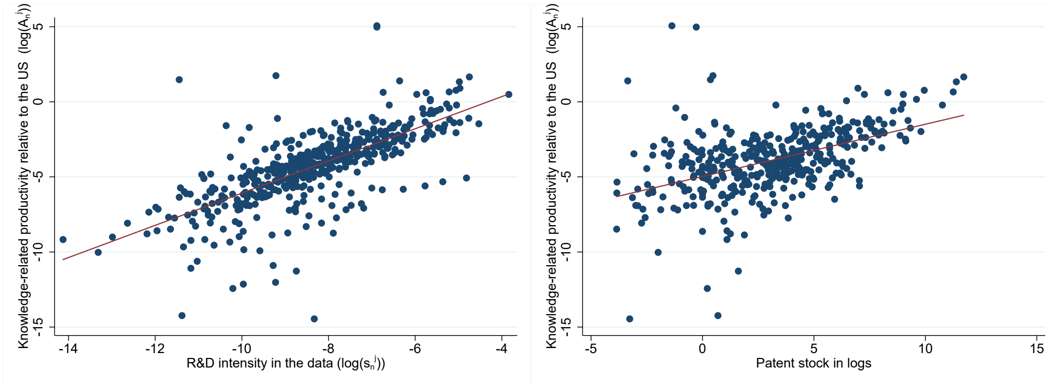


(a) R&D intensity (b) Stock of patents

Figure 5: The productivity of innovation, R&D intensity and the stock of patents

The correlation between R&D intensity, s_n^j, and the parameter in the efficiency of innovation, λ_n^j, is around 0.45. However, the cross-sector correlations between these two variables are heterogeneous across countries. We find that lower-income countries, such as Slovakia, Slovenia and Estonia, do not allocate R&D across sectors according to the exogenous component of the efficiency of innovation, λ_n^j, and have a correlation below 0.2. In contrast, in countries such as United States and Japan or the United Kingdom, the correlation is almost 1. Note that the efficiency of innovation that determines R&D intensity is actually a function of the parameter λ_n^j and the stock of knowledge A_n^j. The stock of knowledge of a country-sector has two main components: (i) knowledge developed in that country-sector through the country's own innovation and (ii) knowledge developed somewhere else that has been diffused to that particular country-sector. In countries with a low correlation between R&D intensity and λ_n^j, R&D intensity is determined by the second component of the stock of knowledge. Diffusion is a key channel for promoting R&D in those countries and sectors. In what follows, we describe how we calibrate A_n^j and analyze its correlation with R&D intensity.

Given these parameter values, and using again the properties of the Frobenius theorem, the associated eigenvector to the growth rate of $g_A = 0.25$ corresponds to the normalized knowledge-related productivity $\hat{A}_n^j$. Figure 6 shows that there is a strong positive relation between the knowledge-related productivity, relative to the United States in sector J, $\hat{A}_i^j$, and both R&D intensity and the stock of patents.



(a) R&D intensity (b) Stock of patents

Figure 6: Research-related productivity and innovation

The correlation between A_n^j and s_n^j is around 0.72. The larger the stock of knowledge that the country-sector has accumulated, the larger the R&D intensity. This relation is heterogeneous across countries and depends on how much has been invested in R&D versus how much knowledge has been accumulated from other countries and sectors through diffusion. The relation also varies by sectors. Figure 7 shows, by industry, the relationship between R&D intensity and the knowledge-related productivity across sectors. We observe a strong positive correlation between the two measures. The correlation is larger for the machinery and equipment, computer, electronic and optical equipment, and electrical machinery sectors.

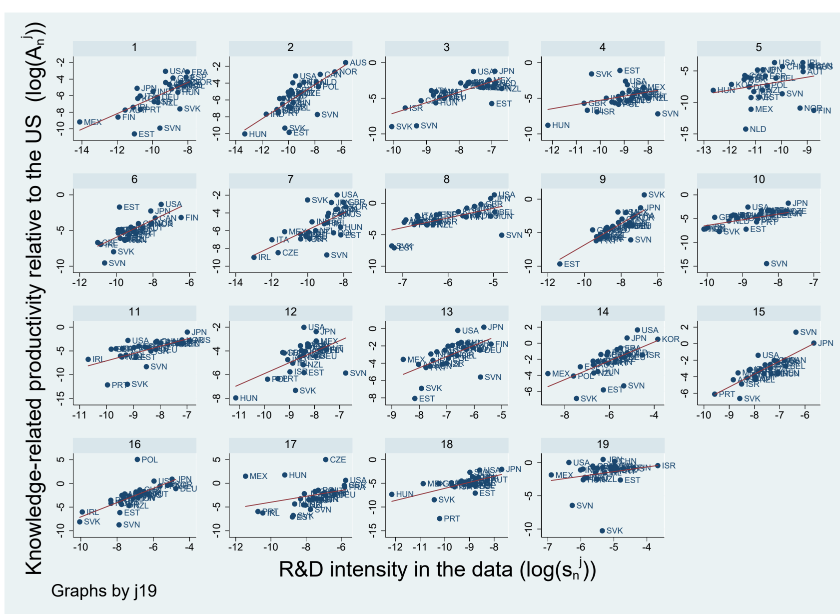


Figure 7: Research-related productivity and innovation by sector

5.1.4 The Algorithm

The calibration of the parameters of innovation, $\{\lambda_n^j, \beta_r, \hat{A}_n^j\}$ follows a recursive algorithm. First, knowing $\{\gamma^j, \gamma^{jk}, \alpha^j, \sigma, \hat{T}_n^j, d_{in}^j\}$, we use the trade structure of the model to obtain wages, prices, expenditures, trade shares, and output, from equations (58)-(67) in Appendix B.

Then, given $\{\varepsilon_{in}^{jk}, g_A, s_n^j\}$, we iterate over equation (22) to obtain $\{\lambda_n^j, \beta_r\}$. We do given guessed values of λ_n^j and β_r, and we use R&D data, s_n^j, and keep iterating until $g_A = 0.25$. We use equation (68), (69) and the Frobenious theorem. The Frobenius theorem guarantees that there is a unique balanced-growth path in which all countries and sectors grow at the same rate g_A. The expression for the growth rate can be expressed in matrix form as

$$g_A A = \Delta(g_A) A.$$

If the matrix $\Delta(g_A)$ is definite positive, then there exists a unique positive balanced-growth rate of technology $g_A > 0$ given research intensities. Associated with that growth rate is a vector A (defined up to a scalar multiple), with every element positive, which reflects each country and sector's relative level of knowledge along that balanced-growth path. We update β_r so that $g_A = 0.25$, and we update λ_n^j so that R&D intensity matches the data. Then we obtain $\hat{A}_n^j$ from the eigenvector associated to $\Delta(g_A = 0.25)$. Knowing $\hat{T}_n^j$ from the gravity regressions and $\hat{A}_n^j$ from the Frobenius theorem, we can obtain $T_{p,n}^j$ from equation (19).

5.2 Counterfactual Analysis

We perform a uniform and permanent reduction of trade barriers, d_{in}^j, of 40% for all country-pairs i, n and sector j. All other parameters are kept fixed at their calibrated values. We analyze the effect of this trade liberalization on innovation and welfare gains from trade across the baseline and the counterfactual BGPs. First, we describe briefly the algorithm that we develop to compute the counterfactual BGP. Different from the calibration algorithm, which could be solved in two stages—first characterizing the competitive equilibrium taking as given $\hat{T}_n^j$ (static equilibrium), and second solving for the innovation and diffusion parameters (dynamic equilibrium)—the algorithm to compute the transition is slightly more involved in that it requires us to solve for the static and dynamic parts of the model simultaneously. After having described the algorithm, we report our main results for our multi-country and multi-sector endogenous growth model featuring heterogeneous interlinkages in production and knowledge flows. First, we characterize welfare gains from trade in the baseline model and describe how changes in sectoral innovation and RCA across counterfactuals help shape gains from trade. Then we explore the role of the main two channels in our model: (i) the presence of knowledge spillovers and (ii) the multi-sector structure of the model.

