Technology Adoption, Innovation, and Inequality in a Global World*

Florian Trouvain
University of Oxford

First version: September 22, 2022
Current version: August 19, 2025
[most recent version here]

Abstract

I develop a tractable semi-endogenous multi-country growth model with a technology adoption margin. Innovation and adoption are skill-intensive activities, and a tradeoff arises whether skilled labor is used to push out the technological frontier or to adopt existing technology. I use the theory to revisit the effect of market integration on growth among asymmetric countries with large differences in innovative capacity. Transitional dynamics and long-run effects implied by the model differ substantially from benchmark endogenous growth models and jointly explain stellar per capita growth in emerging markets and the disappointing growth performance of advanced economies after rising global market integration since the 1990s.

*E-mail: florian.trouvain@economics.ox.ac.uk. This study uses the weakly anonymous Establishment History Panel 1975-2019 data of the German Federal Employment Agency (BA) at the Institute for Employment Research (IAB) in Nuremberg. Data access was provided via on-site use at the Research Data Centre (FDZ) of the IAB and remote data access. I thank Sandra Dummert, Heiner Frank, Lisa Schmidtlein, and Philipp vom Berge for expert research support from the IAB. For generous advising I am extremely grateful to Dominick Bartelme, John Leahy, Dmitriy Stolyarov, and Linda Tesar. For their insights I thank Mark Aguiar, Andres Blanco, Charlie Brown, John Bound, Mike Blank, Paco Buera, Antonio Ciccone, Maarten De Ridder, Max Dvorkin, Jonathan Eaton, Hartmut Egger, John Fernald, Cecilia Fieler, Carlos Garriga, Josh Hausman, Elhanan Helpman, Rishabh Kirpalani, Sam Kortum, John Laitner, Moritz Lenel, Paolo Martellini, Josh Martin, Eduardo Morales, Emir Murathanoglu, Ezra Oberfield, Pablo Ottonello, Stephen Parente, Michael Peters, Lukasz Rachel, B. Ravikumar, Steve Redding, Paulina Restrepo-Echavarria, Richard Rogerson, Hannah Rubinton, Juan Sanchez, Ana Maria Santacreu, Karthik Sastry, Brit Sharoni, Yongs Shin, Sebastian Sotelo, Jagadeesh Sivadasan, Gianluca Violante, Mark Wright, and Fabrizio Zilibotti. I thank the St. Louis Fed and the International Economic Section at Princeton University for their hospitality, and the German Academic Scholarship Foundation for financial support.

1 Introduction

Over the past three decades, globalization has dramatically expanded markets for innovation, driving record levels of patenting, R&D spending, and technological progress. Yet in many advanced economies — including the United States and Western Europe — productivity growth has slowed, wage growth has stagnated, and inequality has surged. This paradox lies at the heart of modern globalization: how can innovation be accelerating while incomes for many are falling behind? In this paper, I develop a new general equilibrium growth model that explains why globalization has been accompanied by increasing inequality and unexpectedly weak growth in the very countries driving the world's technological frontier.

The key novelty is to embed a skill-intensive technology adoption margin into an otherwise standard model of endogenous long-run growth. By adoption I mean a costly and purposeful effort to use a specific technology, which is different from diffusion that happens automatically, or imitation related to the question of the ownership of technology. Adoption here concerns the mundane but essential activity of implementing new technologies into a firm's production process. While often an afterthought, it is clear that frontier innovation is inconsequential without adoption. The main assumption in the model is that both the invention of new ideas and the adoption of existing ones requires skilled labor.

A tradeoff between adopting technology vis-a-vis pushing out the technological frontier emerges as both activities ultimately have to draw on the same scarce factor, skilled labor. This tension becomes particularly acute during an episode of market integration between advanced economies and emerging markets. In the integrated equilibrium, emerging markets adopt advanced economies' technology. Foreign technology adoption raises the returns to innovation, and induces a reallocation of skilled labor within advanced economies from domestic adoption toward innovation for a larger global market. Since aggregate growth depends on both the technological frontier and the degree of technology adoption, positive innovation effects can be entirely undone by weak technology adoption, explaining uneven and sluggish growth in advanced economies as emerging markets join the world economy. This innovation-adoption tradeoff is distinct from the more commonly studied tradeoff between innovation vs. production because production is easily outsourced to emerging markets while local technology adoption is not. Formalizing and exploring this mechanism by means of a tractable general equilibrium model is the main contribution of the paper.

My analysis proceeds in three steps. In the first step, I develop a closed economy version of Romer (1990)'s benchmark two-sector growth model with endogenous technology adoption. Firms in the research sector invent new technology using skilled labor, while firms in the production sector combine capital goods with production labor to produce output. My key departure is that production sector firms solve a dynamic adoption problem: they decide how much skilled labor they devote to adopting new capital goods, which embody frontier technology. The solution pins down an endogenous technology adoption gap.

The closed economy setup highlights two countervailing forces that determine the decentralized equilibrium. On the one hand, the model features a complementarity between adoption and innovation on the market for ideas. Ideas only become profitable after they are adopted so that higher adoption effort in the production sector pushes up the net present value of developing an idea where I use the term idea and innovation interchangeably. Intuitively, innovators would like their ideas to penetrate the market as quickly as possible. On the other hand, by virtue of modeling both innovation and adoption as skill-intensive activities, a factor market rivalry arises as innovators and adopters compete for the same scarce resource, skilled labor. If one activity expands, it does so at the expense of the other in this general equilibrium setting.

Despite the fact that innovation and technology adoption are forward-looking interdependent activities, the steady state is surprisingly tractable and the equilibrium adoption gap is a negative power function of the skill premium. This simple relationship unlocks all key results in the paper, and captures the intuition that when skill becomes relatively more scarce, production sector firms endogenously adopt less technology.

In the second step, I generalize the model to a multi-country setting to study economic growth in an integrated world. I embed the model into a semi-endogenous growth framework à la Jones (1995) to allow for a realistic degree of scale effects. The model is consistent with key stylized facts of economic growth endogenous in open economies. First, the long run equilibrium is characterized by a stationary cross-country income distribution where all countries grow at the same rate. Yet, growth spurts are possible when countries join the global economy and gain access to the world technological frontier. Both aspects follow from a model with endogenous technology adoption that features an advantage of backwardness. Second, countries with higher research productivity, larger populations, or greater shares of skilled labor contribute more to the global technological frontier. Third, while the technological frontier exhibits scale effects and depends on the total size of the global research force, a country's individual productivity is unrelated to its size as it adopts technologies invented elsewhere. Incorporating these features is necessary for any coherent treatment of technological change in a globalized world. This is a difficult task in models that only allow for innovation, but becomes trivial once innovation and adoption are modeled jointly.

I next establish a sharp theoretical distinction between symmetric market integration among similar countries—such as the formation of the European Common Market in the late 20th century—and asymmetric integration between emerging markets and advanced economies. Symmetric integration amongst advanced economies has no adverse effects on inequality or technology adoption and delivers the standard gains from integration. The reason is that foreign adoption, which raises the incentive to innovate, and foreign innovation, which reduces it, exactly cancel. Underlying this result is a classic Ricardo-Heckscher-Ohlin type logic where for identical technology and factor endowments prices remain unchanged, while gains from integration are driven by standard scale effects as each country uses the other country's technology.

This contrasts with the globalization shock of the 1990s and 2000s, marked by a pronounced asymmetry between advanced economies ("West") and emerging markets ("East") ushering in a period of unprecedented market integration. Consider the following thought experiment: each economy is on a balanced growth path, and an unanticipated liberalization in the East moves each country from autarky to free trade in final goods and technology, reflecting the fall of the Iron Curtain or China's ascension to the World Trade Organization. Assuming that the West develops frontier technology and can partially protect its intellectual property, fast technology adoption of emerging markets in the integrated equilibrium raises the returns to innovation in the West. This leads to heightened research effort in the West and temporarily faster growth of the technological frontier, which is the standard mechanism emphasized in the literature. While this effect is present in my framework, there is an important twist to this logic: productivity not only depends on the technological frontier but also on its adoption.

An expanding research sector induces a rise in the skill premium, causes a reallocation of skilled labor away from the production sector, and ultimately hurts technology adoption. The chain of events is most easily understood through the lens of the Heckscher-Ohlin model where skill becomes scarce in advanced economies due to heightened specialization in innovation in the integrated equilibrium, which coincides with a rising skill premium and a widening technology adoption gap.

Since aggregate growth in the West depends on both the evolution of the technological frontier and the degree of technology adoption, the impact of an asymmetric globalization shock is ex-ante unclear. Whether positive or negative effects dominate depends on externalities in innovation and technology adoption. I explore this by contrasting the efficient allocation with the decentralized equilibrium. I find that when ideas are harder to find (Bloom et al., 2020) and there are mild learning spillovers in adoption, underinvestment in adoption relative to innovation occurs. In that case, integration with a foreign economy that adopts technology aggressively but contributes little frontier technology themselves induces a growth slowdown in advanced economies. Note that sluggish growth in the West is consistent with exceptional growth in the East, which is often ignored in the rich-country focused debate around the productivity slowdown. In fact, the mechanism relies on it as key driving force since technology adoption in the East induces the bias in favor of global innovation vs. local adoption in the West.

In the third step, I combine cross-country data supplemented with worker and establishment-level micro data to quantify the impact of globalization on growth and inequality. Solving for the entire transition path and focusing on the West, I find that the asymmetric globalization shock generates a sizable increase in innovation and employment of skilled labor in the research sector. The skill premium increases by 25%, which leads production sector firms to cut down on skilled labor for adoption purposes, ultimately reducing technology adoption by 20%. The calibration predicts a cumulative drop in real wages of production workers of 13%, relative to the counterfactual balanced growth path in autarky. While skilled labor gains in real terms, after aggregating up worker incomes, I arrive at a cumulative growth drag of about 9% for the aggregate wage bill. The impact on total per capita income is around minus 5%, and the difference is accounted for by the role of asset accumulation and royalties earned abroad. High innovative activity and an increase in the valuation of technology coincides with sluggish per capita growth as global innovation crowds out domestic technology adoption.

Transition dynamics are crucial to understanding these results. While productivity growth in the very long run tracks growth of the technological frontier, along the transition path, they are divorced. A widening adoption gap exerts downward pressure on realized productivity growth rates despite elevated growth of the technological frontier.

In contrast to this bleak scenario, full convergence of the East in terms of skill endowments and innovative capacity would restore the powerful pro-growth effects of market integration while simultaneously reducing inequality in the West, consistent with the theoretical results on symmetric integration. The model thus suggests that the rise of India or China as research powerhouses would be extremely desirable from the point of view of unskilled workers in the West, which contrasts with popular fears about emerging markets' abilities to compete in high-tech.

I conclude by citing additional evidence in favor of the main mechanism and discussing related issues such as skill biased technological change and automation. Overall, little attention has been paid to the role of technology adoption in a globalized world,[1] perhaps in part due to the complexity of incorporating both innovation and adoption at once. I provide a tractable framework –directly nesting benchmark growth models–[2] to study the intricate interplay between innovation and technology adoption that should be relevant to a host of issues at the nexus of growth and inequality.

Relationship to the literature. I build on the literature on endogenous growth (Romer, 1990; Grossman and Helpman, 1991b; Aghion and Howitt, 1990), and incorporate a technology adoption margin into Jones (1995). A large literature uses variants of these models to study the aggregate productivity slowdown,[3] focusing on falling population growth and declining business dynamics as drivers.[4] Relative to this literature, I focus on the distinction between innovation and technology adoption, and emphasize the latter to understand patterns of global growth.[5]

[1]Jovanovic (1997) argues that costs of technology adoption as a share of GDP exceed costs associated with innovation by an order of magnitude.

[2]The framework is equally applicable to models of vertical innovation (Grossman and Helpman, 1991b; Klette and Kortum, 2004), see work in progress by Trouvain and Violante (2025).

[3]The productivity slowdown is a robust feature of the data, although its onset differs somewhat across countries. Fernald (2015) and Cette, Fernald, and Mojon (2016) point out that this slowdown started before the financial crisis.

[4]See Peters and Walsh (2021), Jones (2020), Engbom et al. (2019), and Hopenhayn, Neira, and Singhania (2018) on works that highlight the role of population growth on productivity and business dynamics. An alternative explanation for the productivity slowdown focuses on firm dynamics and biased technology shocks, see De Ridder (2019), Olmstead-Rumsey (2019), Rempel (2021), Akcigit and Ates (2019), and Aghion et al. (2019), building on Klette and Kortum (2004). Akcigit, Ates, and Impullitti (2018) consider a two-country model, and explore the impact of trade policy on business dynamics and growth. Incorporating an endogenous technology adoption margin along the lines proposed here should be complementary to the overall agenda in this subfield.

[5]The theory shares some predictions with models of directed technological change (Acemoglu, 2002; Acemoglu, 2003; Ace-

Second, a large literature studies the role of technology adoption for cross-country income differences, see Parente and Prescott (1994), Lucas (2009a), Comin and Hobijn (2010), Lucas and Moll (2014), or Comin and Mestieri (2014). Barro and Sala-i-Martin (1997) and Acemoglu, Aghion, and Zilibotti (2006) study developing economies' choice between adoption and innovation.

A small number of papers has modeled innovation and adoption jointly. Konig et al. (2021), building on König, Lorenz, and Zilibotti (2016), as well as Benhabib, Perla, and Tonetti (2021), Hopenhayn and Shi (2020), and Sampson (2023) develop heterogeneous firm models where high productivity firms innovate, while laggard firms imitate high productivity firms. Comin and Gertler (2006) and Anzoategui et al. (2019) study business cycle fluctuations in a model with both innovation and adoption. The complementarity between innovation and adoption on the market for ideas is related to but distinct from Comin and Hobijn (2007), which focuses on initial implementation of new technologies. Relative to these works, I draw out an innovation-adoption tradeoff embedded in an otherwise standard semi-endogenous growth model with realistic scale effects.[6]

Third, I relate to the literature on trade and growth. Rivera-Batiz and Romer (1991) highlight the strong pro-growth effect of market integration. Eaton and Kortum (1999) develop a model of global growth with exogenous diffusion of ideas.[7] Sampson (2016; 2023) and Perla, Tonetti, and Waugh (2021) highlight the role of learning in heterogeneous firm settings, similar to Buera and Oberfield (2020)'s quantitative model of international technology diffusion. The dynamic gains from trade are large in much of the literature and hard to square with the growth slowdown. Modeling both innovation and adoption, and recognizing the inherent tension this generates in a general equilibrium model, can break the strong pro-growth effects in the case of asymmetric integration whilst preserving the positive link between market size and innovation.[8]

The rest of the paper proceeds as follows. Section 2 presents a model of innovation and adoption. Section 3 models the open economy. Section 4 quantifies the model and studies transition dynamics. Lastly, I discuss empirical evidence consistent with the key mechanism and conclude.

moglu, Gancia, and Zilibotti, 2015), models of appropriate technology (Atkinson and Stiglitz, 1969; Basu and Weil, 1998; Acemoglu and Zilibotti, 2001; Caselli and Coleman, 2006), and models of technological revolutions (Greenwood and Yorukoglu, 1997; Caselli, 1999).

[6]The innovation-adoption tradeoff contrasts with the previous literature, which focuses on heterogeneous firm setups where leading firms innovate and less productive firms learn from leading firms. Since innovators don't internalize the learning spillover, innovation is undersupplied, see König, Lorenz, and Zilibotti (2016), Benhabib, Perla, and Tonetti (2021), and Sampson (2023). In such settings, globalization has strong pro-growth effects as it favors larger innovative companies that can overcome the fixed cost of exporting (Melitz, 2003), and partially corrects for the market failure as in (Sampson, 2016; Perla, Tonetti, and Waugh, 2021).

[7]A quantitatively focused trade literature builds on this line of work, see Lind and Ramondo (2022), Cai, Li, and Santacreu (2022), and Somale (2021). Relative to this literature, I keep the model simple to draw out a novel innovation-adoption tradeoff in the presence of an asymmetric market integration shock but the framework developed here in a steady states is as tractable as quantitative trade models so one could incorporate many countries and bilateral trade cost.

[8]A related literature in international trade has largely focused on the uneven distributional aspects of international trade, which have been found to be modest, see Helpman (2016). In this literature the closest paper Arkolakis et al. (2018), which embeds export platforms into a multi-country Melitz model to study specialization in innovation and production. In contrast, I provide a dynamic model with endogenous technology adoption focused on the interplay between frontier innovation and adoption.

2 Closed Economy

2.1 Environment

I outline the economic environment next, which is cast in continuous time. All derivations and proofs are deferred to the appendix.

Households. A representative household supplies their labor inelastically, which leads to an economy wide endowment of L units of production labor and H units of high skilled labor. The household's labor endowment grows at rate $g_H = g_L > 0$. Factors earn income at wage rates $w_{L,t}$ and $w_{H,t}$, and the skill premium is defined as $s_t := \frac{w_{H,t}}{w_{L,t}}$. The household solves a consumption-saving problem

$$\begin{aligned} U &= \max_{\{c_t, B_t\}_{t\geq 0}} \int_0^\infty e^{-(\rho - g_L)t} \log c_t \, dt \\ \text{s.t.} & \\ \dot{B}_t &= r_t B_t + w_{H,t} H_t + w_{L,t} L_t - C_t, \end{aligned} \tag{1}$$

with the usual transversality condition in place. Total assets in the economy are denoted by B, changes in total assets $\dot{B}$ represent net savings and r is the return on assets. Per capita consumption growth follows from the solution to (1) and equals $\frac{\dot{c}}{c} = r_t - \rho$.

Production. A competitive final goods producer combines differentiated intermediate goods using a Benassy (1996)-CES aggregator

$$Y_t = M_t^{-\frac{1}{\sigma-1}} \left(\int_{i \in \Omega_{M_t}} y_{i,t}^{\frac{\sigma-1}{\sigma}} \, di \right)^{\frac{\sigma}{\sigma-1}}, \tag{2}$$

where M_t is the mass of intermediate goods Ω_{M_t}.[9] Final output, which serves as the numeraire, can be consumed or turned into an investment good, $Y_t = C_t + I_t$, where C_t and I_t are aggregate consumption and investment. Physical capital depreciates at rate δ_k giving rise to the standard neoclassical capital accumulation, $\dot{K} = I_t - \delta_k K_t$.

I assume that intermediate goods producers compete monopolistically, after paying a fixed entry cost f_{E} in terms of unskilled labor. Entry occurs up until the cost of entry matches the value of entry, V_{E}, which is an endogenous object and derived below. The only force inducing firm exit is an exogenous death shock arriving at rate δ_{X}.

Infinitesimal firms $i \in \Omega_{M_t}$ produce differentiated intermediate goods using capital goods $x_{ij,t} \in$

[9] This assumption ensures that there are no scale effects in the production sector. An alternative setup building on Hopenhayn (1992) is presented in the appendix where no such assumption is needed to maintain a production sector that is constant-returns-to-scale. The payoff here is that constant-returns-to-scale production on the firm level makes the exposition easier.

$\Omega_{A_{i,t}}$ and production labor $l_{i,t}$ according to

$$y_{i,t} = \left(\int_{j\in\Omega_{A_{i,t}}} \left(\frac{x_{ij,t}}{\alpha}\right)^{\alpha} dj\right)\left(\frac{l_{i,t}}{1-\alpha}\right)^{1-\alpha}. \quad (3)$$

Production labor is hired at rate $w_{L,t}$, and differentiated capital goods are rented at rate $p_{x_{j,t}} = p_{x_t}$ since capital goods are assumed to be symmetric. This leads to a simple static cost minimization problem where the firm takes as given the set $\Omega_{A_{i,t}}$ and factor prices to minimize production cost given (3).

The static firm problem is standard and deferred to the appendix. The key feature is that firm optimally spreads its capital expenditure evenly across different capital goods, $p_{x_t} x_{ij,t} = \frac{p_{x_t} X_{i,t}}{A_{i,t}}$ where $A_{i,t} \in R^+$ is the measure of set $\Omega_{A_{i,t}}$ and $X_{i,t} = \int x_{ij,t} dj$ is total physical units of capital in firm i. This leads to a setup where $A_{i,t}$ can be interpreted as firm productivity operating through a variety effect as in the original Romer (1990) model. In a monopolistically competitive setting,[10] firm operating profits, which are the difference between sales relative to the cost of production, are a simple log-linear function of aggregate demand $Y_{i,t}$, factor prices, markup, and firm productivity

$$\pi^o_{i,t} = \frac{Y_{i,t}}{\sigma}\left(\frac{\sigma}{\sigma-1} p^{\alpha}_{x,t} w^{1-\alpha}_{L,t}\right)^{1-\sigma} A^{(1-\alpha)(\sigma-1)}_{i,t}. \quad (4)$$

Technology Adoption. Technology adoption is equivalent to making new capital goods available on the firm level, i.e., expanding the set $\Omega_{A_{i,t}}$. Relative to the set $\Omega_{A_{i,t}}$, the superset $\Omega_{A_{\mathrm{F},t}}$ contains all existing capital goods including recent inventions not adopted yet, and the associated measure $A_{\mathrm{F},t}$ represents the technological frontier. Following Nelson and Phelps (1966), I make two crucial assumptions. First, technology adoption requires skilled labor. Second, technology adoption features an "advantage of backwardness". Equation (5) formalizes these concepts

$$\dot{A}_{i,t} = \nu A^{1-\theta}_{\mathrm{F},t} A^{\theta}_{i,t} h^{\beta}_{i,t}, \quad (5)$$

where $h_{i,t}$ is the amount of skilled labor hired by firm i for adoption purposes, $1-\theta \in (0,\infty)$ governs the advantage of backwardness, and the parameter $\beta \in (0,1)$ induces diminishing returns to technology adoption at a point in time. The parameter $\nu > 0$ is a constant assumed to be sufficiently small to avoid a corner solution at $A_{i,t} = A_{\mathrm{F},t}$.[11]

[10] Since technology adoption is costly and firms produce according to a constant-returns-to-scale technology, I have to deviate from the benchmark competitive production sector. This is, of course, the same argument put forth in Schumpeter (1942) and Romer (1990) as to why constant-returns-to-scale competitive economies are not consistent with theories of endogenous growth.

[11] An alternative is to use the original Nelson-Phelps specification $\dot{A}_{i,t} = (A_{\mathrm{F},t} - A_{i,t})\, g(h_{i,t})$ for some monotone function g, which does not change any qualitative insights of the model but ensures that no matter how much skilled labor is used, the firm never hits a corner solution. My specification has an additional degree of freedom in θ which allows me to match the speed

The setup gives rise to a dynamic tradeoff between the cost of skilled labor today, and the future benefit of adopting technology, which can be formalized using a HJB equation. It is convenient to use a normalized level of technology relative to the technological frontier $z_{i,t} := \frac{A_{i,t}}{A_{\mathrm{F},t}}$, and normalize the firm value function of an incumbent firm V by the wage $w_{L,t}$, i.e., $v_t := \frac{V_t(z_t)}{w_{L,t}}$ where I drop the i subscript after recognizing that the only way in which firms can potentially differ is in their normalized productivity z, which is the key individual state variable to keep track of in the dynamic problem

$$\begin{aligned} v_t\left(r_t - g_{w_L,t} + \delta_{\mathrm{X}}\right) &= \max_{h_t} \tilde{\pi}_t\left(z\right) - s_t h_t + \dot{z}_t \partial_z v_t + \dot{v}_t \\ \text{s.t.} & \\ \dot{z}_t &= \nu z_t^{\theta} h_t^{\beta} - g_{\mathrm{F},t} z_t, \end{aligned} \tag{6}$$

where $g_{w_L}, \tilde{\pi}, \partial_z v, \dot{v}, g_{\mathrm{F}}$ represent production worker wage growth, normalized profits $\tilde{\pi} = \frac{\pi^o}{w_L}$, partial derivative of the normalized value function with respect to z, partial time derivative of the normalized value function, and the growth rate of the technological frontier. I now drop time subscripts for ease of notation.

An interior solution to (6) satisfies the first order condition

$$\left[\frac{\beta \nu z^{\theta} \partial_z v}{s}\right]^{\frac{1}{1-\beta}} = h, \tag{7}$$

which is the technology adoption problem at the heart of the paper. The skill premium in (7) appears as the key relative price associated with the cost of technology adoption while the term $\beta \zeta z_i^{\theta} \partial_z v$ captures the marginal benefit.

There is a straightforward link to the q-theory of investment that reveals the key economic intuitions. First, consider the effect of a marginal increase in normalized technology on firm value.

Proposition 1. *The value of an increase in normalized productivity equals*

$$\partial_z v = \int_t^{\infty} e^{-\left[r - g_{w_L} + \delta_X + (1-\theta) g_F - \frac{\dot{s}}{s}\right](u-t)} \left(\frac{z_u}{z_t}\right)^{\theta} \partial_z \tilde{\pi} du \tag{8}$$

where $r - g_{w_L} + \delta_X + (1-\theta)\, g_F - \frac{\dot{s}}{s}$ is an effective discount factor that combines standard inter-temporal discounting and firm death $r - g_{w_L} + \delta_X$ with new terms that depend on the advantage of backwardness $(1-\theta)\, g_F$ and relative skill price growth $\frac{\dot{s}}{s}$. Moreover, the term $\left(\frac{z_u}{z_t}\right)^{\theta}$ captures that for $\theta > 0$ past adoption makes future adoption easier. The term $\partial_z \tilde{\pi}$ captures how productivity improvements raise profits.

Relative to standard q-theory, a new term emerges that accounts for the advantage of backwardness $(1-\theta)\, g_{\mathrm{F}}$. Intuitively, if the advantage of backwardness is strong, delaying investment in tech-

of convergence across countries. See Lucas (1993), Parente and Prescott (1994), and Sampson (2023) for similar formulations.

nology adoption is beneficial, which appears as a higher effective discount factor. Similarly, if there is an anticipated increase in the relative price of skilled labor $\frac{\dot{s}}{s}$, investing in technology adoption is especially advantageous today relative to tomorrow, pushing down the effective discount factor. Lastly, θ is inversely related to the advantage of backwardness, and captures the idea that adopting technology is easier for firms that already use sophisticated technology. This parameter will be key in generating realistic TFP convergence dynamics across countries.

Second, and in contrast to standard q-theory, note that the mapping from skilled labor to technology adoption is non-linear, even in the long run since $\beta < 1$. This ensures that production sector firms always hire a positive but finite amount of skilled labor for technology adoption. An implication is that changes in the skill premium induce finite adjustments on technology adoption and skilled labor demand, in contrast to the standard q-theory where investment is perfectly elastic in the long-run.

I show in the appendix how to derive the optimal dynamic path of skilled labor devoted to technology adoption by solving a standard Hamiltonian yielding the following differential equation

$$\frac{\dot{h}}{h} = \frac{1}{1-\beta}\left\{r - g_{w_L} + \delta_{\mathrm{X}} + (1-\theta)\, g_{\mathrm{F}} - \frac{\dot{s}}{s} - \frac{\beta h^{\beta-1}\nu z^{\theta}}{s}\frac{\tilde{\pi}\,(1-\alpha)\,(\sigma-1)}{z}\right\}, \tag{9}$$

where $\frac{\tilde{\pi}(1-\alpha)(\sigma-1)}{z} = \partial_z \tilde{\pi}$ captures the partial effect of productivity on profits and follows from differentiating (4) with respect to productivity. Along a balanced growth path, equation (9) will deliver a closed-form solution for firms' technology adoption expenditure shown below.

Entry and Adoption Spillover. I close the production sector by assuming free entry after paying a fixed cost f_{E} in terms of unskilled labor. Entrants draw their productivity from the distribution of incumbents written in normalized form

$$v_{\mathrm{E}} = \int v\,(z)\, dF_t\,(z)\,, \tag{10}$$

$$v_{\mathrm{E}} \leq f_{\mathrm{E}} \tag{11}$$

where $v\,(z)$ is the value of a production sector firm of normalized productivity z, dF_t is the incumbent distribution over z, and the expected present discounted value of entry must not exceed the entry cost.

Equation (10) embodies the key spillover in technology adoption, which is that entrants learn from incumbents. This assumption features prominently in the idea flow literature, especially Luttmer (2007), and is even more natural in the context of technology adoption: a spillover is necessary in any growth model where private firms' technology adoption is a key ingredient of long-run growth whenever there is positive firm entry. To see why, imagine that incumbents improve their productivity at a constant rate by adopting new technology, while entrants enter at some fixed productivity level. In that case the firm size distribution would be ever diverging.

A convenient implication of the spillover embedded in the entry technology in (10) is that the model collapses to a homogeneous firm model. Since entrants are, on average, as productive as incumbents, and technology adoption in (5) has no exogenous firm-specific parameters, the only fixed point is one where firms make identical choices.[12] In that case, the value function admits a closed form solution

$$v = \frac{\tilde{\pi} - sh}{r - g_{w_L} + \delta_{\mathrm{X}}}, \tag{12}$$

valid off and on a balanced growth path as long as the free entry condition is binding. The value function in (12) takes into account that the normalized cost of technology adoption, sh, needs to be subtracted from operating profits, and then applies the common discount factor consisting of interest rate, wage growth, and exogenous exit rate.

The creation of production sector firms depends on how much unskilled labor is devoted to entry, L_{E}, and exit rate δ_{X} as follows

$$\dot{M} = \frac{L_{\mathrm{E}}}{f_{\mathrm{E}}} - \delta_{\mathrm{X}} M, \tag{13}$$

where L_{P} denotes unskilled labor devoted to output production.

Research Sector. I largely follow the setup in Romer (1990) where innovators pay a fixed cost in skilled labor to create new ideas, which enter the economy in the form of differentiated capital goods. The key difference to Romer (1990) is that innovators can only monetize their innovation if production firms have adopted their technology. Forward-looking innovators thus have to take into account when their ideas are adopted, which generates a feedback between innovation and adoption absent from the benchmark model.

For simplicity, I assume that older ideas are adopted first.[13] The waiting time for an innovation to become profitable, i.e., to be adopted, is defined as $\tau_t \in R^+$. Denote by $\pi_{\mathrm{I}}(u)$ the flow profits generated by an adopted innovation so the present discounted value equals

$$V_{\mathrm{I},t} = \int_{t+\tau_t}^{\infty} \exp\left(-\int_t^u r_v dv\right) \pi_{\mathrm{I}}(u)\, du, \tag{14}$$

where the discount factor runs from t onward although profits accrue only from $t+\tau_t$ with τ_t referring to the waiting time until adoption. The link between waiting time and profitability is simple: if it took forever for an idea to be adopted, the value of the innovation would be zero.

The assumption that ideas are adopted according to when they were invented induces a simple law of motion for the waiting time, which makes this forecasting problem tractable. Formally, define

[12] I show in the appendix how to allow for a weaker spillover at entry, for example where the entry distribution $F_t(z|\mathrm{E})$ is modeled as a linear function of average incumbent productivity, $z_{\mathrm{E}} = \lambda_{\mathrm{E}} \int z dF_t(z)$, with $\lambda_{\mathrm{E}} \in (0,1)$. This gives rise to a heterogeneous firm model where entrants have a lower productivity than incumbents, and productivity is increasing in firm age as older firms have adopted more technology.

[13] Implicit in this model is that all firms in the production sector adopt technology in the same order. I develop a version with stochastic adoption in the appendix, which relaxes this assumption without changing any of the main insights of the model. What matters for the key mechanism to go through is that the average expected waiting time is a function of adoption effort in the production sector.

the measure of ideas which stands between the adoption of some cohort t's innovation $W(t,t) := A_{\mathrm{F},t} - A_t$. The first argument in W refers to the time when cohort t paid the fixed cost to innovate, and is constant from cohort $t's$ point of view. The second argument is changing over time, and by construction $W(t, t+\tau_t) = 0$, i.e., the calendar time of adoption for inventor cohort t is $t + \tau_t$. To characterize the evolution of W over time, note that while new ideas may be invented, they will only be adopted after cohort t and are thus irrelevant for cohort t's waiting time τ_t. What matters for cohort t is how quickly the measure W melts away, which depends on the flow of adopted ideas, $Ag_A dt$ where $g_A := \frac{\dot{A}}{A}$ is the growth rate of technology on the firm level. The reduction in W over time thus obeys the differential equation

$$\dot{W} \quad = \quad -g_A A, \tag{15}$$

with boundary condition $W(t,t) = A_{\mathrm{F},t} - A_t$ and $W(t, t+\tau_t) = 0$. Together with a a trajectory of A based on firms' adoption choices, I can solve this differential equation to obtain τ_t.