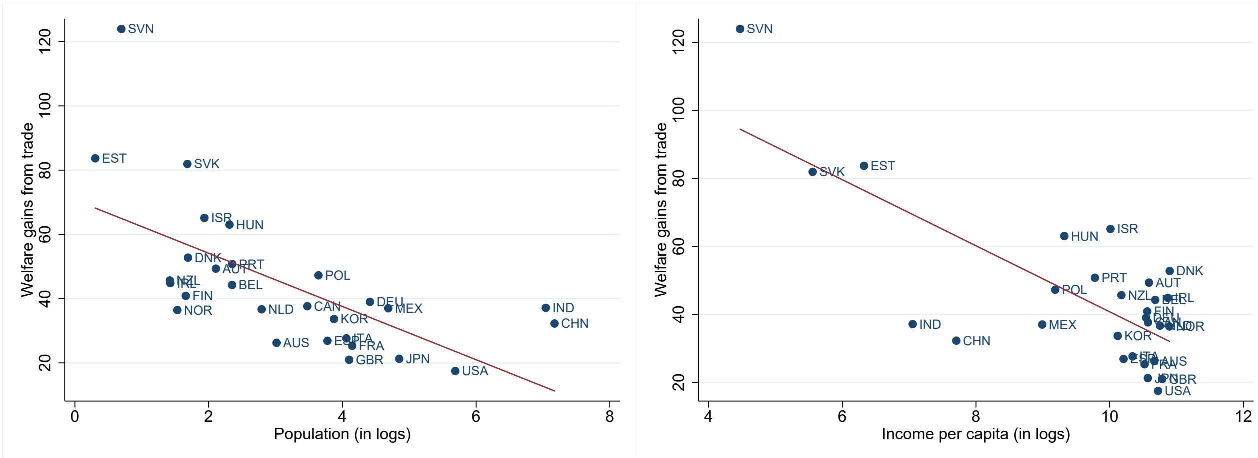
5.2.1 The Algorithm

In our calibration, we took the level of technology, $\hat{T}_n^j$, as given by the estimated values from the gravity regressions. However, when there are changes in trade costs, $\hat{T}_n^j$ will change across counterfactuals to the extent that $\hat{A}_n^j$ also changes. In our model, there are changes in $\hat{A}_n^j$ that are induced by changes in the innovation intensity, s_n^j, and by knowledge diffusion. Our algorithm to solve for the counterfactual equilibrium uses the properties of the Frobenius theorem and allows $\hat{T}_n^j$ to evolve over time through changes in $\hat{A}_n^j$. First, we take $\{\gamma^j, \gamma^{jk}, \alpha^j, \sigma, T_{p,n}^j, \hat{T}_n^j, \beta_r, \lambda_n^j\}$ as given and compute the static equilibrium that corresponds to the new trade barriers, d_{in}^j. With that equilibrium, we compute the new optimal R&D intensity s_n^j and use the Frobenius theorem to obtain the new g_A and associated eigenvector $\hat{A}_n^j$. We do this by iterating over equation (22) until $g_A(t-1) = g_A(t)$. The new $\hat{A}_n^j$ delivers a new $\hat{T}_n^j$ (we keep $T_{n,p}^j$ constant across counterfactuals). We then repeat the procedure until $\hat{T}_n^j$ converges.

5.2.2 Welfare Gains from Trade, Innovation and RCA

We compute welfare gains from trade using equation (23). Welfare gains across two BGPs depend on two components: (1) changes in consumption, $\hat{C}_i$, and (2) changes in the growth rate, g. From equation (24), our model implies that the change in consumption depends on two additional factors: (i) the change in income per capita, $\hat{Y}_i$, and (ii) the change in R&D intensity s_i^k, which has an effect both on the growth rate g and on income per capita $\hat{Y}_i$.

We find that welfare gains from trade are heterogeneous across countries, ranging from 17.5% in the United States to 124% in Slovenia, with a cross-country average gain of 44.6%. The gains are larger for smaller countries and countries with lower income per capita (see Figure 8), which is consistent with the findings in Waugh (2010).



(a) Population

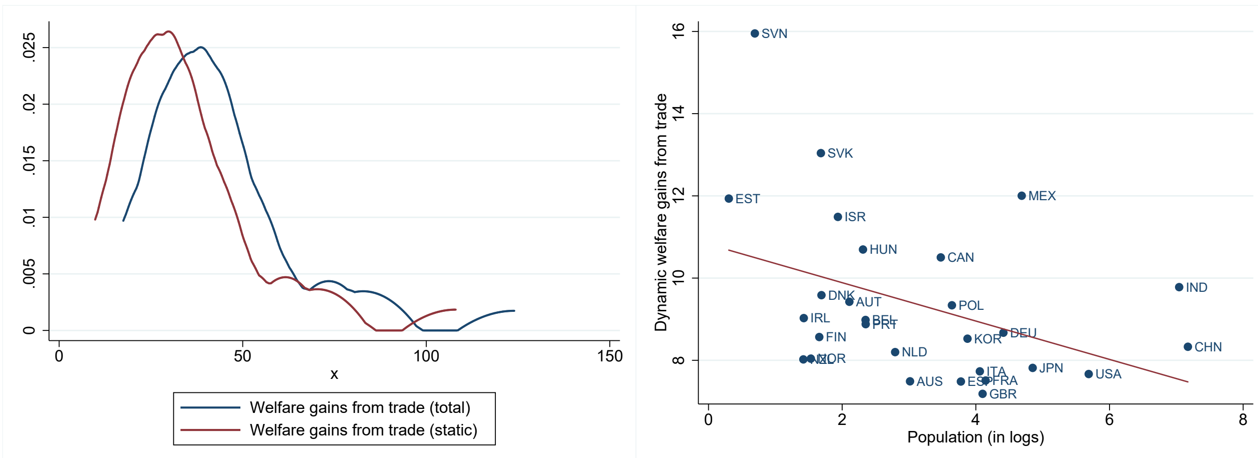
(b) Income per capita

Figure 8: Welfare gains from trade against population and income per capita

Welfare gains from trade can be divided into static and dynamic gains. Static gains correspond to those obtained in a model where the stock of knowledge, $\hat{A}^j_{nt}$, is not allowed to change over time. These are the gains that are obtained in standard static models of trade and are driven by increased specialization and comparative advantage. Dynamic gains take into account the effect of R&D and knowledge spillovers on the stock of knowledge. Both allow the stock of knowledge to increase over time. Higher innovation allows countries to increase their income per capita, which has an unambiguously positive effect on dynamic gains. Knowledge diffusion has two opposite effects on dynamic gains from trade. On the one hand, it increases the stock of knowledge of a country-sector as it can benefit from innovation created in other country-sectors. This has an additional effect on the efficiency of innovation, from equation (11), which reinforces the innovation channel. On the other hand, knowledge spillovers may generate convergence of comparative advantage over time, dampening the total welfare gains from trade that are driven by differences in comparative advantage. We explore this point further in Section 5.2.3.

To compute static welfare gains, we simulate our model keeping $\hat{T}^j_i$ and $\hat{A}^j_i$ constant across counterfactuals. Because we are analyzing only changes across BGPs, dynamic gains do not include the transition. We call them dynamic in that they reflect the gains that account for changes in the stock of knowledge across counterfactuals. Therefore, these gains are computed by letting $\hat{T}^j_i$ and $\hat{A}^j_i$ vary across counterfactuals. We then compare consumption in the initial and counterfactual BGPs.

Figure 9a compares welfare gains from trade in our baseline model to those static gains in which the stock of technology is kept constant across counterfactuals. The difference between the two gains is a measure of dynamic gains from trade (Figure 9b). The cross-country distribution of static gains is shifted to the left, which implies that dynamic gains are positive in every country. After the trade liberalization, the average country experiences dynamic gains from trade that are around 9%, with a minimum of 7.1% in the United Kingdom and a maximum of 15.95% in Slovenia. As in the case of total welfare gains from trade, dynamic gains are also larger for smaller countries (see Figure 9b.)



(a) Welfare gains from trade: Baseline against static

(b) Dynamic welfare gains from trade

Figure 9: Welfare gains from trade against population and income per capita

Dynamic gains are driven by innovation and knowledge diffusion, which generates increases in the growth rate and in income per capita. The growth rate increases from 3% in the initial BGP to 3.1% in the counterfactual BGP. Here we analyze the direct role of innovation on dynamic gains, and we leave the analysis on the role of knowledge diffusion to Section 5.2.3. After a trade liberalization, there is a reallocation of R&D and comparative advantage across sectors within a country. The reallocation translates into changes in innovation at the country level. Except for Finland, New Zealand and Norway, all countries in our sample experience an increase in R&D spending after a trade liberalization. However, the increase in R&D is far from uniform across sectors within a country. In general, R&D increases are correlated with increases in revealed comparative advantage (RCA) at the sector level. We find that this correlation is stronger for those countries that experience larger gains from trade (see Figure 10).

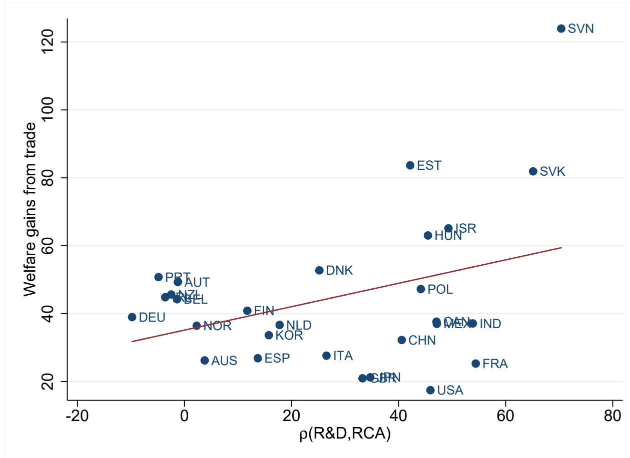
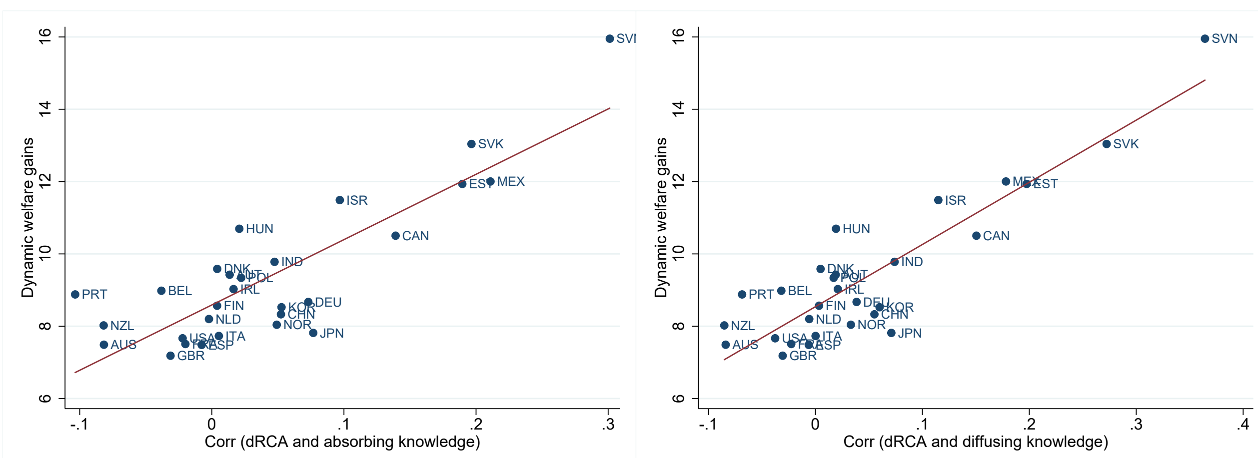


Figure 10: Welfare gains from trade and correlation of R&D and RCA

Note: In the X-axis, $\rho(R\&D, RCA)$ is the correlation of the change in R&D intensity and the change in RCA across country-sectors.