Proposition 2. *The endogenous waiting time τ_t equals*

$$\tau_t = -\frac{\log z_t}{\frac{\int_t^{t+\tau_t} g_A(x)dx}{\tau_t}}. \tag{16}$$

First, note that (16) only implicitly defines τ_t as it appears both on the left and right hand side. Along a balanced growth path with constant adoption gap and productivity growth, however, it is easy to see that the wait time simplifies to $\tau = -\frac{\log z}{g_A}$ where g_A is the productivity growth rate, and $-\log z$, the technology adoption gap. The waiting time is thus determined by production sector firms as their adoption effort pins down how large the gap between current and frontier technology, $\log z$, is. This measure of physical distance is then divided by the average speed at which new technology is adopted, g_A, which yields a statistic measured in units of time τ.

The rest of the research sector is standard. Innovators produce a flow $\frac{1}{f_{\mathrm{R},t}} A^{\phi}_{\mathrm{F},t}$ of new ideas with one unit of skilled labor, where f_{R} is a fixed entry cost, A^{ϕ}_{F} represents a knowledge spillover and $\phi \in (-\infty, 1)$ governs the strength of the spillover following Jones (1995). This induces to the following law of motion of the technological frontier

$$\dot{A}_{\mathrm{F},t} = \frac{1}{f_{\mathrm{R},t}} A^{\phi}_{\mathrm{F},t} H_{\mathrm{F},t}, \tag{17}$$

where H_{F} denotes the total amount of skilled labor devoted to innovation.

The fixed cost $f_{\mathrm{R},t} = \frac{\left(\frac{H_{\mathrm{F},t}}{L_t}\right)^{1-\lambda}}{\gamma}$ depends on exogenous research productivity γ, and a congestion externality $\left(\frac{H_{\mathrm{F},t}}{L_t}\right)^{1-\lambda}$ parameterized by $\lambda \in (0,1]$. For $\lambda < 1$, innovation becomes harder as the share of labor devoted to innovation increases.[14] After invention, innovators hold on to an infinitely lived

[14]The congestion force is slightly different from the one in Jones (1995) where $f_{\mathrm{R},t} = \frac{(H_{\mathrm{F},t})^{1-\lambda}}{\gamma}$ represents the possibility

patent. They rent capital K from households to supply differentiated capital goods X_j that embody the innovator's idea. The marginal cost of renting out such capital goods to the production sector firms thus equals the rental rate inclusive of depreciation $r+\delta_k$. Given the intermediate goods' producers production function with constant elasticity across differentiated capital goods in (3) and innovators' market power, the usual markup over marginal cost formula applies and leads to the following rental rate that production sector firms face, $p_x = \frac{1}{\alpha}\left(r+\delta\right)$.

For adopted technology, the profits or royalty π_{I} simply equal $\pi_{\mathrm{I}} = \frac{L_{\mathrm{P}} w_L}{A}$.[15]

The equilibrium amount of skilled labor devoted to innovation H_{F} is pinned down by a free entry condition, $V_{\mathrm{I}} \leq \frac{f_{\mathrm{R}} w_H}{A_{\mathrm{F}}^{\phi}}$. Whenever the free entry condition is binding, the inventor value function takes the closed-form

$$V_{\mathrm{I}} = \frac{\exp\left(-\int_t^{t+\tau} r_v dv\right)\pi_{\mathrm{I},t+\tau}\cdot\left[1+\dot{\tau}_t\right]}{r-g_{w_L}-g_s-\left(1-\lambda\right)\left(g_{H_{\mathrm{F}}}-g_L\right)+\phi g_{\mathrm{F}}}, \tag{18}$$

where a standard discount factor $\frac{1}{r-g_{w_L}-g_s-(1-\lambda)\left(g_{H_{\mathrm{F}}}-g_L\right)+\phi g_{\mathrm{F}}}$ accounts for firm death, time discounting, and appreciation of the value of an innovation implied by free entry into the research sector. Non-standard is the term $\exp\left(-\int_t^{t+\tau} r_v dv\right)\pi_{\mathrm{I},t+\tau}\cdot\left[1+\dot{\tau}_t\right]$, which accounts for the fact that profits $\pi_{\mathrm{I},t+\tau}$ arrive with a delay. The delay itself could change over time, $\dot{\tau}_t$, and an increase in the waiting time would necessitate higher firm profits to respect the break-even condition implied by free entry.

Market clearing. Capital market, final goods market, and labor market clearing is imposed

$$\begin{aligned} K &= X \\ Y &= C + I \\ L &= L_{\mathrm{E}} + L_{\mathrm{P}} \\ H &= H_{\mathrm{F}} + H_{\mathrm{D}}, \end{aligned}$$

where H_{D} denotes the aggregate amount of skilled labor devoted to adoption, and is given by $h_i \cdot M$ where the i subscript makes clear that h_i refers to a single production sector firm's skilled labor demand.

of useless duplication, i.e., two researchers coming up with the same idea at the same time. The specification is chosen with an eye toward the open economy, and delivers a well-behaved open economy equilibrium.

[15]To derive $\pi_{\mathrm{I}} = \alpha\frac{L_{\mathrm{P}} w_L}{A}$, note that innovator profits are a constant fraction of capital expenditure, $\pi_{\mathrm{I}} = \frac{1}{A} K p_x \left(1-\frac{r+\delta_k}{p_x}\right) = \frac{1}{A}\left(1-\alpha\right)p_x K$, which can be rewritten in terms of unskilled labor using Cobb-Douglas production $\pi_{\mathrm{I}} = \frac{1}{A}\left(1-\alpha\right)\frac{\alpha}{1-\alpha} w L_{\mathrm{P}}$.

2.2 Balanced Growth Path

I focus on a balanced growth path for expositional purposes, and will consider transition dynamics in the quantitative section. To obtain a stationary system, I introduce the normalized variables $v_{\mathrm{I}} := \frac{V_{\mathrm{I}}}{A_{\mathrm{F}}^{-\phi} w_L}$, $a_{\mathrm{F}} := \frac{A_{\mathrm{F}}^{1-\phi}}{L}$, $\frac{L_{\mathrm{P}}}{L} := l_{\mathrm{P}}$, $h_{\mathrm{F}} := \frac{H_{\mathrm{F}}}{L}$, $h_{\mathrm{D}} = \frac{H_{\mathrm{D}}}{L}$, $h_{\text{tot}} = \frac{H}{L}$, $m = \frac{M}{L}$. The normalized value of an innovation and the free entry condition simplify to

$$v_{\mathrm{I}} = \frac{\exp\left(-\int_t^{t+\tau}\left(r_v + g_A - g_{w_L} - g_L\right)dv\right)}{r_t - g_{w_L} - g_s - (1-\lambda)\left(g_{H_{\mathrm{F}}} - g_L\right) + \phi g_{\mathrm{F}}} \frac{\alpha l_{\mathrm{P},t+\tau}}{z a_{\mathrm{F}}} \cdot \left[1 + \dot{\tau}_t\right] \tag{19}$$

$$v_{\mathrm{I}} \leq \frac{s h_{\mathrm{F}}^{1-\lambda}}{\gamma}. \tag{20}$$

The term $\frac{\alpha l_{\mathrm{P},t+\tau}}{z a_{\mathrm{F}}}$ captures a market size effect where a larger production sector $l_{\mathrm{P},t+\tau}$ raises profits, while more competition by other innovators, $z a_{\mathrm{F}}$, lowers them.

Normalized laws of motion for firm creation in the research and production sector read

$$\dot{a}_{\mathrm{F}} = (1-\phi)\left\{\gamma h_{\mathrm{F}}^{\lambda} - a_{\mathrm{F}} \frac{g_L}{1-\phi}\right\}, \tag{21}$$

$$\dot{m} = \frac{l_{\mathrm{E}}}{f_{\mathrm{E}}} - \left(g_L + \delta_{\mathrm{X}}\right) m. \tag{22}$$

Equilibrium Definition, Existence, and Uniqueness. The concept of a balanced growth path is very similar to Jones (1995). The main difference is that equilibrium in the production sector gives rise to a constant technology adoption gap defined as $\Gamma := -\log z$.

Proposition 3. *Along a balanced growth path wages, per capita consumption, productivity, and the technological frontier grow at rate $g_A = g_{w_L} = g_{w_H} = g_F = \frac{1}{1-\phi} g_L$, while population grows at exogenous rate $g_L = g_H > 0$.* The endogenous variables $z, s, a_F, m, l_E, l_P, h_D, h_F$ *are constant and the interest rate equals $r = \rho + g_F$.*

Except for pathological corner solutions, equilibrium is unique and well behaved after imposing the following parameter restrictions.

Assumption 1. *Suppose that the following inequality holds*

$$\rho + \delta_X + (1-\theta)\, g_F > \beta\,(1-\alpha)\,(\sigma-1)\, g_F.$$

This assumption ensures that the benefit of adoption is not too large. If, for instance, varieties were too substitutable ($\sigma \to \infty$), firms' incentive to adopt technology to capture the whole market are too strong, leading to corner solutions.

In a stationary equilibrium I have $\dot{z} = 0$ which, using (6), implies

$$z = \left(\frac{\nu h^{\beta}}{g_{\mathrm{F}}}\right)^{\frac{1}{1-\theta}}. \tag{23}$$

Together with (9) and setting $\dot{h} = 0$, the equilibrium demand of skilled labor of a firm in the production sector takes the simple form

$$h = \frac{1}{s}\frac{\beta\,(1-\alpha)\,(\sigma-1)\,g_{\mathrm{F}}}{\rho+\delta_{\mathrm{X}}+(1-\theta)\,g_{\mathrm{F}}}\tilde{\pi}, \tag{24}$$

which implies that spending on technology adoption is a constant fraction of operating profits. Using the free entry condition, $f_{\mathrm{E}} = \frac{\tilde{\pi}-sh}{\rho+\delta_{\mathrm{X}}}$, I can solve for firm-level skilled labor demand in terms of exogenous parameters and the skill premium, i.e., $h = \frac{1}{s}\Lambda_h$ where Λ_h is a constant, and I can pin down the normalized measure of firms m.[16]

Skill Price and Technology Adoption Gap. The equilibrium adoption gap directly follows by combining (23) and (24)

$$z = \left(\frac{\nu\left(\frac{1}{s}\Lambda_h\right)^{\beta}}{g_{\mathrm{F}}}\right)^{\frac{1}{1-\theta}},$$

which establishes an immediate link between the skill premium and the technology adoption gap.

Proposition 4. *An increase in the skill premium, ceteris paribus, leads to an increase in the technology adoption gap*

$$-\frac{\partial \log z}{\partial \log s} = \frac{\beta}{1-\theta}.$$

Proposition 4 is the central result in the paper on which all other implications hinge: an increase in the skill premium leads to an increase in the adoption gap. This partial equilibrium result is general in the sense that any shock that leads to an increase in the skill premium, which does not directly impact the adoption technology, causes a widening gap between the technological frontier and current productivity. A market integration shock, which I will analyze in the open economy section, can precisely cause such an increase in the skill premium due to specialization in innovation.

Complementarity between Innovation and Adoption. Having understood how equilibrium adoption is pinned down along the balanced growth path, I zoom in on the complementarity between innovation and adoption on the market for ideas. The two are tied together through the endogenous waiting time $\tau = -\frac{\log z}{g_A}$. Substituting out τ when computing the normalized present discounted value

[16]Note $Y\frac{\sigma-1}{\sigma}(1-\alpha) = w_L L_{\mathrm{P}}$ due to Cobb-Douglas production, and $\pi^o = \frac{Y}{M}\frac{1}{\sigma}$, implies $\tilde{\pi} = \frac{l_{\mathrm{P}}}{m}\frac{1}{(1-\alpha)(\sigma-1)}$. Using this together with free entry $f_{\mathrm{E}}(\rho+\delta_{\mathrm{X}}) = \tilde{\pi} - sh$, the law of motion of firm creation $\dot{m} = \frac{l_{\mathrm{E}}}{m f_{\mathrm{E}}} - (\delta_{\mathrm{X}} + g_L) = 0$ in the steady state, and $l_{\mathrm{P}} = 1 - l_{\mathrm{E}}$, implies $m = \frac{1}{f_{\mathrm{E}}((\rho+\delta_{\mathrm{X}})(1-\alpha)(\sigma-1)\Lambda_\pi+\delta_{\mathrm{X}}+g_L)}$ where $\Lambda_\pi = \left(1-\frac{\beta(1-\alpha)(\sigma-1)g_{\mathrm{F}}}{\rho+\delta_{\mathrm{X}}+(1-\theta)g_{\mathrm{F}}}\right)^{-1}$.

of an innovation from (19) and defining the effective household discount factor $\tilde{\rho} = \rho - g_L$ leads to[17]

$$v_{\mathrm{I}} = \frac{1}{\tilde{\rho} + g_{\mathrm{F}}} \frac{\alpha l_{\mathrm{P}}}{a_{\mathrm{F}}} z^{\frac{\tilde{\rho}}{g_A}}, \tag{25}$$

where the term $z^{\frac{\tilde{\rho}}{g_A}}$ is the key difference to a standard Romer model. It captures the fact that a technology adoption gap reduces the present discounted value of an innovation. When there is no adoption ($z \to 0$), the value of an innovation is zero. As ideas are adopted more quickly, the value of innovation increases, and when $z \to 1$ the expression nests the present discounted value in Romer (1990).

While innovators take this adoption gap as given, it should be clear that it is an endogenous outcome depending on adoption effort in the production sector. Increasing technology adoption, ceteris paribus, leads to rising innovation. To see this, combine equation (25) with the free entry condition, and the resource constraint for $\dot{a}_F$ to obtain equilibrium demand for skilled labor from the research sector[18]

$$h_{\mathrm{F}} = \frac{1}{s} \frac{g_{\mathrm{F}}}{\tilde{\rho} + g_{\mathrm{F}}} \alpha l_{\mathrm{P}} z^{\frac{\tilde{\rho}}{g_A}}. \tag{26}$$

Both the positive effect of adoption on innovation through z, and the direct negative effect of the relative cost of skilled workers on innovation, i.e., the skill premium s, appear in (26). The next proposition formalizes the complementarity between adoption and innovation.

Proposition 5. *A reduction in the technology adoption gap, ceteris paribus, increases the measure of frontier technology by a constant elasticity*

$$\frac{\partial \log A_F}{\partial \log z} = \frac{\lambda}{1-\phi} \frac{\tilde{\rho}}{g_F}.$$

The term $\frac{\tilde{\rho}}{g_{\mathrm{F}}}$ characterizes the passthrough from adoption to net present discounted value of an innovation, while the term $\frac{\lambda}{1-\phi}$ characterizes the passthrough from net present discounted value to innovation, which depends on the entry technology in (17).

Both proposition 4 and 5 are partial equilibrium results, and it is clear by now that the relationship between innovation and adoption is complex: on the one hand, there is a complementarity on the market for ideas where higher adoption effort raises the returns to innovation, and hence entry into research, through shorter waiting times. On the other, both activities draw on skilled labor, which

[17] I use the fact that the discounting term simplifies along a balanced growth path, $e^{-\int_t^{t+\tau}(r_v + g_A - g_{w_L} - g_L)dv} = e^{-\tau(\tilde{\rho}+g_A)} = e^{\log z\left(\frac{\tilde{\rho}}{g_A}+1\right)} = z^{\frac{\tilde{\rho}}{g_A}+1}$, and partly cancels with $\frac{1}{z}$ in the denominator in (19). The fact that a wider adoption gap raises profits through its effect in the denominators represents weak competition once an idea is adopted.

[18] Since $\dot{a}_{\mathrm{F}} = 0$ along a balanced growth path, the resource constraint implies $\frac{\gamma h_{\mathrm{F}}^{\lambda}}{g_{\mathrm{F}}} = a_{\mathrm{F}}$. Inserting this into the free entry condition, $v_{\mathrm{I}} = \frac{s h_{\mathrm{F}}^{1-\lambda}}{\gamma}$, yields normalized demand for skilled labor in the research sector.

means that the expansion of one activity has to come at the expense of the other in general equilibrium. It is thus important to consider the effect of shocks in a fully-specified general equilibrium model.