We also find that RCA tends to increase more in those sector that have faster knowledge diffusion, both absorbing (citing) and diffusing (cited) knowledge. This correlation is larger in countries that experience larger dynamic gains from trade (see Figure 11). Furthermore, knowledge diffusion allows countries and sectors to access a larger pool of ideas, increasing their efficiency of innovation.

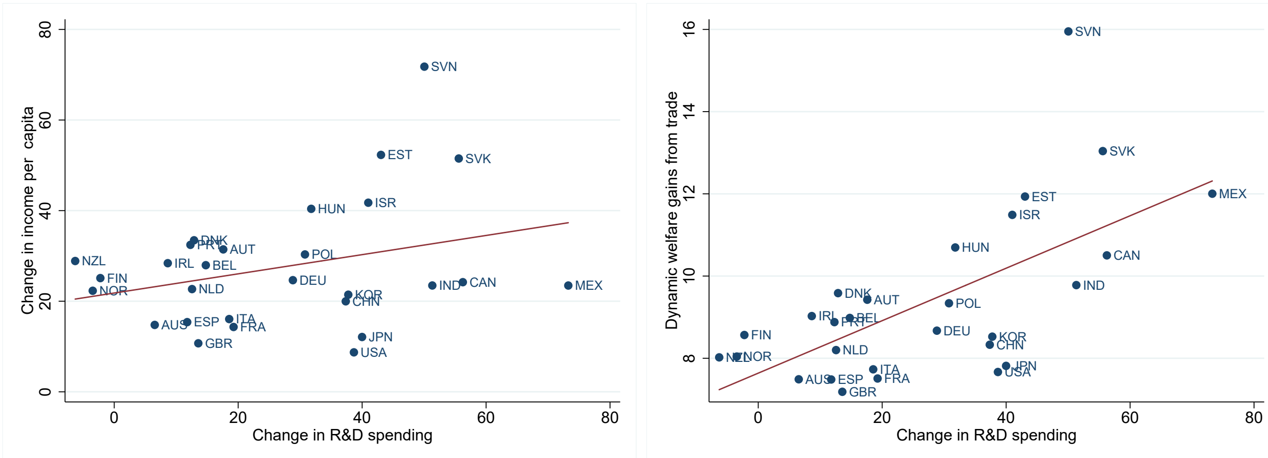


(a) Population

(b) Income per capita

Figure 11: Dynamic welfare gains and the correlation of RCA and knowledge flows

Because of innovation and knowledge diffusion, income per capita increases after a trade liberalization and, from equation (24), consumption and welfare also increase. Countries with larger increases in R&D spending also experience larger increases in income per capita (see Figure 12a) and larger dynamic welfare gains from trade (see Figure 12b).



(a) Income per capita

(b) Dynamic welfare gains from trade

Figure 12: The effect of changes in innovation on dynamic welfare gains

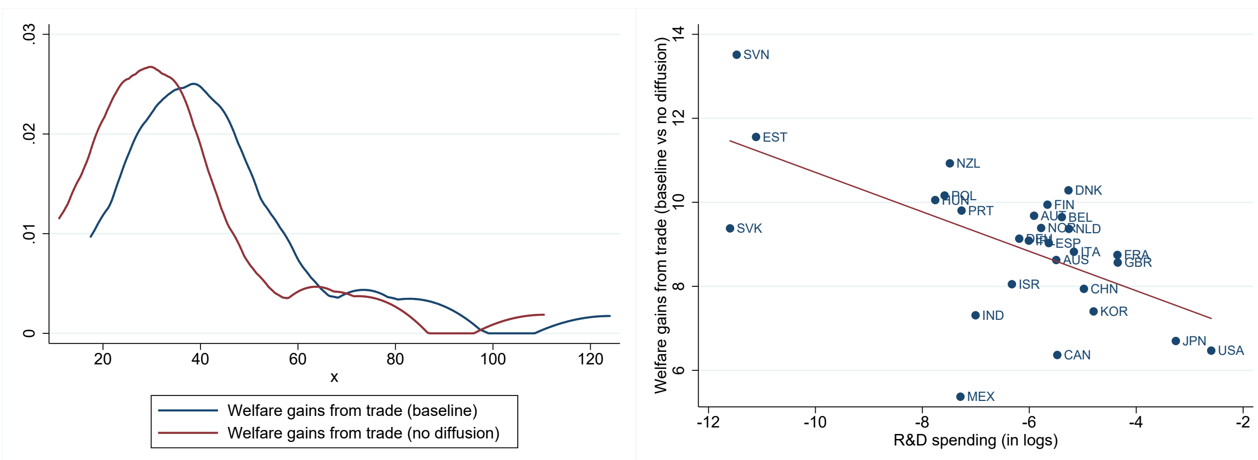
5.2.3 Welfare Gains from Trade: The Role of Knowledge Spillovers

We study the role of knowledge diffusion on welfare gains from trade. To do that, we recalibrate our baseline model by setting the diffusion parameters ε_{ni}^{jk} to a very small value of 0.0001, for all $i \neq n$ and $k \neq j$ (we set $\varepsilon_{nn}^{jj} \rightarrow \infty$; that is, we assume instantaneous diffusion within the same country-sector pair).[10] This recalibration does not affect the first-stage calibration that solved for the competitive equilibrium of the model. However, we need to recalibrate the second-stage parameters, β_r and λ_n^j, by using the same input-output linkage parameters $\{\alpha^j, \gamma^j, \gamma^{jk}\}$, estimated technology, T_n^j, R&D intensity, s_n^j, and growth rate, g_A, values than in the baseline model. We now obtain $\beta_r = 0.28$ and a mean λ_n^j of 0.18 with a standard deviation of 2.2. We then perform the same uniform and permanent trade liberalization of a 40% reduction of trade barriers.

Knowledge diffusion has two opposite effects on welfare. On the one hand, it may lead to productivity convergence across countries and sectors, which would make countries and sectors more similar to each other. This effect would dampen the gains from trade. On the other hand, it provides stronger dynamic gains, as countries and sectors have access to a larger knowledge pool. This effect increases the stock of knowledge of a country-sector, and hence the efficiency of innovation from equation (11).

[10]The Frobenius theorem is only valid if there is at least some diffusion across all country-sector pairs. Setting ε_{ni}^{jk} to a very low number allows us to make use of the properties of the Frobenius theorem while allowing for very slow to virtually no diffusion.

Figure 13 shows that the welfare gains from trade in a model with very low diffusion are smaller than in the baseline model. The direct effect on the stock of knowledge and the effect on the efficiency of innovation that allows countries and sectors to benefit from other country-sector innovations through diffusion dominate the comparative advantage effect by which convergence in technology would reduce the gains from trade as countries become more similar (see Levchenko and Zhang (2016)). The differences in welfare gains between the baseline model and a model with very low diffusion are larger for less-innovative countries, such as Slovenia, Estonia and Slovakia. These are the countries that benefit more from knowledge flows from other countries and less from innovations in their own sectors. In Japan, the United States, and Korea, the differences between the gains are smaller, as these countries are already very innovative and have less to gain from knowledge diffusion from other countries.



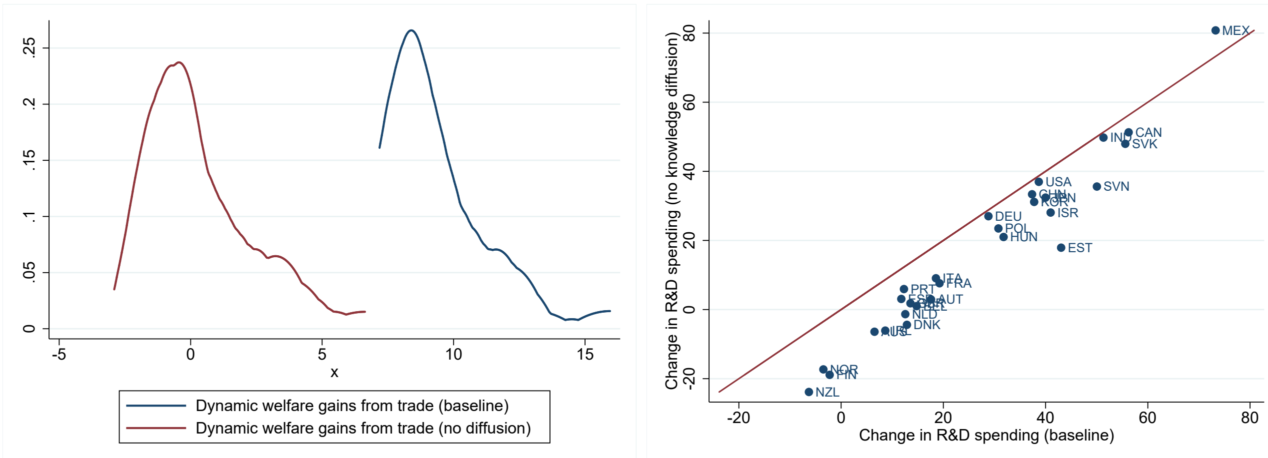
(a) Welfare gains from trade

(b) Welfare gains from trade and R&D

Figure 13: Relative welfare gains from trade (baseline versus no knowledge diffusion)

Static welfare gains from trade are the same in the baseline model and in the model with low knowledge spillovers, as they are computed keeping the level of technology constant across counterfactuals. In both cases, the initial stock of knowledge is the same and is from section 5.1.1. Therefore all the differences in welfare gains in the two cases are driven by differences in dynamic gains (see Figure 14a). Dynamic gains are driven by both innovation and knowledge spillovers. Figure 14b compares, for a cross-section, the change in R&D in the baseline model with the corresponding change in R&D in a model with negligible knowledge spillovers. The figure shows that there are substantial differences in how R&D changes in the two models. All countries, except for Mexico, experience larger increases in R&D in the baseline model. This suggests that the larger dynamic gains from trade in

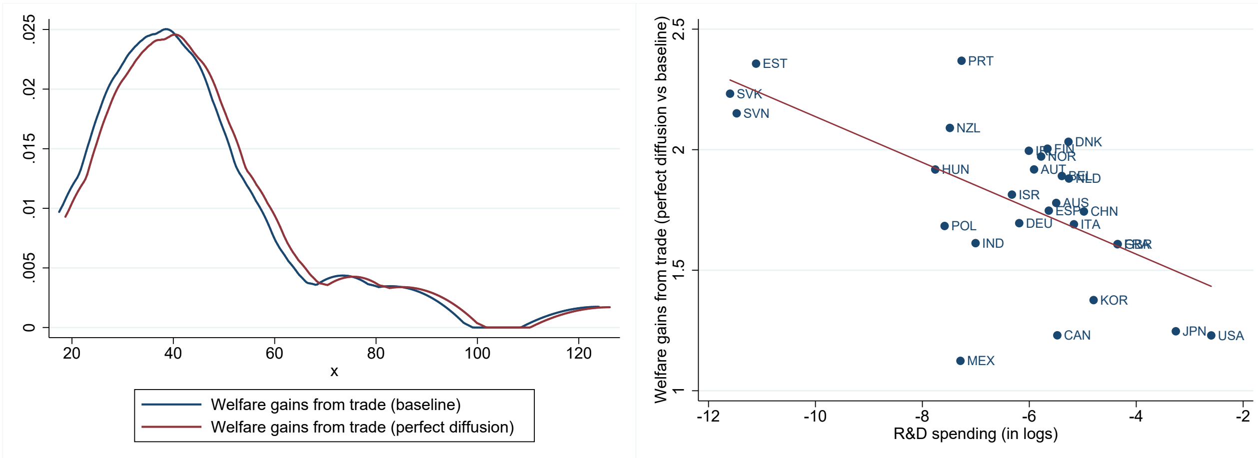
Mexico are entirely driven by knowledge flows and the effect that knowledge diffusion has on the efficiency of innovation.



(a) Dynamic welfare gains from trade (b) Changes in R&D spending

Figure 14: Dynamic welfare gains from trade with negligible knowledge spillovers

We then look at the extreme case of instantaneous (perfect) diffusion across all sectors and countries. In this exercise we set $\varepsilon_{ni}^{jk} \to \infty$ for all $i \neq n$ and $k \neq j$. All other parameters are kept fixed as in the baseline calibration. We then recalibrate the parameters of innovation and obtain $\beta_r = 0.19$ and λ_n^j with mean 0.0015 and standard deviation 0.0045. We find that the welfare gains are larger than in the baseline model, but the differences are small. We also find that these differences are much smaller than those between the baseline model and the model with negligible knowledge diffusion. This finding suggests that the world is closer to almost perfect diffusion than to no diffusion, as changes in trade frictions have lower effects on welfare gains when there is perfect diffusion (see Figure 15). Similarly to the case of very low diffusion, the difference in gains between the baseline and the model with perfect diffusion are larger in countries with lower total R&D spending, that is, countries in which the stock of knowledge is lower.



(a) Welfare gains from trade

(b) Welfare gains from trade (cross-section)

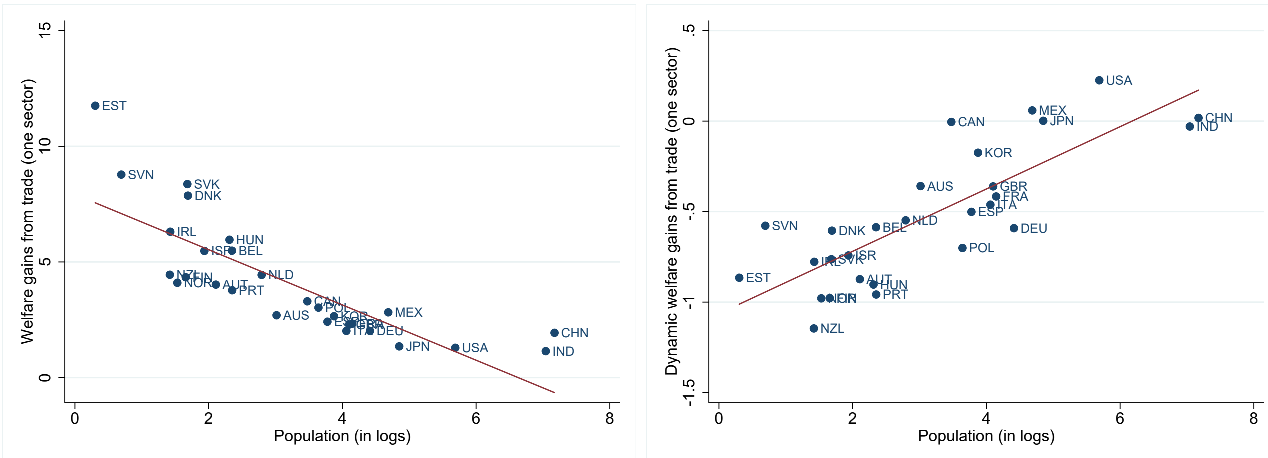
Figure 15: Welfare gains from trade with negligible knowledge spillovers

5.2.4 Welfare Gains from Trade: The Role of Multiple Sectors

In our baseline model, we have emphasized the importance of the role of multiple sectors in reallocating R&D intensity and RCA across sectors. This reallocation was important to understand the role of innovation and knowledge diffusion in generating dynamic gains from trade. We now recalibrate our baseline model to a one-sector model in which there are no production and knowledge diffusion interlinkages across sectors.

We re-estimate the technology parameters, $\hat{T}_n$ by running gravity equations at the country level. Note that now there is no a sector j dimension in the model. The production and knowledge linkages parameters are also recalibrated at the country level. We set $\alpha^j = 1$, $\gamma^j = 1$ and $\gamma^{jk} = 0$ for all j and k. We obtain country-level data for R&D intensity, s_n. Then, assuming the same g_A as in the baseline model, we obtain a $\beta_r = 0.28$ and λ_n with mean 0.22 and standard deviation 0.13. We find that this model delivers substantially lower gains from trade than our multi-sector growth model with production and knowledge interlinkages.

In the multi-sector version of the model, R&D reallocates towards sectors that experience larger increases in comparative advantage. This reallocation channel is no longer present in a one-sector model, and R&D increases are much smaller. As in the baseline model, welfare gains from trade are larger for smaller countries (see Figure 16a). Dynamic gains, however, are negative for most countries and positive only for larger countries (see Figure 16b). In a multi-sector model, small countries can specialize in a few sectors in which they concentrate their R&D. Their stock of knowledge in those sectors is larger than in other sectors, and that has an important effect on innovation at the country level and welfare gains. In the one-sector model, however, there is no RCA or specialization and no reallocation of R&D across sectors. In this case, large countries have a larger stock of knowledge than small countries and hence do even more R&D than smaller countries. Furthermore, in our one-sector model with knowledge diffusion, dynamic gains are driven by the existence of royalty payments, which flow to larger countries with a larger stock of knowledge. Knowledge diffusion with royalties generates an additional source of gains from trade for larger countries.