The key market clearing condition is the one for skilled labor,

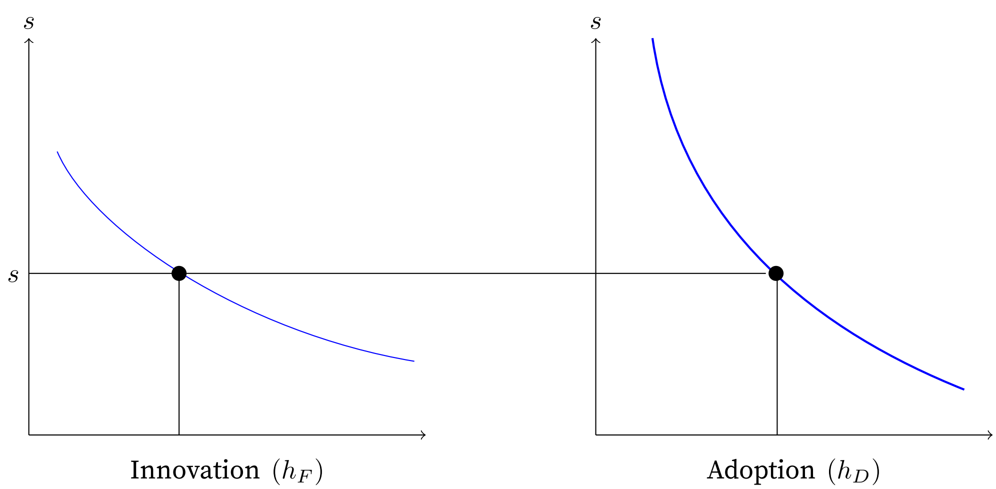
$$h_{\mathrm{F}} + h_{\mathrm{D}} = h_{\mathrm{tot}}.$$

Combining this with (24) and (26) pins down the skill premium

$$\frac{1}{s}(z)^{\frac{\tilde{\rho}}{g_A}}\Lambda_{\mathrm{F}} + \frac{1}{s}\Lambda_{\mathrm{D}} = h_{\mathrm{tot}}, \tag{27}$$

where Λ_{F} and Λ_{D} are constant along a balanced growth path.[19] A simple plot in figure 1 illustrates their interactions. Both adoption activity and innovation activity are downward sloping in the skill premium. While aggregate labor supply is fixed, it is perfectly elastic within sector and equilibrium is reached when the relative demand for skilled labor matches aggregate supply.[20]

Figure 1. Market Clearing for Skilled Labor



Given a solution for the skill premium, the relative share of labor devoted to adoption vis-a-vis

[19]Throughout the paper I focus on equilibria where h_{tot} is sufficiently scarce so that $s > 1$.

[20]Note that (27) is a non-linear equation because z itself is a function of s, which is the only aspect of the closed-economy model requiring a numerical solution.

innovation in the decentralized allocation reads

$$\frac{h_{\mathrm{D}}}{h_{\mathrm{F}}}=\frac{\beta}{1-\theta}\frac{1}{\alpha}\left[\frac{\tilde{\rho}}{g_{\mathrm{F}}}+1\right]\frac{z^{-\frac{\tilde{\rho}}{g_A}}}{1+\frac{\tilde{\rho}+\delta_{\mathrm{X}}+g_L}{g_{\mathrm{F}}(1-\theta)}}. \tag{28}$$

Two aspects of (28) are noteworthy. First, the allocation is unrelated to research productivity γ. Consequently, neither the allocation of skilled labor across sectors nor the skill premium respond to changes in research productivity in the long run. Changes in fundamental research productivity thus do not lead to a reallocation of skill across sectors.

Second, note that a negative shock to the relative supply of skilled labor would reduce both innovation and adoption activity, but the effect on the research sector would be larger due to second round effects: an increase in s first makes the key input, skilled labor, more expensive reducing both activities. Furthermore, a widening technology gap would further decrease innovation through a second round effect as the value of innovation is falling due to a longer waiting time. Vice versa, an expansion in skilled labor will simultaneously reduce the technology adoption gap and boost the share of skilled labor devoted to innovation.

To complete the description of the decentralized equilibrium, note that the equilibrium interest rate $r=\rho+g_{\mathrm{F}}$ determines the price of capital goods $p_x=\frac{r+\delta_k}{\alpha}$. The capital-effective-labor ratio $\frac{K}{zA_{\mathrm{F}}L_{\mathrm{P}}}=B_k\cdot\left(\frac{\alpha}{\rho+g_{\mathrm{F}}+\delta_k}\right)^{\frac{1}{1-\alpha}}$ where $B_k=\alpha^{-\frac{\alpha}{1-\alpha}}\left(1-\alpha\right)^{-1}\left(\frac{\sigma-1}{\sigma}\cdot\alpha\right)^{\frac{1}{1-\alpha}}$ is a constant.[21] The production side aggregates up to the familiar Cobb-Douglas production function

$$Y=\left(\frac{zA_{\mathrm{F}}L_{\mathrm{P}}}{1-\alpha}\right)^{1-\alpha}\left(\frac{K}{\alpha}\right)^{\alpha}, \tag{29}$$

and the unskilled wage is given by

$$w_L=B_w\cdot zA_{\mathrm{F}} \tag{30}$$

where B_w picks up constant parameters. Note that the unskilled wage depends on both the overall technological frontier A_{F} and the degree of technology adoption captured in z.

Lastly, total assets B in the economy consist of physical capital, ownership of production sector firms, and ownership of research firms, i.e., intellectual property

$$B=K+M\cdot V+A\cdot V_{\mathrm{I}}^{\text{adopted}}+\left(A_{\mathrm{F}}-A\right)\cdot\int_0^{\overline{\tau}}V_{\mathrm{I}}\left(\tau\right)dF\left(\tau\right),$$

where the second term adds the value of firms in the production sector, the third term includes the value of adopted ideas, and the final term adds the value of ideas not yet adopted, taking into account that innovations closer to adoption are more valuable with $\overline{\tau}=\sup\{\tau_j\}_{j\in\Omega_{A_{\mathrm{F}}}}$ and $dF\left(\tau\right)$ keeps track

[21] Using total labor $H+L$ instead of unskilled labor engaged in production to define capital intensity only impacts the constant term.

of the distribution of waiting times.

Efficiency. The allocation of skilled labor will be generically inefficient, and depends on externalities in both innovation and technology adoption. In the appendix, I derive the socially optimal allocation of skilled labor where the only margin the social planner decides on is the allocation of skilled labor across sectors, taking endogenous firm entry in the production sector as given. [22] To provide the key intuition, I focus on the case with low effective discount rate, i.e., $\tilde{\rho} \to 0$, which pins down the constrained efficient allocation of skilled labor in closed form, where SP denotes the planner solution and DC the decentralized one

$$\left(\frac{h_{\mathrm{D}}}{h_{\mathrm{F}}}\right)^{\mathrm{SP}} = \frac{\beta}{1-\theta}\frac{1}{\lambda}\left(1-\phi\right), \tag{31}$$

$$\left(\frac{h_{\mathrm{D}}}{h_{\mathrm{F}}}\right)^{\mathrm{DC}} = \frac{\beta}{1-\theta}\frac{1}{\alpha}\left(1\right)\frac{1}{1+\frac{\delta_{\mathrm{X}}+g_L}{(1-\theta)g_{\mathrm{F}}}}. \tag{32}$$

Private and planner allocation differ in four ways. Three of which are standard and arise due to inefficiencies in the research sector, as elaborated in Jones (1995): a markup distortion for $\alpha < 1$, a negative dynamic knowledge externalities for $\phi < 0$, and a congestion externality for $\lambda < 1$.[23]

The novel feature is the externality associated with technology adoption, whereby spillovers at entry lead to too little adoption. The reason is that incumbents do not internalize that entrants benefit from their adoption efforts. This mechanism is captured in (32) by the term $\frac{\delta_{\mathrm{X}}+g_L}{(1-\theta)g_{\mathrm{F}}} > 0$. The inefficiency disappears when $\delta_{\mathrm{X}} = -g_L$, which can be interpreted as incumbents creating additional spin-off establishments at an exogenous rate, which exactly matches the overall population growth rate so that gross entry is zero. If so, firms fully internalize the benefits of adoption.[24]

Whether the autarky decentralized allocation is efficient or not is crucial to understanding the impact of market integration on growth, see Atkeson and Burstein (2010) and Perla, Tonetti, and Waugh (2021). When there is underinvestment in technology adoption to begin with, which is then amplified by specialization in innovation in the open economy, then the model can generate sluggish and uneven growth patterns in the aftermath of market integration across asymmetric countries.[25]

[22] The model is not well-suited to study this margin. For one, the fact that adoption has to happen on the firm level means there is a strange incentive to curtail firm entry. Moreover, I took out scale effects in the production sector using the productivity shifter $M_t^{-\frac{1}{\sigma-1}}$, which introduces inefficiencies unrelated to the key tradeoff of allocating skilled labor to innovation or adoption.

[23] Note that Jones (1995) finds that for reasonable parameterizations the decentralized allocation features too little innovation. Jones' conclusion is partly driven by the assumption that the production sector is efficient. In a model with technology adoption, this tradeoff becomes more interesting since technology adoption features externalities as well.

[24] The influential model of technology adoption of Parente and Prescott (1994) can be understood as an efficient knife edge case where $\delta_{\mathrm{X}} = g_L = 0$ with exogenous frontier growth.

[25] In contrast to learning spillovers at entry, Lucas (2009b), Lucas and Moll (2014), Perla and Tonetti (2014), Perla, Tonetti, and Waugh (2021) highlight knowledge spillovers among incumbents. Two comments are in order. First, note that the spillovers in Perla and Tonetti (2014) and Perla, Tonetti, and Waugh (2021) could be re-interpreted as spillovers to entrants if the lowest productive firms, which are the adopters in their setting, were to be replaced by new entrants. The point is, the mechanics of the spillover are very similar. The framework of Lucas (2009b) and Lucas and Moll (2014) is different as agents continuously

Extensions. The appendix explores several extensions, including heterogeneous firms with imperfect knowledge spillover, complementary public investments to spur technology adoption, the role of immigration, skilled labor in the cost of entry in the production sector, skilled labor directly in the production function, and endogenous skilled labor supply. Movements in the skill premium are key to generating adverse technology adoption effects, and any extension that weakens this channel weakens the mechanism's strength. Of course, however elastic skilled labor is, it appears not elastic enough to undo the substantial increase in the skill premium observed over the past couple of decades.

3 Open Economy

In this section, I generalize the setup to a multi-country open economy model. Countries differ in only two dimensions: research productivity γ_n, relative skill endowment $h_{\text{tot},n}$, where $n \in \{1, ..., N\}$ is a country index of a total of N countries.[26] Non-research related technology including the adoption technology are identical across countries. We shall see that this minimal amount of heterogeneity is sufficient to generate a realistic cross-country income distribution.

International Trade. Countries frictionlessly trade an undifferentiated final good.[27] Capital goods are produced locally to avoid the complex issues of offshoring. Importantly, even if a domestic idea is embodied in foreign capital abroad, the domestic inventor still receives profits coming from the markup applied to the capital good. Trade thus occurs to compensate foreign technology owners for access to their intellectual property, similar to Burstein and Monge-Naranjo (2009) and McGrattan and Prescott (2010).

Adoption in the Open Economy. The problem of a production sector intermediates goods producer, conditional on prices, remains unchanged in the sense that along a balanced growth path firms spend a constant fraction of their operating profits on technology adoption, and I still have $h_{i,k} = \frac{\Lambda_h}{s_k}$. This makes the model very tractable as all local feedback effects from innovation are captured in the relative price of skill, s_k. The only difference for production sector firms is that they now adopt technology from the world technological frontier, which is the sum of all ideas in each country, $A_{\text{F}}^{\text{W}} := \sum_n A_{\text{F},n}$.

A desirable feature of the framework is that it delivers a coherent theory of scale effects across countries, which in this context concerns the link between country size and productivity. Note that most standard growth (and trade models) predict a positive link between country size and productivity

learn from another. I note that adding additional adoption spillovers among incumbents amplifies the main mechanism and renders inefficient adoption more likely.

[26] Allowing for heterogeneous country size L_n is an easy extension pursued in the quantitative section.

[27] Intermediate goods trade a la Krugman (1980) can be added without complication. Exporting neither raises nor reduces normalized profits of firms producing differentiated varieties, so this margin doesn't interact with the incentives to adopt technology.

at odds with the data.[28]

Proposition 6. *A country's relative productivity is determined by the relative amount of skilled labor devoted to technology adoption, and the overall importance of skilled labor in adopting technology parametrized by the advantage of backwardness* $(1-\theta)$ *and static diminishing returns to adoption* (β)

$$\left(\frac{h_{i,k}}{h_{i,j}}\right)^{\frac{\beta}{1-\theta}} = \left(\frac{h_{\mathrm{D},k}}{h_{\mathrm{D},j}}\right)^{\frac{\beta}{1-\theta}} = \frac{A_k}{A_j} = \frac{w_k}{w_j} \quad \forall\, k, j \in C. \tag{33}$$

The proposition follows from noting that the normalized measure of firms, m, will be identical across countries. Consequently, the relative share of skilled labor devoted to adoption, h_{D}, is proportional to the actual number of skilled workers within each firm, h_i. This, in turn, pins down the technology adoption gap and thus productivity. Note that productivity here refers to the rate at which production sector firms turn physical inputs into physical output, which is a simple metric that also determines unskilled worker wages.[29] The model highlights the elevated role of skill in technology adoption, which is consistent both with causal micro evidence (Foster and Rosenzweig, 1996; Foster and Rosenzweig, 2010) and the aggregate relationship between skill and GDP per capita. The setup thus reconciles the findings Mankiw, Romer, and Weil (1992), which focuses on human capital differences across countries, with Parente and Prescott (1994) emphasizing technology adoption and TFP differences as the key driver of cross-country inequality.

While I maintain the non-rivalry of ideas, equation (33) makes clear that there is no link between country size and productivity precisely because smaller countries can adopt technology invented elsewhere. This is why the technological frontier has to be modeled as a global one: If countries were to adopt only their own local innovations, a skill-rich but small country like Denmark would be a poor place. Technically, the key to generating realistic scale effects is to note that endogenous firm entry in the economy ensure that a country that is twice as large requires twice as many skilled workers to keep the amount of skilled labor devoted to technology adoption constant within each firm.[30]

[28] See Alesina, Spolaore, and Wacziarg (2005) for a summary of empirical work on the issue, and Ramondo, Rodríguez-Clare, and Saborío-Rodríguez (2016) for recent work in international trade. In related work, Gross and Klein (2024) differentiate between local and global ideas to deal with cross-country scale effects.

[29] It need not coincide with an aggregate measure of productivity that aims to incorporate gains from trade, see Burstein and Cravino (2015) for an in-depth discussion. More generally, the two-sector model does not map easily into standard national accounting practice. Note that most of technology adoption on the firm level is likely misattributed to production and implicit in standard wage contracts. Similarly, the notion of R&D is a restrictive one as not every innovation is associated with expenditure aimed for patentable intellectual property. My strategy is the quantitative section is to apply a broadly defined notion of research activity and match the employment share associated with this activity. My focus on the term $A = zA_{\mathrm{F}}$ as productivity is consistent with micro-level TFP estimation on the establishment-level in the production sector. This also turns out to be the key metric for unskilled workers later defined as non-college, which constitute the lion share of employment even in advanced economies.

[30] The role of endogenous entry in dealing with scale effects is reminiscent of Young (1998). It seems a natural assumption in the context of adoption since any one idea needs to be adopted multiple times in different firms highlighting the rivalrous nature of technology adoption. The idea the scale effects disappear in the cross-section of countries once technology is adopted or diffuses is not new, and already implicit in Nelson and Phelps (1966). The contribution here is to provide an explicit micro-foundation that tractably embeds this into a model where technology is endogenous, and responds to technology adoption

One caveat is in order. If skill is key, and skill is scarce in poor countries, it ought to be true that the price of skill is high in poorer countries. While skill premia tend to be higher in poor countries, they are not as high as a standard model of neoclassical production with two types of labor like Mankiw, Romer, and Weil (1992) would suggest, see Caselli and Coleman (2006). In the open economy, this concern will be largely irrelevant due to endogenous specialization in innovation related to the classic factor price equalization theorem in international trade due to Stolper and Samuelson (1941).[31]

Technology adoption makes the global idea stock useful for the local economy, and in particular raises labor productivity and real wages of unskilled workers. I emphasize that the allocation of skilled labor to adoption, h_i, is an endogenous outcome in the framework and not only depends on the relative skilled labor endowment, h_{tot}, but also on endogenous specialization in innovation in the open economy creating a tradeoff between between innovation and adoption that impacts both growth and inequality, which I turn to next.

Innovation in the Open Economy. The knowledge spillover is global in the integrated equilibrium,[32] but otherwise the same law of motion as in the closed economy applies, $\dot{A}_{\text{F},k} = \frac{\left(A_{\text{F}}^{\text{W}}\right)^{\phi} H_{\text{F},k}}{f_{\text{R},k}}$ where W denotes world aggregates. Entry is subject to the familiar local congestion externality $f_{\text{R},k} = \frac{\left(\frac{H_{\text{F},k}}{L_k}\right)^{1-\lambda}}{\gamma_k}$. Using the same normalized variables as in the closed economy, the variable $\chi_k := \frac{A_{\text{F},k}}{A_{\text{F}}^{\text{W}}}$ denotes the share of ideas developed in country k along a balanced growth path [33]

$$\chi_k = \frac{\gamma_k h_{\text{F},k}^{\lambda}}{g_{\text{F}} a_{\text{F}}^{\text{W}}} \tag{34}$$

where $a_{\text{F}}^{\text{W}} = \left(\frac{A_{\text{F}}^{\text{W}}}{L}\right)^{1-\phi}$ and $h_{\text{F},k} = \frac{H_{\text{F},k}}{L}$.

In the absence of trade cost the value of an innovation, V_{I}, is unrelated to where the idea was invented so there is no country subscript. Consequently, I can use the free entry condition, $V_{\text{I}} = f_{\text{R},k} w_{H,k} \left(A_{\text{F}}^{\text{W}}\right)^{-\phi}$, to derive countries' relative research effort as a function of skilled worker wages and research productivities $\left(\frac{h_{\text{F},k}}{h_{\text{F},n}}\right)^{1-\lambda} = \frac{\gamma_k}{\gamma_n} \frac{w_{H,n}}{w_{H,k}}$ $\forall n$. Note that the ratio of skilled wages across countries is a function of the relative skill premia, $\frac{w_{H,k}}{w_{H,n}} = \frac{s_k w_{L,k}}{s_n w_{L,n}} = \left(\frac{s_k}{s_n}\right)^{\frac{1-\theta-\beta}{1-\theta}}$ since $w_{L,k} \propto z_k A_{\text{F}}^{\text{W}} \propto s_k^{-\frac{\beta}{1-\theta}} A_{\text{F}}^{\text{W}}$. Together with the resource constraint in (34), differences in skill premia and fundamental research productivities are the only determinants of the share of ideas χ_k,

choices.

[31] The more challenging problem is to generate low skill premia when the emerging market is in autarky. A skill-intensive endogenous technology adoption choice provides some help in this regard as I show in the quantitative section.

[32] See Grossman and Helpman (1991a) for an in-depth discussion of this issue. Global knowledge spillovers seem a natural assumption in a model of long-run growth. The knowledge spillover is assumed to be local in autarky.

[33] Note $\dot{A}_{\text{F},k} = \gamma_k \left(A_{\text{F}}^{\text{W}}\right)^{\phi} h_{\text{F},k}^{\lambda} L$ so along a balanced growth path $\frac{g_{\text{F}}}{\gamma_k} = h_{\text{F},k}^{\lambda} L \frac{A_{\text{F}}^{\text{W}}}{A_{\text{F},k}} \left(A_{\text{F}}^{\text{W}}\right)^{\phi-1} \Rightarrow \chi_k = \frac{\gamma_k h_{\text{F}}^{\lambda}}{g_{\text{F}} a_{\text{F}}^{\text{W}}}$.

$$\chi_k = \frac{\gamma_k^{\frac{1}{1-\lambda}} s_k^{-\frac{1-\beta-\theta}{1-\theta}\frac{\lambda}{1-\lambda}}}{\sum_n \gamma_n^{\frac{1}{1-\lambda}} s_n^{-\frac{1-\beta-\theta}{1-\theta}\frac{\lambda}{1-\lambda}}}. \tag{35}$$

Equation (35) takes the familiar logit-form pervasive in the trade literature since Eaton and Kortum (2002) depending on productivities and prices.

To guarantee a well-behaved open economy equilibrium, I impose the following assumption ensuring that the share of ideas produced in a country is falling in the skill premium.

Assumption 2. *Assume that $\beta < 1 - \theta$.*

Once imposed, the model is well behaved. Since the knowledge spillover is global, and the congestion force local, no one country will capture the entire market for ideas.

To solve for the total amount of skilled labor devoted to innovation, I have to pin down the actual value V_{I}, which now depends on innovator profits across all countries but is otherwise unchanged

$$V_{\mathrm{I}} = \frac{1}{\tilde{\rho} + g_{\mathrm{F}}} \frac{\alpha L_{\mathrm{P}}}{A_{\mathrm{F}}^{\mathrm{W}}} \sum_n w_n z_n^{\frac{\tilde{\rho}}{g_A}}. \tag{36}$$

Equation (36) highlights the role of adoption gaps across countries in determining the value of innovation. There is a direct negative effect of adoption gaps through wait times, $z^{\frac{\tilde{\rho}}{g_A}}$, and an indirect effect through the wage $w_n \propto z_n A_{\mathrm{F}}^{\mathrm{W}}$ where higher real wages imply a larger market.

I show in the appendix how to combine (34), (35), and(36) to obtain the relative share of skilled labor in innovation in country k that is consistent with free entry in the integrated equilibrium.

Proposition 7. *Demand for skilled labor in innovation equals*

$$h_{\mathrm{F},k} = \Lambda_{FO} \frac{\gamma_k^{\frac{1}{1-\lambda}} s_k^{-\frac{1-\beta-\theta}{1-\theta}\frac{1}{1-\lambda}}}{\sum_n \gamma_n^{\frac{1}{1-\lambda}} s_n^{-\frac{1-\beta-\theta}{1-\theta}\frac{\lambda}{1-\lambda}}} \sum_n s_n^{-\frac{\beta}{1-\theta}\left(1+\frac{\tilde{\rho}}{g_A}\right)} \tag{37}$$

where Λ_{FO} picks up terms that are constant along a balanced growth path.

Market Clearing. Combining (37) with aggregate relative demand for skilled labor in adoption, $h_{\mathrm{D},k} = \frac{\Lambda_{\mathrm{D}}}{s_k}$, delivers the key market clearing condition that pins down all other endogenous variables

$$h_{\mathrm{tot},k} = \frac{\Lambda_{\mathrm{D}}}{s_k} + \frac{\Lambda_{\mathrm{FO}} \gamma_k^{\frac{1}{1-\lambda}} s_k^{-\frac{1-\beta-\theta}{1-\theta}\frac{1}{1-\lambda}}}{\sum_n \gamma_n^{\frac{1}{1-\lambda}} s_n^{-\frac{1-\beta-\theta}{1-\theta}\frac{\lambda}{1-\lambda}}} \sum_n s_n^{-\frac{\beta}{1-\theta}\left(1+\frac{\tilde{\rho}}{g_A}\right)}. \tag{38}$$

Given a set of skill premia solving (38), each countries' research share and technology adoption gap follows directly $\{\chi_n, z_n\}_{n\in\{1,..,N\}}$, as well as the normalized world technology $a_{\mathrm{F}}^{\mathrm{W}} = \frac{\sum_n \gamma_n h_{\mathrm{F},n}^{\lambda}}{g_{\mathrm{F}}}$.[34]

[34] To pin down the open economy equilibrium, a simple algorithm based on market clearing in equation (38) works well

Patterns of Trade. Trade is unbalanced. Countries more specialized in innovation earn royalties abroad, matched by final goods flows from the net importers of technology

$$net\_trade\_inflow_k = \frac{\sigma - 1}{\sigma} \alpha (1 - \alpha) Y^{\mathrm{W}} \left[\chi_k - \frac{Y_k}{Y^{\mathrm{W}}} \right]$$

where $Y^{\mathrm{W}} = \sum_n Y_n$ is world GDP. The patterns of net trade depend on the vector of fundamental research productivity $\{\gamma_n\}$ and skill endowment $\{h_{\mathrm{tot},n}\}$, allowing for both Ricardian technology differences and Heckscher-Ohlin factor endowment differences. In an equilibrium where technology is largely produced in the West due to higher research productivity or greater skill endowments, final goods flow from East to West as compensation for technology usage.

3.1 Globalization, Growth, and Inequality

Two points are noteworthy about the relationship between international trade and growth. First, in this simple model the trade elasticity of the final good is infinite, which implies that final goods trade is not important in itself. However, trade allows countries to access foreign technology by compensating the owners of foreign technology with shipments in terms of the final good. Second, gains from integration can be large even when trade flows are relatively small because access to ideas doesn't require that the entire capital good is shipped across countries. It suffices to pay a royalty to the foreign country, constituting only a fraction of the value of the capital good.[35]

The model admits a simple sufficient statistic for the long-run wage effects of market integration, similar to Arkolakis, Costinot, and Rodríguez-Clare (2012) (henceforth ACR).

Proposition 8. *The ratio of unskilled worker wages in the open economy vs. closed economy for some country*

---

where the skill premium is raised whenever there is excess demand for skilled labor in a country using

$$s'_k = \frac{\Lambda_{\mathrm{D}}}{h_{\mathrm{tot},k}} + \frac{1}{h_{\mathrm{tot},k}} \frac{\Lambda_{\mathrm{FO}} \gamma^{\frac{1}{1-\lambda}} s_k^{-\frac{\lambda(1-\theta)-\beta}{1-\theta} \frac{1}{1-\lambda}}}{\sum_n \gamma_n^{\frac{1}{1-\lambda}} s_n^{-\frac{1-\beta-\theta}{1-\theta} \frac{\lambda}{1-\lambda}}} \sum_n s_n^{-\frac{\beta}{1-\theta}\left(1+\frac{\tilde{\rho}}{g_A}\right)}$$

where the new starting point $s_{k,next}$ uses a simple relaxation scheme $s_{k,next} = \tilde{\alpha} s'_k + (1 - \tilde{\alpha}) s_k$, $\tilde{\alpha} \in (0, 1)$. I verify that this procedure converges to the same solution for different starting values. I am not aware of a general uniqueness proof for the multi-country version of this model. After inspecting the market clearing condition in (38), the reader will note that neither the gross substitutes property used in Alvarez and Lucas (2007), nor uniqueness proofs based on non-homogenous integral equations as in Allen and Arkolakis (2014) apply, because of the term $\sum_n s_n^{-\frac{\beta}{1-\theta}\left(1+\frac{\tilde{\rho}}{g_A}\right)}$ which causes an interdependency across countries absent from standard trade models.

[35] A country in my baseline calibration would have to export 14% of its final output when its entire idea stock is held by foreigners. In contrast, in a symmetric equilibrium where every country holds an even share of global technology with sufficiently rich financial markets, no goods would have to be shipped whatsoever as net trade flows are zero and trade in financial asset would render trade in final goods redundant.

k reads

$$\frac{w_k^{open}}{w_k^{closed}} = \underbrace{\left(\frac{h_{\mathrm{F},k}^{open}}{h_{\mathrm{F},k}^{closed}}\right)^{\frac{\lambda}{1-\phi}} \left(\frac{1}{\chi_k}\right)^{\frac{1}{1-\phi}}}_{\textit{Innovation margin}} \underbrace{\left(\frac{s_k^{open}}{s_k^{closed}}\right)^{-\frac{\beta}{1-\theta}}}_{\textit{Adoption margin}}, \tag{39}$$

where I netted out long-run growth. The effect on skilled worker wages follows $\frac{w_H^{open}}{w_H^{closed}} = \frac{w^{open}}{w^{closed}} \cdot \frac{s^{open}}{s^{closed}}$. The key distinction to ACR is that the change of the skill premium forms part of the sufficient statistic accounting for movements in the technology adoption gap.

The impact of market integration on the technological frontier are captured in increasing innovative effort in the home economy $\left(\frac{h_{\mathrm{F},k}^{open}}{h_{\mathrm{F},k}^{closed}}\right)^{\frac{\lambda}{1-\phi}}$ and gains from ideas developed in other countries, $\left(\frac{1}{\chi_k}\right)^{\frac{1}{1-\phi}}$, which depend on a constant scale elasticity $\frac{1}{1-\phi}$ as well as the share of ideas developed in the home economy. If this share is small, the gains are large. The logic is the same as in ACR where a large import share suggest a country has much to lose if it fell back to autarky. The novel feature in (39) is the endogenous adoption margin which shows up in the skill price ratio $\left(\frac{s^{open}}{s^{closed}}\right)^{-\frac{\beta}{1-\theta}}$. An increase in the skill price ratio, ceteris paribus, hurts unskilled workers. The reason is that a rising skill premium leads to less domestic technology adoption, which allows for a richer response of market integration on growth. In particular, the effects are now dependent on whether the country integrates with similar or very different trading partners in terms of research productivity and skill endowments, $\{\gamma_n, h_{\mathrm{tot},n}\}$.

Consider first integration between symmetric countries with $\gamma_n = \gamma$, $h_{\mathrm{tot},n} = h_{\mathrm{tot}}\ \forall n$. This benign scenario delivers the standard variety gains from trade without negative distributional effects, and no changes in the adoption gap, summarized in the following proposition.

Proposition 9. *Symmetric integration leaves the skill premium and the technology adoption gap unchanged relative to autarky, and raises income by a factor of* $N^{\frac{1}{1-\phi}}$.

The proposition follows from noting that for symmetric countries the open economy market clearing condition in (38) collapses to the closed economy. Consequently, skill premium and adoption gap don't respond in the long run. By symmetry, every country produces a fraction $\frac{1}{N}$ of ideas, and plugging this into (39) completes the proof.[36] In the case of symmetric integration, real wage gains accrue to all workers without adverse inequality effects. The strength of this response depends on ϕ, i.e., how much harder ideas are to find. This result contrasts with the case of asymmetric integration described next.

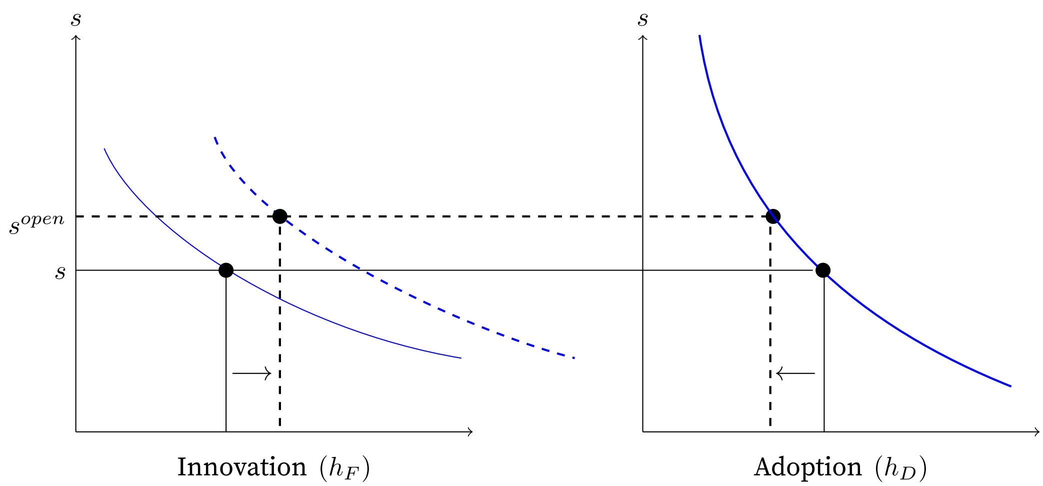
In the case of asymmetric integration, the impact of market integration is ambiguous. First, note that the term $\left(\frac{h_{\mathrm{F},k}^{open}}{h_{\mathrm{F},k}^{closed}}\right)^{\frac{\lambda}{1-\phi}} \left(\frac{1}{\chi_k}\right)^{\frac{1}{1-\phi}}$ will lead to the usual gains driven by technological frontier growth. The non-standard effect arises from the negative link between technology adoption and

[36] This result is reminiscent of Krugman (1980) where trade integration induces variety gains but leaves the measure of varieties in each country unchanged.

the skill premium captured in the last term. If market integration raises the skill premium, then the overall effect on the absolute wages of unskilled labor can easily turn negative, and productivity of production sector firms may fall. Even so, it is easy to prove that skilled workers always gain given $\beta < 1 - \theta$.