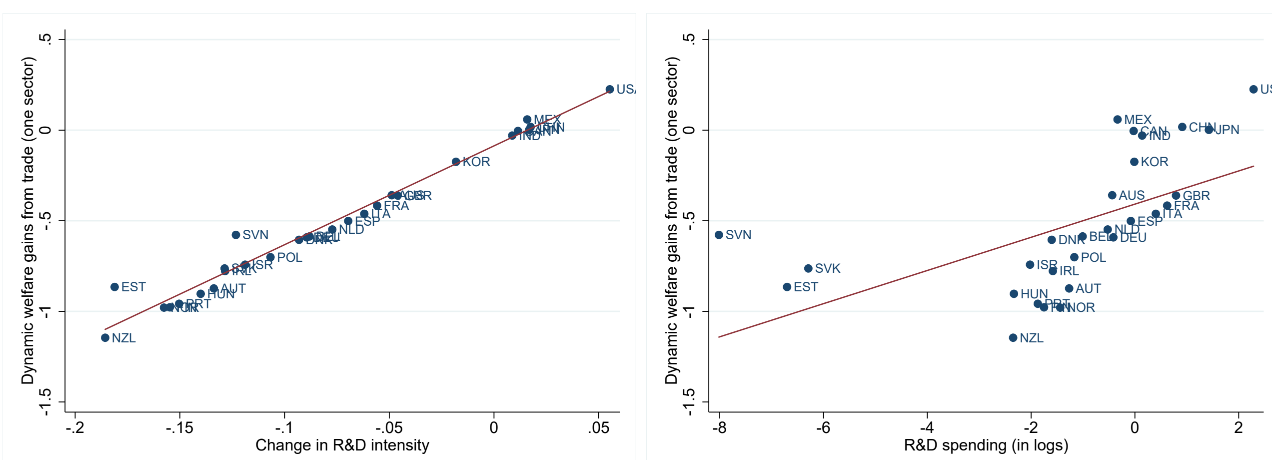


(a) Total gains

(b) Dynamic gains

Figure 16: Welfare gains from trade in a one-sector model

Figure 17a shows that more-innovative countries are the ones who experience larger dynamic welfare gains from trade in the one-sector model. These are the countries with larger increases in R&D intensity after a trade liberalization. Negative dynamic gains are driven by decreases in innovation after a trade liberalization in small countries.



(a) Change in R&D intensity

(b) R&D spending

Figure 17: Welfare gains from trade in a one-sector model

In our baseline model, changes in trade costs do generate changes in R&D intensity even when we consider a one-sector model, because of the presence of knowledge diffusion and royalties. These effects are, however, much smaller than in the multi-sector model. Standard one-sector models of trade and innovation in which trade has no effect on innovation do not feature royalty payments (see Eaton and Kortum (1996, 1999) , Buera and Oberfield (2016), Atkeson and Burstein (2010)). In Appendix F we show that in a one-sector model without royalties, changes in trade barriers have no effect on R&D or the growth rate of the economy, hence no effects on dynamic welfare gains from trade.

6 Concluding Remarks

We develop a quantitative framework to study the effect of interlinkages among trade, knowledge flows and production on innovation and welfare. We distinguish between static gains from trade, driven by increased specialization, and dynamic gains from trade, driven by innovation and knowledge diffusion. Changes in trade barriers have a quantitatively important effect on innovation and welfare. Knowledge diffusion amplifies the effect, as sectors in a country benefit from a larger pool of ideas, increasing dynamic welfare gains. A one-sector version of our model delivers negligible, or even negative in some countries, dynamic welfare gains. This result reinforces the importance of modeling sectoral production and knowledge interlinkages when studying the effect of trade liberalizations on innovation and welfare.

In order to expose the role of knowledge spillovers and multiple sectors, we analyzed a reduction in trade frictions that is uniform across sectors and countries. Our quantitative framework can also be used to analyze changes in trade policy of a particular sector in a country. For instance, we could study the effect of the accession of China to the WTO in 2001 that involved reduction in import quotas on apparel and textiles had on innovation and welfare in several sectors of the United States.

Our model can be extended to study other important issues in macroeconomics and international trade. If the production structure of the economy is CES rather than Cobb-Douglas, a trade liberalization will shift production shares across sectors, hence inducing structural change. Moreover, we assume perfect enforcement of royalty payments. If, instead, international royalty payments are partially enforced, we could extend the model to include intellectual property rights and study their welfare implications. We leave these questions for future research.

References

Alvarez, Fernando E, Francisco J Buera, and Robert E Lucas Jr. 2008. "Models of idea flows." Tech. rep., National Bureau of Economic Research.

———. 2013. "Idea flows, economic growth, and trade." Tech. rep., National Bureau of Economic Research.

Arkolakis, Costas, Arnaud Costinot, and Andrés Rodríguez-Clare. 2012. "New trade models, same old gains?" *The American Economic Review* 102 (1):94–130.

Arkolakis, Costas, Natalia Ramondo, Andrés Rodríguez-Clare, and Stephen Yeaple. 2013. "Innovation and production in the global economy." Tech. rep., National Bureau of Economic Research.

Atkeson, Andrew and Ariel Tomas Burstein. 2010. "Innovation, Firm Dynamics, and international Trade." *Journal of political economy* 118 (3):433–484.

Bernard, Jonathan Eaton, Bradford Jensen, and Samuel Kortum. 2003. "Plants and Productivity in International Trade." *The American economic review* 93 (4).

Bernard, Andrew B and Charles I Jones. 1996a. "Comparing apples to oranges: productivity convergence and measurement across industries and countries." *The American Economic Review* :1216–1238.

———. 1996b. "Technology and convergence." *The Economic Journal* :1037–1044.

Bottazzi, Laura and Giovanni Peri. 2003. "Innovation and spillovers in regions: Evidence from European patent data." *European economic review* 47 (4):687–710.

Buera, Francisco J and Ezra Oberfield. 2016. "The Global Diffusion of Ideas." Tech. rep., National Bureau of Economic Research.

Cai, Jie and Nan Li. 2016. "The Composition of Knowledge and Cross-Country Income Differences." .

———. 2017. "Growth through intersectoral knowledge linkages." .

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

Cameron, Gavin, James Proudman, and Stephen Redding. 2005. "Technological convergence, R&D, trade and productivity growth." *European Economic Review* 49 (3):775–807.

Chor, Davin. 2010. “Unpacking sources of comparative advantage: A quantitative approach.” *Journal of International Economics* 82 (2):152–167.

Costinot, ARNAUD, DAVE Donaldson, and IVANA Komunjer. 2012. “What Goods Do Countries Trade? A Quantitative Exploration of Ricardo’s Ideas.” *The Review of Economic Studies* :581–608.

David, Paul A. 1990. “The dynamo and the computer: an historical perspective on the modern productivity paradox.” *The American Economic Review* 80 (2):355–361.

Eaton, Jonathan and Samuel Kortum. 1996. “Trade in ideas: Productivity and patenting in the OECD.” *Journal of International Economics* 40 (3-4):251–278.

———. 1999. “International technology diffusion: Theory and measurement.” *International Economic Review* 40 (3):537–570.

———. 2002. “Technology, geography, and trade.” *Econometrica* 70 (5):1741–1779.

———. 2006. “Innovation, diffusion, and trade.” Tech. rep., National Bureau of Economic Research.

Eaton, Jonathan, Samuel Kortum, Brent Neiman, and John Romalis. 2016. “Trade and the global recession.” *The American Economic Review* 106 (11):3401–3438.

Ethier, Wilfred J. 1982. “Dumping.” *The Journal of Political Economy* :487–506.

Griliches, Zvi. 1981. “Market value, R&D, and patents.” *Economics letters* 7 (2):183–187.

Hausmann, Ricardo, Jason Hwang, and Dani Rodrik. 2007. “What you export matters.” *Journal of economic growth* 12 (1):1–25.

Hausmann, Ricardo and Bailey Klinger. 2007. *The structure of the product space and the evolution of comparative advantage.* Center for International Development at Harvard University.

Hidalgo, César A, Bailey Klinger, A-L Barabási, and Ricardo Hausmann. 2007. “The product space conditions the development of nations.” *Science* 317 (5837):482–487.

Jaffe, Adam B, Manuel Trajtenberg, and Rebecca Henderson. 1993. “Geographic localization of knowledge spillovers as evidenced by patent citations.” *the Quarterly journal of Economics* 108 (3):577–598.

Keller, Wolfgang. 2004. “International technology diffusion.” *Journal of economic literature* 42 (3):752–782.

Kortum, Samuel S. 1997. “Research, patenting, and technological change.” *Econometrica* :1389–1419.

Levchenko, Andrei A and Jing Zhang. 2016. “The evolution of comparative advantage: Measurement and welfare implications.” *Journal of Monetary Economics* 78:96–111.

Mancusi, Maria Luisa. 2008. “International spillovers and absorptive capacity: A cross-country cross-sector analysis based on patents and citations.” *Journal of International Economics* 76 (2):155–165.

Perla, Jesse, Christopher Tonetti, and Michael E Waugh. 2015. “Equilibrium technology diffusion, trade, and growth.” Tech. rep., National Bureau of Economic Research.

Proudman, James and Stephen Redding. 2000. “Evolving patterns of international trade.” *Review of international economics* 8 (3):373–396.

Rosenberg, Nathan. 1982. *Inside the black box: technology and economics*. Cambridge University Press.

Sampson, Thomas. 2016. “The Global Productivity Distribution and Ricardian Comparative Advantage.” .

Santacreu, Ana Maria. 2015. “Innovation, diffusion, and trade: Theory and measurement.” *Journal of Monetary Economics* 75:1–20.

Somale, Mariano. 2014. “Comparative advantage in innovation and production.” Tech. rep., Mimeo, Princeton University.

Waugh, Michael. 2010. “International trade and income differences.” *The American Economic Review* 100 (5):2093–2124.

Appendix

A Model Equations

There are 14 endogenous variables and we need 14 equations. The endogenous variables are

$$\{\pi^j_{in}, T^j_i, c^j_i, W_i, P^j_n, X^j_{ni}, X^j_n, P_n, Y_n, \Phi^j_n, C_n, s^j_n, V^j_n, A^j_n\}$$

The corresponding equations are as follows:

(1) Probability of Imports

$$\pi^j_{ni} = T^j_i \frac{\left(c^j_i d^j_{ni}\right)^{-\theta}}{\Phi^j_n}, \tag{43}$$

with

$$T^j_i = A^j_i T^j_{p,i}. \tag{44}$$

(2) Import shares

$$X^j_{ni} = \pi^j_{ni} X^j_n. \tag{45}$$

(3) Cost of production

$$c^j_n = \Upsilon^j W^{\gamma^j}_{nt} \prod_{k=1}^{J} (P^k_n)^{\gamma^{jk}}. \tag{46}$$

(4) Intermediate good prices in each sector

$$P^j_n = A^j \left(\Phi^j_n\right)^{-1/\theta}. \tag{47}$$

(5) Cost distribution

$$\Phi^j_n = \sum_{i=1}^{M} T^j_i \left(d^j_{ni} c^j_i\right)^{-\theta}. \tag{48}$$

(6) Price index

$$P_n = \prod_{j=1}^{J} \left(\frac{P^j_n}{\alpha^j}\right)^{\alpha^j}. \tag{49}$$

(7) Labor market clearing condition

$$W_n L_n = \sum_{j=1}^{J} \gamma^j \sum_{i=1}^{M} \pi_{in}^j X_i^j. \tag{50}$$

(8) Sector production

$$X_n^j = \sum_{k=1}^{J} \gamma^{kj} \sum_{i=1}^{M} X_i^k \pi_{in}^k + \alpha^j P_n Y_n. \tag{51}$$

(9) Final production

$$P_n Y_n = W_n L_n + \frac{\sum_{j=1}^{J} \sum_{i=1}^{M} \pi_{in}^j X_i^j}{1+\theta}. \tag{52}$$

(10) Resource constraint

$$Y_n = C_n + \sum_{k=1}^{J} s_n^k Y_n. \tag{53}$$

(11) Innovation

$$\dot{A}_{nt}^j = \sum_{i=1}^{M} \sum_{k=1}^{J} \varepsilon_{ni}^{jk} \int_{-\infty}^{t} e^{-\varepsilon_{ni}^{jk}(t-s)} \alpha_{is}^k \left(s_{is}^k\right)^{\beta^k} ds. \tag{54}$$

(12) R&D expenditures

$$\beta^j \lambda_{nt}^j V_{nt}^j \left(s_{nt}^j\right)^{\beta^j - 1} = P_{nt} Y_{nt}. \tag{55}$$

(13) Value of an innovation

$$V_{nt}^j = \int_t^{\infty} \left(\frac{P_{nt}^j}{P_{ns}^j}\right) e^{-\rho(s-t)} \frac{\Pi_{ns}^j}{A_{ns}^j} ds + \sum_{i=1 i\neq n}^{M} \sum_{k=1 k\neq j}^{J} \int_t^{\infty} \left(\frac{P_{it}^k}{P_{is}^k}\right) e^{-\rho(s-t)} (1 - e^{-\varepsilon_{in}^{kj}(s-t)}) \frac{\Pi_{is}^k}{A_{is}^k} ds, \tag{56}$$

$$\Pi_{nt}^j = \frac{1}{(1+\theta)} \sum_{i=1}^{M} X_{it}^j \pi_{int}^j. \tag{57}$$

B Model Equations (Normalized)

In what follows, we report the equations of the model after normalizing the endogenous variables.

(1) Probability of imports

$$\pi_{ni}^{j} = \hat{T}_{i}^{j} \frac{\left(\hat{c}_{i}^{j} d_{ni}^{j}\right)^{-\theta}}{\hat{\Phi}_{n}^{j}}, \tag{58}$$

where $\hat{T}_{n}^{j} = \frac{T_{n}^{j}}{T_{M}^{J}}$ and $\hat{\Phi}_{n}^{j} = \frac{1}{T_{M}^{J}} \frac{\Phi_{n}^{j}}{(W_M)^{-\theta}(T_{M}^{J})^{\Lambda j}}$ with Λ^{j} defines in Appendix E.

(2) Import shares

$$\hat{X}_{ni}^{j} = \pi_{ni}^{j} \hat{X}_{n}^{j}. \tag{59}$$

(3) Cost of production

$$\hat{c}_{n}^{j} = \Upsilon^{j} \hat{W}_{n}^{\gamma^{j}} \prod_{k=1}^{J} (\hat{P}_{n}^{k})^{\gamma^{jk}}. \tag{60}$$

(4) Intermediate good prices in each sector

$$\hat{P}_{n}^{j} = B \left(\hat{\Phi}_{n}^{j}\right)^{-1/\theta}. \tag{61}$$

(5) Cost distribution

$$\hat{\Phi}_{n}^{j} = \sum_{i=1}^{M} \hat{T}_{i}^{j} \left(d_{ni}^{j} \hat{c}_{i}^{j}\right)^{-\theta}. \tag{62}$$

(6) Price index

$$\hat{P}_{n} = \prod_{j=1}^{J} \left(\frac{\hat{P}_{n}^{j}}{\alpha^{j}}\right)^{\alpha^{j}}. \tag{63}$$

(7) Labor market clearing condition

$$\hat{W}_{n} L_{n} = \sum_{j=1}^{J} \gamma^{j} \sum_{i=1}^{M} \pi_{in}^{j} \hat{X}_{i}^{j}. \tag{64}$$

(8) Sector production

$$\hat{X}_{n}^{j} = \sum_{k=1}^{J} \gamma^{kj} \sum_{i=1}^{M} \pi_{in}^{k} \hat{X}_{i}^{k} + \alpha^{j} \hat{Y}_{n}. \tag{65}$$

where $\hat{Y}_{n} = \frac{P_n Y_n}{W_M}$.

(9) Final production

$$\hat{Y}_{n} = \hat{W}_{n} L_{n} + \frac{\sum_{j=1}^{J} \sum_{i=1}^{M} \pi_{in}^{j} \hat{X}_{i}^{j}}{1+\theta}. \tag{66}$$

(10) Resource constraint

$$\hat{Y}_n = \hat{C}_n + \sum_{k=1}^{J} s_n^k \hat{Y}_n. \tag{67}$$

(11) R&D expenditures

$$\beta_r \lambda_n^j \hat{V}_n^j \left(s_n^j\right)^{\beta_r - 1} = \hat{Y}_n. \tag{68}$$

(12) Value of an innovation

$$\hat{V}_n^j = \frac{1}{\rho - g_y + g_A} \frac{\Pi_n^j}{\hat{A}_n^j} + \sum_{i=1}^{M} \sum_{k=1}^{J} \left(\frac{1}{\rho - g_y + g_A} - \frac{1}{\rho - g_y + \varepsilon_{in}^{kj} + g_A} \right) \frac{\hat{\Pi}_i^k}{\hat{A}_i^k}, \tag{69}$$

with

$$\hat{\Pi}_{nt}^j = \frac{1}{(1+\theta)} \sum_{i=1}^{M} \hat{X}_{it}^j \pi_{int}^j. \tag{70}$$

C Data Description and Calculation

This appendix describes the data sources and the construction of various variables for the paper. Twenty-eight countries are included in our analysis based on data availability (mostly constrained by the availability of the R&D data): Australia, Austria, Belgium, Canada, China, Czech Republic, Estonia, Finland, France, Germany, Hungary, India, Israel, Italy, Japan, Korea, Mexico, the Netherlands, New Zealand, Norway, Poland, Portugal, Slovenia, Spain, Slovakia, Slovenia, the United Kingdom, and the United States. The model is calibrated for 2005. Eighteen tradable sectors and one aggregate nontradable sector are under consideration and reported in Table C.1.

Bilateral trade flows at the sectoral level Bilateral trade data at the sectoral level (expenditure by country n of sector j goods imported from country i, X_{ni}^j) are obtained from the OECD STAN Bilateral Trade Dataset. Values are reported in thousands of U.S. dollars at current prices. Sectors are recorded at the ISIC (rev. 3) 2-3 digit level and are aggregated into the 19 sectors as listed in Table C.1. We use the importer reported exports in each sector as the bilateral trade flows because it is generally considered to be more accurate than the exporter reported exports.

Value added and gross production Domestic sales in sector j, X_{nn}^j, are estimated based on the *domestic* input-output table provided by the OECD STAN database, which contains data at the ISIC 2-digit level that can be easily mapped into our 19 sectors. OECD provides separate IO tables for domestic output and imports. We sum up the values for a given row before the column "Direct purchases abroad by residents (imports)" to obtain X_{nn}^j. We compare this way of estimating the domestic expenditure on domestic product with an alternative calculation based on $X_{nn}^j = Y_n^j - \sum_{i \neq n}^M X_{in}^j$, where both gross production of country n in sector j, Y_n^j, and the total exports from n to i in sector j, $\sum_{i \neq n}^M X_{in}^j$, are from the OECD STAN Database for Structural Analysis. The first method proves to be superior, as the second generates a number of negative observations for some country-sectors. However, data are missing for India, for which we use the INDSTAT (2016 version) provided by United Nations Industrial Development Organization (UNIDO).