Figure 2 summarizes the main argument of this paper graphically where a demand-driven expansion of the research sector in a globalized world leads to a reduction in technology adoption. In general equilibrium these interdependencies are transmitted through an increase in the skill premium. Whether positive innovation or negative adoption effects dominate is a quantitative question. Moreover, the previous discussion was restricted to long-run level effects ignoring transitional dynamics. I next calibrate a more general version of the model to explore these questions quantitatively.

Figure 2. Open Economy Market Clearing for Skilled Labor

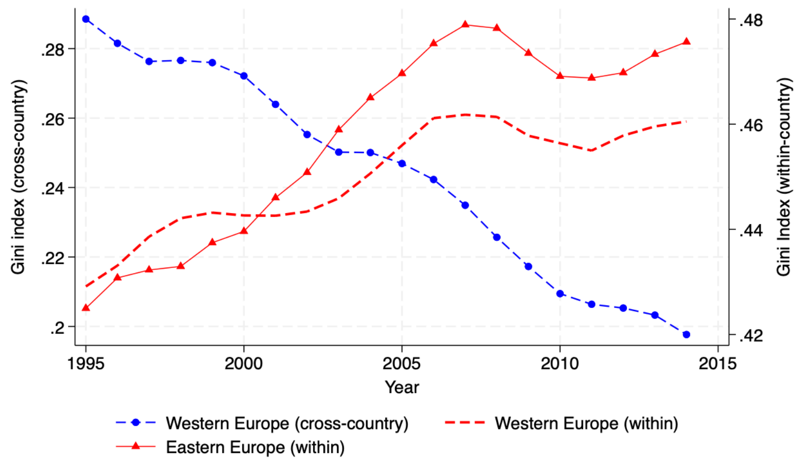


4 Quantification

I quantify a version of the model to explain three striking features of global growth since the mid 90s, two of which are depicted in figure 3.

First, strong growth in emerging markets, especially Eastern Europe and China, have lead to a decline in global inequality measured as dispersion in income per capita across countries. Using PPP-adjusted GDP per capita form the PWT, the cross-country Gini index declines by around seven

Figure 3. Cross-Country Convergence and Within-Country Divergence



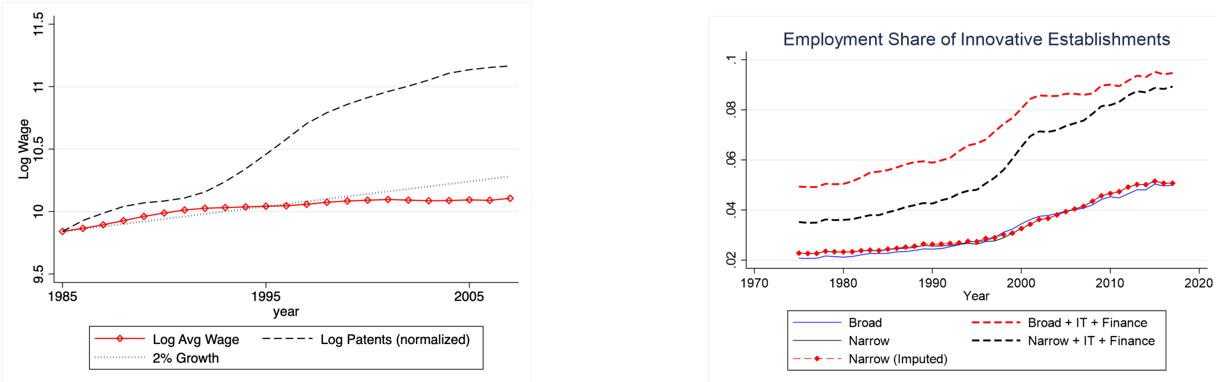
The data is based on the World Inequality Database, see Alvaredo et al. (2020). The gini index is computed over the whole population and uses pre-tax income, split concept. Aggregates are employment-weighted averages within each country group, where I compute the within-country Gini index using a 3-year moving average. Country income is based on ppp-adjusted output per capita using PWT V10. Country list for Western Europe: Austria, Belgium, Denmark, Finland, France, Germany, Iceland, Luxembourg, Netherlands, Norway, Sweden, Switzerland, United Kingdom; Country list for Eastern Europe: Albania, Bosnia and Herzegovina, Bulgaria, Croatia, Czech Republic, Estonia, Hungary, Latvia, Lithuania, Montenegro, Poland, Romania, Serbia, Slovakia, Slovenia. I do not consider Southern European countries.

pp. from 1995 to 2015 when focusing on a set of West and East European economies.[37] Second, within-country income inequality went up substantially. When measured in terms of pre-tax per capita income the rise in within-country inequality matches the decline in cross-country inequality.

Third, while per capita growth was exceptional in Eastern Europe and some parts of Asia, per capita growth in the West was abysmal during this episode of rising market integration. This is particular puzzling against the backdrop of rising research effort and strong patent growth, a common proxy for innovation. I showcase this for the German economy, which has long-run administrative matched employer-employee data including information on worker education, and is a major producer of frontier technology. The left panel in figure 4 contrasts strong patent growth against weak wage growth. While wage growth and patent growth moved in lock-step at the beginning of the sample, they diverge since the 1990s when patenting takes off and wage growth begins to slow down. The right panel in figure 4 documents the increasing total employment share in the research sector, here broadly defined including headquarter services, IT, and Finance. The model developed can account account qualitatively for all these patterns using one simple framework that takes into account the

[37] Convergence is not uniform, and many countries, especially in Africa, have seen much less catchup growth, if at all. See Milanovic (2016) for an in-depth discussion of global convergence. The focus on Europe is simply due to easily comparable data series.

Figure 4. Growth, Patents, and Inequality in Germany



Data for patents comes from the Crios Patstata database, see Coffano and Tarasconi (2014). Wage data is computed based on the PWT version 09, combining real national gdp (not PPP) with their measure of the labor share and dividing thorough by the total population. Patents are normalized so that the wage level and patent level coincide in 1984. GDP per capita growth does better than wages, but still grows substantially below trend, leading to an overall growth slowdown. Employment shares in production and research sector are based on IAB data for West Germany with additional details in the appendix. This classification can be carried out using different sectoral codes (broad vs. narrow where the narrower codes are unavailable for earlier periods) and a question arises whether IT and finance belong into the research or production sector.

interplay of innovation and technology adoption in a globalized world.

I next explore how far quantitatively the model can go in rationalizing these empirical patterns. To do so, I generalize the model along three dimensions. First, I introduce exogenous innovator death shocks that arrive at rate δ_{I}. These shocks hit both adopted and unadopted ideas, which innovators take into account when entering the research sector, and lead to a simple generalization of the baseline model with generalized waiting time, $\tau_t = \frac{\log z_t}{\frac{\int_t^{t+\tau} g_A(x)+\delta_{\text{I}} dx}{\tau_t}}$. Idea death shocks lead to more realistic convergence dynamics, as well as entry and exit dynamics consistent with the data. Second, I now allow for differences in country size L_n. Third, I introduce a technology adoption friction $\mu_{\text{D}} < 1$ operating in the emerging market in autarky similar to Parente and Prescott (1994) to avoid a counterfactually large autarky skill premium in the emerging market explained in more detail below.

Calibration. To quantify the aggregate implications of the theory I have to pin down a set of parameters $\Theta = \{\rho, \alpha, \delta_k, \sigma, f_{\text{E}}, \delta_{\text{X}}, \{L_n\}, \{\mu_n\}, \{\gamma_n\}, \delta_{\text{I}}, \nu, \beta, \theta, \lambda, \{h_{\text{tot},n}\}, g_L\}$ where the inner parentheses and subscript n indicate that the parameters are set separately for each country.

Externally set parameters. I set the capital share equal to .4, capital depreciation equals 5%, and the discount rate is 5%, which, together with income growth of one percent, $g_{\text{F}} = 1\%$, implies a long-run real rate of 6% consistent with stock market returns. I set the elasticity of substitution across intermediate goods varieties equal to 2.5, consistent with evidence from Broda and Weinstein (2006). I set firm exit rates in the production and research sector equal to 4%, i.e., $\delta_{\text{X}} = \delta_{\text{I}} = .04$.[38] These numbers line up well with German establishment micro data from the IAB, where I split establishments into research-intensive and production-intensive firms consistent with the two-sector structure of the

[38] Recall that the higher the arrival rate of death shocks (δ_{X}) in the production sector, the stronger is the adoption externality. I choose conservative values that are lower than the empirical estimates from Garcia-Macia, Hsieh, and Klenow (2019) (ranging from 8% to 4% depending on firm age and time period), Peters and Walsh (2019) (5.4%), or Sampson (2023) (10%).

model.

Adoption. The implications of the theory hinge on the importance of technology adoption, which directly relates to the ratio $\frac{\beta}{1-\theta}$. Cross-country inequality and growth patterns are informative about the ratio $\frac{\beta}{1-\theta}$ since real wage differences for unskilled workers across countries in the steady state equal

$$\frac{w_n}{w_k} = \left(\frac{h_{\mathrm{D},n}}{h_{\mathrm{D},k}}\right)^{\frac{\beta}{1-\theta}}. \tag{40}$$

Conditional on a distribution of the relative amount of skilled labor devoted to adoption across countries $\{h_{\mathrm{D},n}\}$, the parameters (θ, β) translate this distribution into observed cross country unskilled wage inequality.

Taking logs of (40) and adding measurement error u allows me to back out $\frac{\beta}{1-\theta}$ by running the following regression

$$\log z_{n,t} = \alpha + \delta_t + \frac{\beta}{1-\theta} \log h_{\mathrm{D},n,t} + u_{n,t}. \tag{41}$$

The slope coefficient equals $\frac{\beta}{1-\theta}$ where I proxy for unskilled worker wages using GDP per capita and I proxy for h_{D} using the relative share of college-educated workers in each country, i.e. h_{tot}. This approach only works for countries that do little innovation and I thus drop countries above the 90th percentile in terms of output per worker, which account for most innovative effort in the world.[39] Figure 5 shows a simple cross-sectional plot between the relative share of skilled labor and output per worker using the PWT and the Barro and Lee (2013), with an R-square of .55. The skill share is measured as the ratio of workers that have complete tertiary education relative to those who don't consistent with the model. The resulting regression coefficient when dropping the top 10% of richest countries equals 0.74 with robust standard error 0.07 .

While this identifies the ratio $\frac{\beta}{1-\theta}$, I need to pin down each parameter individually. To do so, I chose θ such that the model is consistent with catch-up growth convergence patterns across countries. Barro's "Iron law" (Barro, 1991) suggests countries converge at a rate of 2%, i.e., the coefficient in the cross-country convergence regression, after controlling for additional covariates such as human capital is close to $-.02$. I linearize the law of motion of z to show that the advantage of backwardness $1-\theta$ is the key parameter governing this speed of convergence

$$\frac{\dot{z}}{z} \approx \underbrace{(1-\theta)\left(g_{\mathrm{F}}+\delta_{\mathrm{I}}\right)}_{=\hat{\beta}_B}\left(\log z_{ss}-\log z_t\right)+\beta\left(g_{\mathrm{F}}+\delta_{\mathrm{I}}\right)\left(\log h_{ss}-\log h_t\right).$$

If $\delta_{\mathrm{I}} + g_{\mathrm{F}} = 5\%$, a reasonable estimate for θ is 0.55 which ensures that $\hat{\beta}_B \approx -.02$. This is consistent with Lucas (2009a)'s calibration of the same advantage of backwardness parameter. In that case, the

[39] In closed economies the approach would not work since country-specific technological frontiers would confound the link between relative skill endowments and real wages, hence I run this cross-sectional regression in 2015 to capture the post-integration state.

Figure 5. Relative High-Skill Share and Output p.c.

The plot combines ppp-adjusted output per worker from PWT 9.1 with the Barro and Lee (2013) dataset on schooling for a cross-section of countries in 2015. I equate high-skilled labor with the share of completed tertiary education of the population, and low-skilled labor is its complement. The colored dots refer to Congo, Brazil, and Germany in this order from left to right. When implementing the linear regression I use lagged skill ratios as instrument to deal with measurement error in a two-stage least square regression.

implied parameter is $\beta = .35$.

Innovation. The long-run frontier growth rate is endogenous, and equals $\frac{1}{1-\phi} g_L$. To keep the model simple, I will assume a constant long-run labor force growth rate $g_L = 2\%$ with fixed high-skill vs. low-skill shares.[40]

I assume a dynamic knowledge externality of $\phi = -1$ in the baseline calibration. This parameterization is more optimistic than what the evidence in Bloom et al. (2020) suggests closer to $\phi = -1.5$. Two aspects are noteworthy. First, because there is an endogenous technology adoption gap, the mapping between research effort and productivity is more complicated than in Bloom et al. (2020): the framework here highlights that ideas may appear harder to find, in part, because less ideas are adopted. Second, the sluggish response of advanced economies' GDP growth to globalization makes it unlikely that this parameters is large and positive. Calibrating the model using a more optimistic dynamic externality leads to counterfactually strong growth for all countries.

To calibrate the congestion force λ,[41] I note that cross-country specialization in innovation should

[40]Slowing population growth and rising educational attainment represent two confounding factors that push in opposite direction: slowing population growth reduces long-run growth in idea-based growth models (Jones, 1995) while rising educational attainment is pushing the other way. The model developed here could in principled be used to study this tradeoff, but one would have to think harder about selection into schooling, and the substitutability between skilled and unskilled workers in production vs. technology adoption vs. innovation. These are important issues for future work.

[41]My congestion externality differs slightly from Jones (1995)'s original setting with $f_R \propto H_F^{1-\lambda} A_F^{\phi}$, where the frontier growth rate equals $g_F = \frac{\lambda}{1-\phi} g_L$. If countries have different population sizes and the static congestion is local (which is what

contain information on this parameter. If research productivity among innovating countries was identical, $\gamma_n = \gamma$, the model would imply a log-linear relationship between the share of ideas χ, the total number of researchers H_{F}, and country size measured in terms of the unskilled labor force L in steady state

$$\log \chi_n = \alpha_0 + \lambda \log H_{\mathrm{F},n} + (1-\lambda) \log L_n. \tag{42}$$

I assume that research productivity is similar across advanced economies,[42] and I approximate the share of ideas from country n using average internationally protected patents from country n over the period 2011–2019. Note that I use recent flows to proxy for the stock of ideas, which works in steady state. Combining these assumptions with data on the total number of researchers and total employment from the OECD "Science, Technology, and Innovation" database allows me to use (42) to obtain estimates for λ. The fit between total number of researchers and patenting is almost perfect with an elasticity of unity, and a correlation of .97, see figure **??**, providing some justification for assuming identical research productivity across advanced economies.[43] The fact that global patents are tightly related to a country's number of researchers, which, among rich countries, is tightly correlated with country size, highlights once more why a multi-country semi-endogenous growth model is the right model to study long-run growth. Scale matters, and scale differs across countries.

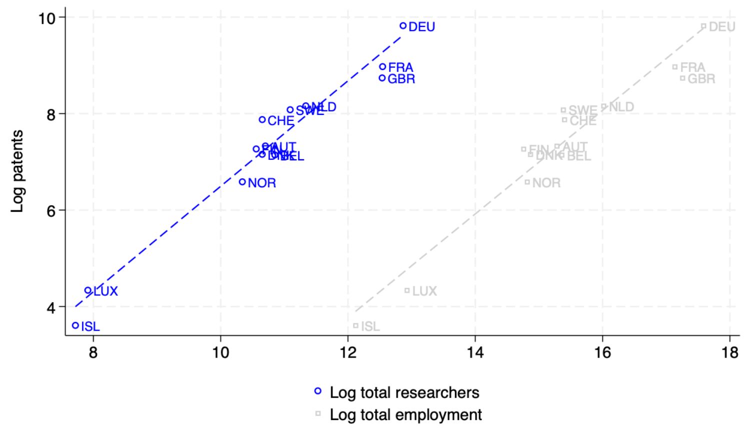
---

one would like for a well-behaved open economy equilibrium), it is preferable to use the ratio of researchers to unskilled labor $\left(\frac{H_{\mathrm{F}}}{L}\right)^{1-\lambda}$. Otherwise, a strong counterfactual bias toward innovation in small countries would emerge since the total number of researchers of a small country is generically smaller.

[42] While Sampson (2023) provides evidence that sectoral research productivity differs across countries, he finds that differences are much larger across emerging vis-a-vis advanced economies. Sectoral heterogeneity likely overstate differences in aggregate research productivity. I thus view his findings as broadly consistent with the spirit of my exercise.

[43] While Sampson (2023) provides evidence that sectoral research productivity differs across countries, he finds that differences are much larger across emerging vis-a-vis advanced economies. Sectoral heterogeneity likely overstate differences in aggregate research productivity. I thus view his findings as broadly consistent with the spirit of my exercise.

Figure 6. Specialization in Innovation



The plot uses data from the OECD "science, technology, and innovation" database. I plot patents flied under the PCT (global patents) against and the total number of researchers similar to Bloom et al. (2020). The set of countries is the same as in the previous plot for the West. Each dot represents a simple average within each country over the period 2011-2019 to avoid the confounding influence of the financial crisis and the pandemic.

Running cross-sectional regressions based on (42) with both total employment and number of researchers is challenging due to high multicollinearity and obvious endogeneity concerns. Nonetheless, I run a linear regression with details deferred to the appendix that suggest a reasonable value of λ around .98. Since small differences in research productivity and skill endowments are likely to induce an upward bias, I make a back of the envelope adjustment and use $\lambda = .9$ as my baseline value.

I am left with the following parameters $\{f_{\text{E}}, \{L_n\}, \{\mu_n\}, \{\gamma_n\}, \nu, \{h_{\text{tot},n}\}\}$. I assume that h_{tot} in the advanced economy is .15, roughly consistent with data from Barro and Lee (2013) over the period 1980 – 2015, and similar to the high skill-low-skill ratio in Acemoglu et al. (2018). I target a skill premium of $s = 1.6$ in the initial equilibrium for the advanced economy based on Buera et al. (2022), and I target a relative productivity level of $z = .75$ in advanced economies in autarky, which means that the waiting time for an idea to be adopted equals 5.7 years. Since I don't have firm heterogeneity, the reader should think of this waiting time as the moment when the product has reached substantial market penetration,[44] and not the first time the capital good is used somewhere in the economy. I then use z and s to pin down the fixed cost of entry in the production sector as well as the shifter in the adoption technology $\{f_{\text{E}}, \nu\} = \{1.25, .21\}$. I normalize unskilled labor in the advanced economy to unity, and assume that the foreign economy is of relative size $\frac{L^*}{L} = .66$, which roughly lines up with the relative size of Eastern Europe.[45] Lastly, I normalize research productivity in the West to

[44]The number lines up well with the study of Golder and Tellis (1997) which finds that it takes on average six years for a new consumer durables to achieve substantial market penetration. Of course, adoption gaps have been associated with virtually any technological innovation, see Griliches (1957) for the classic case of hybrid corn.

[45]The larger the East, the stronger are the uneven effects of market integration in the West. Moreover, note that I abstract

unity without loss of generality.

I will consider different values for γ^* and h^*_{tot} in the quantitative exercise. For the case of asymmetric integration, I assume $\gamma^* \approx 0$, $h^*_{\text{tot}} = .05$, which implies that all innovation is produced in the West. I view this as a central feature of market integration in the 1990s and 2000s, see for instance the OECD study by Khan and Dernis (2006) which documents a large increase in patenting in Europe during this period with almost no patenting activity in Eastern Europe.[46] In an alternative scenario, I will assume skill endowment and research productivity in the East converge to the values in the West. Before I turn to the quantitative exercise, I have to take a stance on the initial equilibrium.

Initial Equilibrium. The initial autarky equilibrium should be consistent with the wage gap across East and West, and the skill premium within each block. Having close to zero research productivity, however, would imply a counterfactually large wage gap. I fix this by assuming that the East's innovation in autarky is easier as it copies technology already invented in the West.[47] Moreover, since skill is relatively scarce in the East, ceteris paribus, the skill premium should be relatively high. In the data, this is not true and I calibrate the model such that the skill premium in East and West are the same in autarky. This can be done by imposing a simple technology adoption wedge similar to Parente and Prescott (1994). Specifically, suppose that there is a market-share reallocation friction $\mu^*_{\text{D,aut}}$ in autarky such that the impact of technology adoption on profits is suppressed

$$\partial_z \pi^*_{\text{autarky}} = \underbrace{\mu^*_{\text{D,aut}}}_{<1} (\sigma - 1)(1 - \alpha),$$

where $\partial_z \pi = (\sigma - 1)(1 - \alpha)$ is the frictionless benchmark always operating in the West, see Hsieh and Klenow (2014) for evidence in favor of such reallocation frictions. In a model with endogenous technology adoption, such frictions reduce the incentive to adopt technology. This suppresses the demand for skilled labor twofold. Fist, directly through lower demand for skilled labor in the production sector, and then indirectly through the equilibrium effect of weak technology adoption on the present discounted value of innovation. Consequently, the skill premium can be arbitrarily depressed.[48] Once markets integrate, this wedge disappears, i.e., $\mu^*_{\text{D,open}} = 1$, which reflects market

away from trade costs and potential bargaining over innovator rents, which would allow me to increase the size of the East and generate the same impact on the West, see the appendix for extensions along these lines.

[46] The contribution of Eastern Europe at the time is so small that it ends up in a residual category. For more recent years, this assumption may be less appropriate, especially with respect to other emerging markets such like China. See

[47] The problem here is that the emerging market block I consider is small with a relative size of .39, and the initial income gap, while large, is not nearly large enough. The model could be applied without this ad-hoc assumption to China, which is a much larger economy starting out from a lower income level. Because the allocation of skilled labor across sectors is unrelated to research productivity in autarky, I simply solve for the allocation in the emerging market using the standard solution routine, and then scale wages by some constant factor, which is meant to capture that copying is easier than innovating. Note that no matter how easy copying is, as long there is some cost, there is no incentive to do so in the integrated equilibrium since Bertrand competition would drive profits to zero.

[48] While this is not the focus of the paper, a skill intensive technology adoption margin provides a promising candidate explanation for low productivity and the weak link between skill scarcity and skill premium. Indeed, Brainerd (1998) summarizes evidence of low levels of inequality within and across worker groups in the Soviet Union, and I loosely follow this evidence

reforms in emerging markets.

Lastly, in a world where ideas are harder to find, it is natural to assume that countries develop the same technology in autarky. While there is no incentive to develop the same idea twice in the integrated equilibrium, varieties could overlap initially so I need to specify which country holds the patent to which variety right after market integration. I assume that the West holds a share ζ of technology while the East holds $1-\zeta$ and provide a micro-foundation based on small quality differences in the appendix. I discipline this initial process of ownership reallocation in the integrated equilibrium with the drop in income in the East right after the fall of the Iron Curtain, which was substantial, and propose a value of $\zeta = .99$.[49] Table 1 summarizes parameters and key moments used for the calibration.

Table 1. Parameterization

Parameter		Value	Target/Source
Household			
ρ	discount factor	0.05	standard value
g_L	pop. growth	0.02	standard value
h_{tot}	rel. skill share	0.15	Barro/Lee (2013)
h^*_{tot}	rel. skill share	0.05	see text
$\frac{L^*}{L}$	rel. size	0.66	see text
$\frac{w}{w^*}$	init. wage gap	4.0	see text
Production/Adoption			
α	capital share	0.4	standard value
δ_k	capital deprecication	0.05	standard value
δ_{X}	exit	0.04	IAB data
σ	substitution	2.5	Broda/Weinstein (2006)
β	static curvature	0.35	see text
θ	adv. backwardness	0.55	see text
f_{E}	entry cost	1.25	skill premium/waiting time
ν	adoption tech.	0.21	skill premium/waiting time
Innovation			
ϕ	dyn. externality	-1.0	see text
λ	congestion	0.9	see text
δ_{I}	exit innovation	0.04	IAB data
γ	research prod.	1	normalization
γ^*	research prod.	0.01	see text
ζ	initial idea share	0.99	initial output drop East

Note: The table provides the parameters used in the baseline calibration, which refers to the case of asymmetric integration.

when calibrating the technology adoption wedge in the East so as to match the skill premium in autarky.

[49] Note that the instantaneous change in GDP in the East at time zero equals approximately $\frac{\Delta Y^*}{Y^*} = -\frac{\sigma-1}{\sigma}\alpha(1-\alpha)$ for ζ close to one, which is the share of output paid to the owners of technology in autarky. For the baseline calibration this implies a drop of 14.4%.

Long-Run Effects. I first report the long-run effects of market integration before I turn attention to transition dynamics. Table 7 reports the results in the form of the cumulative effect, i.e., long-run level effect relative to a balanced growth path in autarky.

In the case of asymmetric integration where all innovation is produced in the West, wages of unskilled workers in advanced economies fall by about 13% in real terms. This contrasts with wage gains for skilled labor of around 12%. Even though the technological frontier increases by 7%, the adoption gap widens by 20% explaining weak wage growth for unskilled labor in advanced economies. I also provide a measure of changes in GDP per capita y.[50] The negative effect of integration is still there, but somewhat weaker than the wage effects, consistent with GDP growth performing better than wage growth data over the past couple of decades, see for instance Autor et al. (2020). The role of intellectual property accumulation and asset income in the open economy explains this difference.

Figure 7. Long-Run Level Effects

Asymmetric Integration

	West	East
χ	1.0	0.0
$\Delta \log w_L$	-0.129	1.276
$\Delta \log w_H$	0.123	1.448
$\Delta \log s$	0.252	0.172
$\Delta \log (y)$	-0.051	1.22
$\Delta \log z$	-0.196	1.209
$\Delta \log A_{\mathrm{F}}^{\mathrm{W}}$	0.067	0.067

Symmetric Integration

	West	East
χ	0.602	0.398
$\Delta \log w_L$	0.253	1.639
$\Delta \log w_H$	0.253	1.584
$\Delta \log s$	0.0	-0.056
$\Delta \log (y)$	0.253	1.614
$\Delta \log z$	-0.0	1.386
$\Delta \log A_{\mathrm{F}}^{\mathrm{W}}$	0.253	0.253

Note. The table summarizes the long-run effect of market integration on income and inequality. The left panel uses the baseline calibration, while the right panel assume that skilled labor endowment and research productivity in the East coincide with the West. Note that the log change in the value of patents is infinite in the case of no innovation in the East in the integrated equilibrium. The adoption gap is defined relative to the global technological frontier $\log z_n = \frac{A_n}{A_{\mathrm{F}}^{\mathrm{W}}}$. For things to add up, I thus defined frontier technological growth as $\Delta \log A_{\mathrm{F}}^{\mathrm{W}} = \log A_{\mathrm{F}}^{\mathrm{W}} - \log A_{\mathrm{F}}$ where A_{F} refers to the frontier in autarky, which resides in the West. Note that I net out exogenous long-run growth $\frac{\Delta \log L}{1-\phi}$ for non-stationary variables.

These negative findings contrast with the growth experience in the East: all workers gain massively albeit skilled labor gains relatively more. The gains are entirely accounted for by the adoption of frontier technology. Because adoption is a skill-intensive activity, there is pressure on the skill

[50] Measuring GDP poses conceptual challenges. An issue arises as to where the value of ideas appears when computing GDP per capita, which itself might depend on firm's profit shifting. I propose the following measure, which corresponds to household consumption per capita. Note that with log utility, households will consume a fraction $\tilde{\rho}$ of their assets. I thus approximate real GDP per capita as $y := \frac{w_L L + w_H H + \tilde{\rho} B}{L+H}$ where B is the value of total household assets. This total value of assets does not change much because an increase in intellectual property roughly cancels with the relative fall in assets held in physical capital and production sector firms. However, because capital income is a substantial share of household income, including this term weakens the overall negative effect.

premium in emerging markets in spite of virtually zero research effort. Importantly, fast technology adoption in the East has a feedback effect on innovation in the West in the integrated equilibrium. This feedback effect induces increasing innovation at the cost of lower technology adoption, with overall negative long-run effects for the West as a whole.

To understand this long-run effects, note that in the baseline calibration the autarky equilibrium in the West is characterized by insufficient adoption. The share of skilled labor devoted to adoption vis-a-vis innovation in the decentralized autarky equilibrium is roughly 40%, whilst a planner would allocate 2/3rds of the skilled work force to technology adoption.[51] This inefficiency is amplified in the open economy explaining the underwhelming impact of globalization on aggregate growth. What is crucial for this outcome is that i) there are learning externalities in adoption, and ii) dynamic knowledge externalities are sufficiently weak. Intuitively, if ideas are harder to find, shifting skilled labor into innovation has only modest effects on frontier growth, which are dominated by the negative impact of reduced technology adoption. An implication of this is that subsidizing innovation in this model would be counterproductive for unskilled workers in advanced economies: the skill premium would widen further, and more labor skilled labor would be diverted away from domestic technology adoption.[52]

This bleak scenario contrasts with the case of symmetric integration where the gains from market integration are positive for all worker groups across all countries. Specifically, wages in advanced economies go up by 25%. This is, of course, just an instance of proposition 9: symmetric integration leaves the allocation of skill across sectors unchanged but delivers desirable scale effects as more skilled workers are engaged in idea production. The pro-growth effects in the case of symmetric integration without any adverse inequality effects illustrate the more complex link between growth and globalization. The model is consistent with the beneficial role of increasing integration across advanced economies in driving strong post-WW2 growth, while simultaneously allowing for the possibility of uneven and sluggish growth in the aftermath of market integration among asymmetric countries. The analysis so far has been restricted to long-run effects and I turn to transition dynamics next.

Transition dynamics. To simplify the transition dynamics, I make two additional assumptions. First,

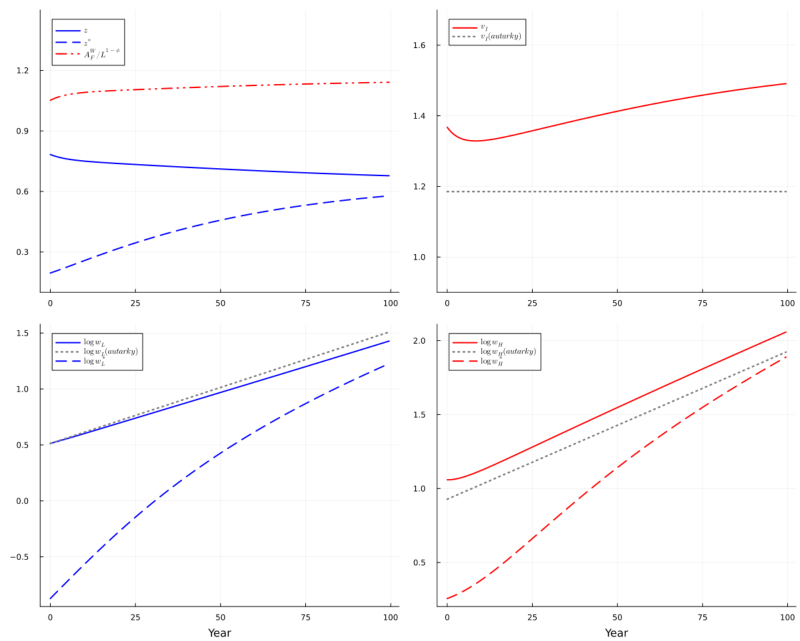
[51] An important assumption underlying this welfare analysis hinges on the planner taking a national perspective. A planner that cares about world output instead faces a different tradeoff. Intuitively, pushing out the frontier helps both the domestic and foreign economies, so more labor is devoted to producing frontier technology. Any welfare statement thus crucially depends on whether the scope of the analysis is global or national.

[52] I find that at for $\phi \approx 0$, all workers gain, and skilled workers gain a lot resurrecting the strong pro-growth effect of integration. Again, the predictions of the theory are directly tied to how much harder ideas are to find. Could it be, then, that all that is needed to reconcile the effects of the recent globalization wave on growth is a model where ideas are harder to find? Not quite. Using Jones (1995)'s model the impact of integration on long-run wages are a log-linear function of the increase in the size of the population. In the integrated equilibrium with only the advanced economy specializing in innovation, market integration delivers an increase in productivity and wages by a factor of $\left(1 + \frac{L^*}{L}\right)^{\frac{\lambda}{1-\phi}}$, which is strictly positive. The assumption that ideas are harder to find is necessary but not sufficient condition, and need to be combined with a model where technology adoption complements frontier innovation.

suppose households reinvest a constant fraction χ_{sav} of the final output good in each country where $\chi_{\text{sav}} = \left(\frac{\sigma-1}{\sigma}\alpha\right) \cdot \alpha \frac{g_L+g_{\text{F}}+\delta_k}{\rho+g_{\text{F}}+\delta_k}$ is set such that the saving rate is consistent with the long-run supply of savings from the household sector.[53] The second simplification is that I keep the normalized measure of firms in the production sector fixed at its long-run level m, which is the same in closed and open economy and across rich and poor countries.[54]

Figure 8 plots the transition dynamics of wages and technology in the case of asymmetric integration. Non-stationary variables are normalized by the long-run growth rate.