Trade barriers and gravity equation variables Data for variables related to trade costs used in gravity equations (such as geographic distance and common border dummies) at the country-pair level are obtained from the comprehensive geography database compiled by CEPII. The WTO's RTA database provides information on regional trade agreements. The currency union indicator is obtained from Rose (2004) and was updated to reflect Euro-area membership.

Wages Average annual wages are reported by the OECD labor statistics at current prices in local currency. They are transformed into U.S. dollars at the 2005 exchange rates to obtain the variable w_n in the model. However, wage data for China, India, and New Zealand are missing in this database, and are obtained from the International Labor Organization (ILO).

Factor shares and final consumption shares In our analysis, we used the U.S. factor shares in 2005 for all countries. Data on the share of materials from sector k used in the production in sector j, γ^{jk}, as well as the labor share of production in sector j, γ^j, come from the Input-Output Database maintained by OECD STAN. The I-O table gives the value of the intermediate input in row k required to produce one dollar of final output in column j. We then divide this value by the value of gross output of sector j to obtain γ^{jk}. Similarly, the labor share is calculated as the ratio of value added to gross output, as capital input does not exist in the model. In addition, the final consumption expenditure shares of each sector, α_n^j also come from the I-O matrix.

R&D data R&D expenditures at the country-sector level are obtained from the OECD database of Business Enterprise R&D expenditure by industry (ISIC Rev 3). Since sectoral R&D data for China, India and Sweden and several sectors in other countries are missing, we obtain estimates of these missing observations using the following approach. First, we run a regression using existing country-sector specific R&D and patent data from USPTO for 2005:

$$\log(R_n^j) = \beta_0 + \beta_1 \log\left(PS_n^j\right) + \mu_n + \gamma_j + \varepsilon_n^j, \tag{71}$$

where R_n^j is the R&D dollar expenditure of country i in sector j and PS_n^j is the patent stock of country i in sector j. μ_i and γ_j are country and sector fixed effects. This relation is built on the observations that (i) in the steady state, R&D expenditure should be a constant ratio of R&D stock and (ii) innovation input (R&D stock) is significantly positively related to innovation output (patent stock). In fact, the coefficient β_1 is large and significant at 99% and the R^2 is close to 0.90. Assuming that the relationship captured by equation (71) holds for China, India, and Sweden, we can obtain the fitted value of their sectoral level R&D expenditure:

$$\log(\widehat{R}_n^j) = \widehat{\beta}_0 + \widehat{\beta}_1 \log\left(PS_n^j\right) + \widehat{\mu}_n + \widehat{\gamma}_j.$$

For these three countries, we have information on all the right-hand-side variables except for the country fixed effects, $\widehat{\mu}_n$. This allows us to compute the *share* of R&D in a given sector for each country as

$$\widehat{r}_n^j = \frac{\widehat{R}_n^j}{\sum_j \widehat{R}_n^j} = \frac{(PS_n^j)^{\widehat{\beta_1}} \exp(\widehat{\mu}_n) \exp(\widehat{\gamma}_j)}{\sum_j (PS_n^j)^{\widehat{\beta}} \exp(\widehat{\mu}_n) \exp(\widehat{\gamma}_j)} = \frac{(PS_n^j)^{\widehat{\beta_1}} \exp(\widehat{\gamma}_j)}{\sum_j (PS_n^j)^{\widehat{\beta_1}} \exp(\widehat{\gamma}_j)}.$$

Second, we obtain the aggregate R&D expenditure as a percentage of GDP, $R\&D/GDP_n^{WB}$, for each country from the World Bank World Development Indicator database. The country-sector specific R&D can then be estimated as $s_n^j = \widehat{r}_n^j \times R\&D/GDP_n^{WB}$. For the countries with missing sectors, we estimate the fitted value using the same procedure. To maintain consistency across countries, we correct the OECD data-generated total R&D with the World Bank total R&D.

$$s_n^j = R\&D/GDP_n^{WB} \times \frac{R_n^{j,OECD}}{\sum_j R_n^{j,OECD}}$$

This estimated s_n^j is the R&D intensity parameter in equations (17) and (12) used in our quantitative analysis.

Table C.1: List of Industries

Sector	ISIC	Industry Description
1	C01T05	Agriculture, Hunting, Forestry and Fishing
2	C10T14	Mining and Quarrying
3	C15T16	Food products, beverages and tobacco
4	C17T19	Textiles, textile products, leather and footwear
5	C20	Wood and products of wood and cork
6	C21T22	Pulp, paper, paper products, printing and publishing
7	C23	Coke, refined petroleum products and nuclear fuel
8	C24	Chemicals and chemical products
9	C25	Rubber and plastics products
10	C26	Other non-metallic mineral products
11	C27	Basic metals
12	C28	Fabricated metal products, except machinery and equipment
13	C29	Machinery and equipment, nec
14	C30T33X	Computer, Electronic and optical equipment
15	C31	Electrical machinery and apparatus, n.e.c.
16	C34	Motor vehicles, trailers and semi-trailers
17	C35	Other transport equipment
18	C36T37	Manufacturing n.e.c. and recycling
19	C40T95	Nontradables

D The Determinants of Diffusion Speed

Table D.2 examines the determinants of cross-country-sector knowledge diffusion speed by estimating a gravity equation extended to include measures of linguistic and religious distance as well as common history variables that potentially affect effectiveness of interaction and communication, all obtained from the CEPII. We also investigate whether trade plays any role in driving the diffusion speed once distances and historical variables are controlled for. Citing and cited country fixed effects are included to control for country-specific characteristics such as size, level of development, and geography. Since we are interested not only on cross-country but also on cross-sector knowledge diffusion, we also include patent stock in citing and cited country-sectors, and directional sector-pair fixed effects to capture the innate knowledge-spillover relationship between different technologies (Cai and Li, 2016) that are independent of the source and destination countries.

Column (1) to (3) show that knowledge in sectors of country i diffuses faster to sectors of country n when the two countries are linguistically closer to each other or share a common language, share a border or are on the same continent, belong to the same regional free trade agreement (FTA) or currency union, were ever in a colonial relationship before 1945, have shared a common colonizer or were once the same country, or have different latitudes. One country being landlocked or one country once being the colony of the other reduces the knowledge diffusion speed. Interestingly, geographic distance does not play a significant negative role and even an insignificant positive role on knowledge diffusion once trade linkages—that is exports between any country-sector pair combinations are controlled for. Trade linkages are significantly and positively associated with knowledge linkages. The size of the knowledge stock, as reflected in the patent stock, also matters. Higher the stock of knowledge the faster the diffusion speed, while countries with similar knowledge structure tend to diffuse slower.

Table D.2: Determinants of Knowledge Diffusion Speed across Countries and Sectors

Dependent variable: $\log \varepsilon_{ni}^{jk}$			
	(1)	(2)	(3)
Georgraphic distance$_{ni}$)	-0.071	0.038	0.014
	(-1.66)	(0.86)	(0.31)
Border$_{ni}$	0.349***	0.355***	0.348***
	(5.64)	(5.73)	(5.61)
FTA$_{ni}$	0.259***	0.214***	0.225***
	(5.66)	(4.67)	(4.88)
Currency union$_{ni}$	0.582***	0.596***	0.585***
	(9.83)	(10.06)	(9.87)
Common language$_{ni}$	1.035***	0.997***	0.988***
	(18.77)	(18.01)	(17.82)
landlock$_{ni}$	-1.091***	-1.154***	-1.162***
	(-9.25)	(-9.80)	(-9.84)
absolute distance in latitude$_{ni}$	0.015***	0.016***	0.016***
	(9.07)	(9.18)	(9.31)
absolute distance in longitude$_{ni}$	-0.000	-0.000	0.000
	(-0.21)	(-0.40)	(0.01)
Common continent$_{ni}$	0.309***	0.289***	0.277***
	(5.00)	(4.68)	(4.47)
Linguistic distance$_{ni}$	-1.732***	-1.663***	-1.663***
	(-9.80)	(-9.40)	(-9.40)
Religious distance$_{ni}$	0.301*	0.263	0.203
	(2.12)	(1.86)	(1.41)
Colony$_{ni}$	-0.551***	-0.555***	-0.552***
	(-7.13)	(-7.17)	(-7.13)
Common colonizer$_{ni}$	1.754***	1.730***	1.722***
	(6.23)	(6.16)	(6.14)
Colony after 1945$_{ni}$	0.998***	0.991***	1.021***
	(5.89)	(5.85)	(6.01)
Same country$_{ni}$	2.185***	2.145***	2.141***
	(20.72)	(20.36)	(20.32)
$\log X_{in}^{jk}$		0.025***	0.027***
		(5.18)	(5.48)
$\log X_{ni}^{jk}$		0.015**	0.013**
		(3.07)	(2.59)
$\log X_{in}^{kj}$		0.012*	0.010*
		(2.52)	(2.02)
$\log X_{ni}^{kj}$		0.029***	0.030***
		(5.86)	(6.17)
log Patent stock$_n^j$			0.025***
			(3.66)
log Patent stock$_i^k$			0.026***
			(3.87)
$Similarity_{ni}$			-0.256**
			(-2.72)
Citing country FEs	Yes	Yes	Yes
Cited country FEs	Yes	Yes	Yes
Sector-pair FEs	Yes	Yes	Yes
R^2	0.52	0.52	0.52
Num of obs	272,916	272,916	272,916

E The Balanced-growth Path

Here, we derive an expression for the growth rate of the economy along the BGP. First, note that through technology diffusion, the level of knowledge-related productivity, A_n^j, grows at the same rate for every country n and sector j. Therefore, we can pick country M and sector J's technology level to normalize every A_n^j and T_n^j. Normalized variables are denoted with a hat. In particular, $\hat{T}_n^j = \frac{T_n^j}{T_M^J}$.

From equation (55), we normalize the value of an innovation as $\hat{V}_n^j = \frac{V_n^j T_M^J}{W_M}$. Then, from equation (57), profits are normalized as $\hat{\Pi}_n^j = \frac{\Pi_n^j}{W_M}$, and from equation (50), X_i^j is normalized as $\hat{X}_i^j = \frac{X_i^j}{W_M}$ for all j. Hence, expenditures grow at a constant rate for all sectors, since π_{in}^j is constant in the BGP (see equations (43) and (48)). From equations (50) and (52), $P_n Y_n$ grow at the rate of W_M. Note that $g_{w_n} = g_w$ for all n.

To derive an expression for the BGP growth rate of the real output per capita, Y_n, we start from the fact that $\frac{W_n}{P_n Y_n}$ is constant in steady-state. Hence,

$$g_{Y_n} = g_w - g_{P_n}.$$

Using equation (49),

$$g_{P_n} = \sum_{j=1}^{J} \alpha^j g_{p_n^j}.$$

We then derive the expression for $g_{p_n^j}$ from equations (46), (47) and (48). First, we rewrite equation (46) as

$$\frac{c_n^j}{W_n} = \prod_{k=1}^{J} \left(\frac{p_n^k}{W_n} \right)^{\gamma_n^{jk}}.$$

In growth rates, it becomes

$$g_{\tilde{c}_n^j} = \sum_{k=1}^{J} \gamma_n^{jk} g_{\tilde{p}_n^k}, \tag{72}$$

where $\tilde{c}_n^j = \frac{c_n^j}{W_n}$ and $\tilde{p}_n^k = \frac{p_n^k}{W_n}$. From equation (48),

$$g_{\Phi_n^j} = g_T - \theta g_{c_n^j} = g_T - \theta g_{c_i^j}.$$

with $g_T = g_A$.

Hence, $g_{c_n^j} = g_{c^j}$ for all n. Normalizing by wages,

$$g_{\tilde{\Phi}_n^j} = g_T - \theta g_{\tilde{c}_n^j}, \tag{73}$$

where $\tilde{\Phi}_n^j = \frac{\Phi_n^j}{W_n^{-\theta}}$

Combining equation (47) and (73) implies that

$$g_{\tilde{p}_n^k} = -\frac{1}{\theta} g_T + g_{\tilde{c}^k}. \tag{74}$$

Substitution into (72) and using $\sum_{k=1}^{J} \gamma^{jk} = 1 - \gamma^j$, we get

$$g_{\tilde{c}^j} = -\frac{(1-\gamma^j)}{\theta} g_T + \sum_{k=1}^{J} \gamma^{jk} g_{\tilde{c}^k}. \tag{75}$$

We can express the previous expression in matrix form so that

$$\begin{bmatrix} g_{\tilde{c}^1} \\ g_{\tilde{c}^2} \\ \vdots \\ g_{\tilde{c}^J} \end{bmatrix} = -\frac{1}{\theta} g_T \begin{bmatrix} 1-\gamma^1 \\ 1-\gamma^2 \\ \vdots \\ 1-\gamma^J \end{bmatrix} + \begin{bmatrix} \gamma^{11} & \gamma^{12} & \dots & \gamma^{1J} & \\ \gamma^{21} & \gamma^{22} & \dots & \gamma^{2J} & \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ \gamma^{J1} & \gamma^{J2} & \dots & \gamma^{JJ} & \end{bmatrix} \begin{bmatrix} g_{\tilde{c}^1} \\ g_{\tilde{c}^2} \\ \vdots \\ g_{\tilde{c}^J} \end{bmatrix} \tag{76}$$

From here

$$\begin{bmatrix} g_{\tilde{c}^1} \\ g_{\tilde{c}^2} \\ \vdots \\ g_{\tilde{c}^J} \end{bmatrix} = -\frac{g_T}{\theta} (I-A)^{-1} \begin{bmatrix} 1-\gamma^1 \\ 1-\gamma^2 \\ \vdots \\ 1-\gamma^J \end{bmatrix} \tag{77}$$

where

$$A = \begin{bmatrix} \gamma^{11} & \gamma^{12} & \dots & \gamma^{1J} \\ \gamma^{21} & \gamma^{22} & \dots & \gamma^{2J} \\ \vdots & \vdots & \vdots & \ddots \\ \gamma^{J1} & \gamma^{J2} & \dots & \gamma^{JJ} \end{bmatrix}$$

Therefore, the cost of production c_n^j can be normalized as

$$\hat{c}_n^j = \frac{c_n^j}{W_M (T_M^J)^{-\frac{1}{\theta}\Lambda_j}}, \tag{78}$$

where Λ_j is the jth entry of the vector $\Lambda = (I-A)^{-1} \begin{bmatrix} 1-\gamma^1 \\ 1-\gamma^2 \\ \vdots \\ 1-\gamma^J \end{bmatrix}$.

With this, we can obtain an expression for the growth rate of real output as

$$g_{Y_n} = g_w - \sum_{j=1}^{J} \alpha^j g_{p_n^j}.$$

From Equation (74), we have

$$g_{Y_n} = g_w - \sum_{j=1}^{J} \alpha^j \left(\frac{-1}{\theta} g_T + g_{c^j} \right).$$

Based on Equation (78), the above equation becomes

$$g_{Y_n} = g_w - \sum_{j=1}^{J} \alpha^j \left(\frac{-1}{\theta} g_T + g_w - \Lambda_j g_T \right).$$

Therefore,

$$g_{Y_n} = \frac{1}{\theta} \left(1 + \sum_{j=1}^{J} \alpha_j \Lambda_j \right) g_T = g_y, \forall n. \tag{79}$$

Note that in a one-sector economy in which $\gamma^{jk} = 0, \forall n, k$ and $\gamma^j = 1, \forall j$, the growth rate is

$$g_y = -\frac{1}{\theta} g_T.$$

as in Eaton and Kortum (1996, 1999). With multiple sectors, however, the growth rate of the economy is amplified by the input-output linkages.

F One Sector Model

We show that in a one-sector version of our model, changes in trade barriers have no effect on the optimal R&D intensity, hence on growth rates along the BGP. In the one-sector model, $\gamma^j = 1$ and $\gamma^{jk} = 0$. The one-sector version of equations (64), (65) and (66) is

$$\hat{W}_n L_n = \sum_{i=1}^{M} \pi_{in} \hat{X}_i, \tag{80}$$

$$\hat{X}_n = \hat{Y}_n, \tag{81}$$

$$\hat{Y}_n = \hat{W}_n L_n + \frac{\sum_{i=1}^{M} \pi_{in} \hat{X}_i}{1+\theta}. \tag{82}$$

Using equations (80) and (82),

$$\hat{Y}_n = \frac{1+\theta}{\theta}\hat{W}_n L_n$$

and

$$\frac{\sum_{i=1}^{M}\pi_{in}\hat{X}_i}{1+\theta} = \frac{\hat{Y}_n}{2+\theta}.$$

From equations (68) and (69) in a one-sector model with royalties

$$\left(s_n^j\right)^{(1-\beta_r)} = \beta_r\lambda_n\frac{\hat{V}_n}{\hat{Y}_n} = \beta_r\lambda_n\frac{1}{\rho-g_y+g_A}\frac{\sum_{i=1}^{M}\frac{\varepsilon_{in}}{g_A+\varepsilon_{in}}\frac{\sum_{m=1}^{M}\hat{X}_m\pi_{mi}}{1+\theta}}{\hat{Y}_n}. \tag{83}$$

Using the previous expression

$$\left(s_n^j\right)^{(1-\beta_r)} = \beta_r\lambda_n\frac{\hat{V}_n^j}{\hat{Y}_n} = \beta_r\lambda_n^j\frac{1}{\rho-g_y+g_A}\frac{1}{2+\theta}\sum_{i=1}^{M}\frac{\varepsilon_{in}}{g_A+\varepsilon_{in}}\frac{\hat{Y}_i}{\hat{Y}_n}. \tag{84}$$

In this case, changes in trade costs have an effect on optimal R&D intensity to the extent that they have an effect on $\frac{\hat{Y}_i}{\hat{Y}_n}$.

If there are no royalties, the above expression becomes

$$\left(s_n^j\right)^{(1-\beta_r)} = \beta_r\lambda_n^j\frac{1}{\rho-g_y+g_A}\frac{1}{2+\theta}. \tag{85}$$

In this case, changes in trade costs do not have an effect on optimal R&D intensity, hence on the growth rate along the BGP.