Figure 8. Wages & Technology along the Transition Path



Note. Based on baseline asymmetric calibration. Wages are plotted in logs, and the frontier level of technology is normalized by $L^{1-\phi}$ to net out exogenous long-run growth. Grey lines indicate the evolution of the variable in autarky from the point of view of advanced economies.

The left upper panel highlights that innovation responds quickly to a larger global market as the technological frontier expands quickly. The overall modest increase in the technological frontier is a consequence of ideas becoming harder to find. Technology adoption slows in advanced economies, leading to a widening adoption gap and a declining relative technology level z. In contrast, fast catch-up growth in emerging market is evident by the increase in the relative technology level z^*. The upper

[53] In principal, one can solve a model with forward-looking household consumption, which would introduce consumption smoothing and inter-temporal trade. I prefer to simplify, which gives rise to a subtle issue: foreigners now may hold domestic physical capital.

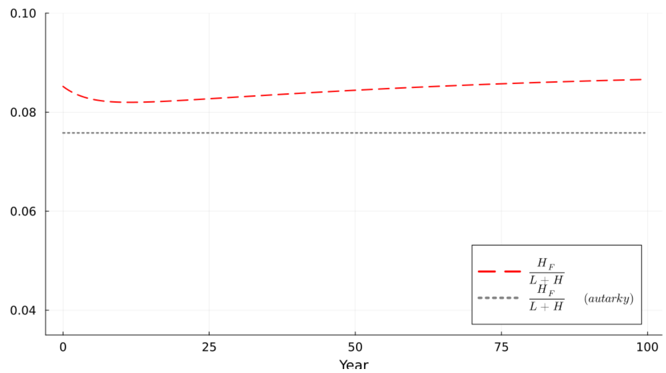
[54] The response of endogenous entry in the production sector is both computationally taxing and matters little quantitatively. I plot the value of firms in the production sector, v, in the appendix, which does not differ by more than 3% compared to its long-run value suggesting the channel is unimportant.

right panel plots a jump in the normalized value of an idea, v_{I}, driven by the initial impact of an increase in market size. This value displays non-monotone dynamics. The slight initial decline is explained by weak adoption in advanced economies, which is a drag on innovator profits. At the same time, the positive impact of the emerging market is growing over time as it travels closer to the technological frontier. Note that technology adoption matters twofold here. First, a rising the real wage makes the foreign market more attractive. Second, the waiting time τ^* is falling over time, further raising the net present discounted value of an innovation.

The lower two panels plot out the wage effects in each country. Unskilled worker wages grow below trend for a long time in advanced economies as technology adoption is relatively slow-moving. In contrast, unskilled workers in emerging markets experience exceptional wage growth driven by the advantage of backwardness. A sufficiently high skill endowment and a relatively large distance to the technological frontier allows for this exceptional growth spurt. Turning attention to high-skilled worker wages, they jump up in both countries. If one were to plot out the skill premium for each country, that skill premium would be falling over time for the emerging market as high returns to technology adoption when far away from the technological frontier drive up the skill premium beyond its long-run value along the transition path. The dynamics in the advanced economy are more subtle, and non-monotone. The skill premium initially spikes, then eases slightly due to sluggish adoption in the advanced economy, but eventually climbs again as technology adoption in the East continues to raise demand for skilled labor in the advanced economy's research sector.

A key statistic in the model is the relative share of labor devoted to research, defined as $\frac{H_{\mathrm{F}}}{H+L}$, which is plotted in figure 9. The dynamics reveal a large initial jump, consistent with the discrete increase in market size and the forward-looking nature of R&D investment followed by a gentle decline which eventually gives rise to steady increases in the research employment share. These non-monotone patterns are driven by a falling technology adoption gap in emerging market and its positive feedback effect on innovation in advanced economies generating long-lasting dynamics. The model predicts an increase of research employment by roughly one percentage point. The magnitude is in the right ballpark but still understates the large increase in research employment in Germany plotted in figure 4. A richer model with endogenous skill-accumulation would certainly be able to generate larger adjustments. Similarly, the right kind of sectoral adjustment costs would generate a monotone increase more similar to the data. Of course, neither is key for the basic point I try to tease out here.

Figure 9. Expanding Research Sector

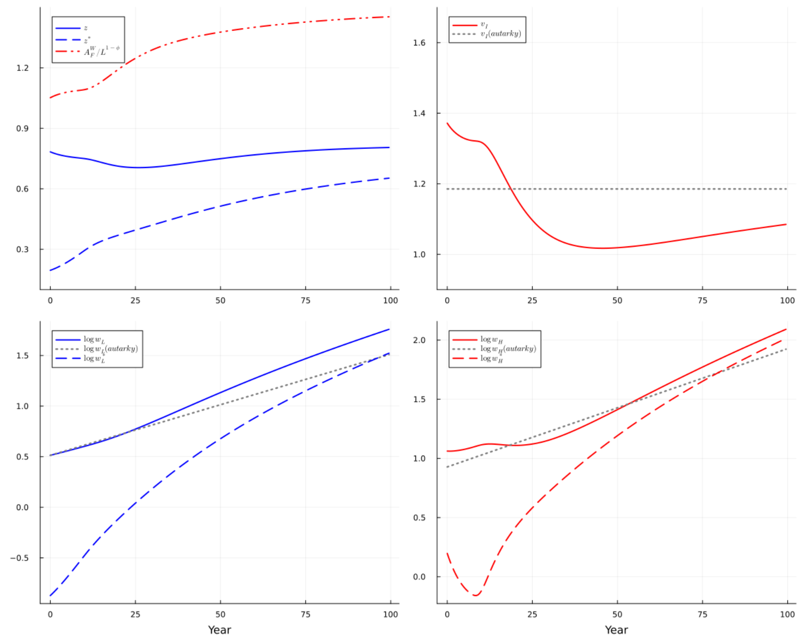


The overall weak wage growth hides a great deal of heterogeneity across worker types with essentially zero growth for low-skilled workers, and robust growth for high skilled workers. Figure 20 in the appendix shows the evolution of the skill premium and the Gini Index in Germany, both of which shoot up in the mid 1990s, consistent with the model and the impact of market integration on the returns to innovation. This pattern of robust innovative activity, weak productivity growth, and stagnation for unskilled worker wages holds broadly across advanced economies and poses a puzzle for benchmark models of endogenous growth. A model that acknowledges the importance of technology adoption, and considers the impact on globalization on the returns to "local adoption" vis-a-vis "global innovation" naturally accounts for the decoupling of innovation and unskilled worker wage growth. Similarly, fast growth in emerging markets, and rising profits and stock market valuations of multinational companies in advanced economies are perfectly consistent with the increasing value of innovation in a globalization world.

This bleak scenario contrasts with the benevolent implication of symmetric market integration as I have argued before. I next plot the transition dynamics for the case when emerging markets converge in terms of fundamental research productivity and skill endowment to the level of advanced economies. I assume convergence in fundamentals takes 30 years and the convergence dynamics mimic the ones implied by a Solow-model with additional details in the appendix.

Initially, symmetric and asymmetric integration do not look very different initially: it takes time for the emerging market to improve their research productivity and skilled labor supply so in the early stages of convergence the emerging market largely adopts technology, which drives up the return to innovation and raises the skill premium, reducing technology adoption in the West. However, in the long-run it becomes apparent that the impact on unskilled labor is fundamentally different. Innovation abroad pushes out the technological frontier, while simultaneously pushing skilled labor back into domestic technology adoption. This is good news for unskilled workers, and even skilled workers are better off in this scenario as the fall in the skill premium is dominated by overall technological growth.

Figure 10. Wages & Technology along the Transition Path



Note. Based on symmetric integration where convergence in innovative capacity and skilled labor endowment takes roughly 30 years, see appendix. Wages are plotted in logs, and the frontier level of technology is normalized by $L^{1-\phi}$ to net out exogenous long-run growth. Grey lines indicate the evolution of the variable in autarky from the point of view of advanced economies.

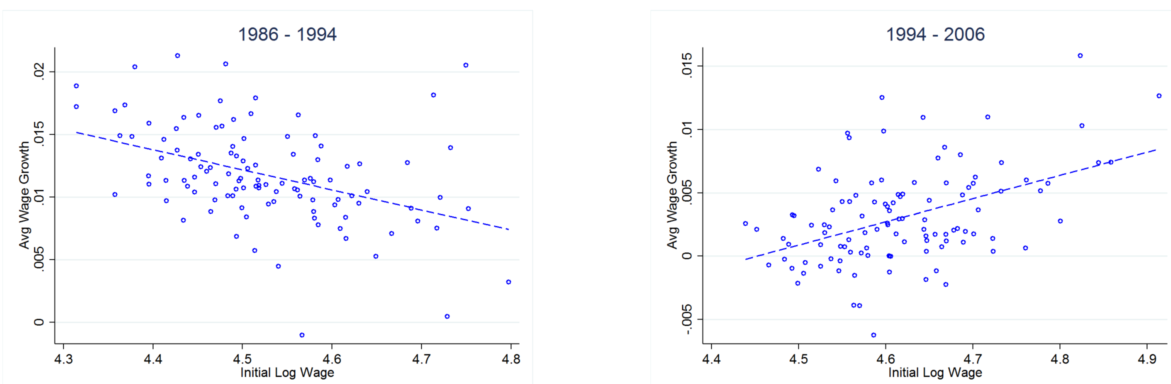
5 Discussion & Conclusion

The model is consistent with a number of empirical facts that have so far only been considered in isolation: rising innovative effort and an increasing skill premium against the backdrop of sluggish productivity growth and overall wage stagnation in advanced economies. The key mechanism that accounts for the puzzling evolution of these secular trends is that innovation in a globalized world potentially comes at the cost of weakened local technology adoption. While a clean causal identification strategy for a theory of global growth seems out of reach, there are a number of distinct predictions that are borne out by the model buttressing its plausibility.

Empirical Evidence. The model has distinct implications for the uneven impact of globalization on innovation and technology adoption. The work of Andrews, Criscuolo, and Gal (2015) and Andrews, Criscuolo, and Gal (2016) seems largely in line with this interpretation as laggard firms, here firms in the production sector, seem to be losing out. Anzoategui et al. (2019) use survey data on the diffusion of specific technologies in the US and the UK. The latter study is consistent with the important role of technology adoption in accounting for the growth slowdown although their focus is ultimately transitory and driven by cyclical variation, in contrast to my framework where a globalization-induced increase in the skill premium creates persistent level effects.

The model also has distinct implications for regional growth patterns within advanced economies based on the logic that specialization in innovation vis-a-vis production is extremely uneven across space: if high-income regions have a persistent advantage in innovation, then market integration ought to further raise growth in innovative regions within advanced economies. This effect would be accompanied by a reallocation of skilled labor away from relatively poorer regions and coincide with weak growth in production-focused regions as technology stalls. I make this argument more carefully in the working paper version (Trouvain, 2023), and focus on the simple cross-regional growth patterns for Germany in this version of the paper. Figure 11 plots average wage growth, defined as the total wage bill of full-time employees over total full-time employment, against the log of the initial average real wage for a local labor market in West Germany.[55] Wage growth in the early period from 1986 – 1994 was, on average, higher for laggard regions. These growth patterns are turned upside down in the 2000s, where high-income places grew relatively fast while laggard regions stagnated.

Figure 11. Regional Convergence in Germany



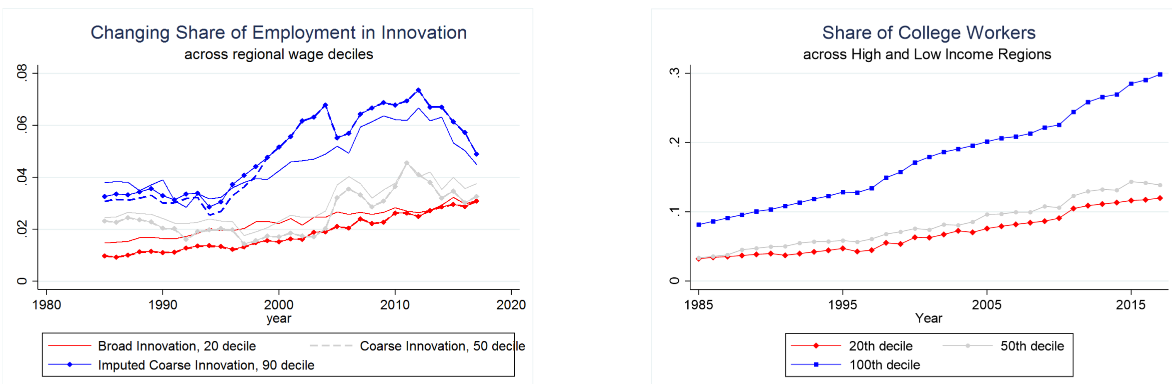
Using data from the BHP establishment sample, the figure plots average wage growth against initial the initial average wage in real terms. The plot shows how growth pre 1994 was biased towards lagging regions, while from 1994 onwards growth was biased towards high income regions. I stop short of the financial crisis, but have looked at convergence patterns from 2006 - 2015 as well which are mostly neutral with a regression coefficient statistically indistinguishable from zero at standard levels of significance. See the appendix for plots for high, middle, and low skilled wages separately. A common concern is that international trade, and in particular import exposure following Autor, Dorn, and Hanson (2013), fully explains weak growth in laggard regions. To consider the effect of import exposure on wage growth, I run a convergence regression with an additional import exposure variable as control. Import competition accounts for virtually none of the stagnation in laggard regions.

High income regions are more specialized in innovation, and the changing growth patterns are

[55] Clearly, German integration poses econometric challenges. However, note that internal integration, at least through the lens of the model, does not generate a bias towards innovation as skilled labor would earn a high return in Eastern Germany to help adopt technology. This appears consistent with the timing of both the rise in inequality and the productivity slowdown that takes hold in the mid 90s and early 2000s.

consistent with rising returns to innovation in the aftermath of global market integration. Figure 12 highlights both the higher initial specialization in innovation, and the further divergence unfolding in the 2000s. Moreover, a reallocation of skilled labor from production to innovation intensive regions occurs as can be seen in the right panel where the skill share of innovative regions diverges from laggard regions. These changing regional convergence patterns are more broadly true across advanced economies including the USA (Berry and Glaeser, 2005; Ganong and Shoag, 2017; Giannone, 2017; Rubinton, 2020), and thus should be uncontroversial. The important difference to this spatially-focused literature is that their explanations are based on skill-biased technological change, which is a force for faster growth. In contrast, in a model with endogenous technology adoption, a meaningful tradeoff emerges. The general equilibrium structure of the model makes clear that the acceleration in skill growth in innovative centers comes at the cost of production-focused regions, which could appropriately be described as brain drain.

Figure 12. Share of College Workers across Regions



These plots compute skill share and employment in innovation across high and low income regions in Germany by grouping regions into wage deciles and computing simple averages. The plots are purely cross-sectional in the sense that I assign labor markets into bins each year so that for example the set of places in the top bin can change every year. In practice, whether one fixed the income ranking in 1994 instead does not change the broad patterns. There is substantial sampling variation within each region, however, and the cross sectional plots is smoother, which is why I prefer it.

Finally, the key mechanism is consistent with and provides a micro-foundation for a number of recent studies: using micro data and a causal estimation design based on cross-regional variation, Lewis (2011), Beaudry, Doms, and Lewis (2010), and Imbert et al. (2022), provide compelling evidence that a change in the local skill mix towards less skilled workers reduces a local labor market's adoption of new technologies.

Alternative Hypotheses. Skill-biased technological change is perhaps the most influential explanations to account for the rise in the skill premium, and wage stagnation of unskilled workers, see Katz and Murphy (1992) and Bound and Johnson (1992). While the evidence in favor of skill-biased technological change is strong, it is difficult to square it with overall wage stagnation because even though skill-biased technical change may favor one worker group over another, it ultimately raises everyone's wage, see Acemoglu and Autor (2011) and Autor (2019) for a discussion of this point. A related literature has focused on the task-content of work and automation, which can generate more adverse effects of technological change on specific worker groups, see Autor, Levy, and Murnane (2003) and Acemoglu and Restrepo (2021). Even so, it is difficult to rationalize overall stagnant wages and slow productivity growth: while automation may reduce wages for some workers, it should raise overall GDP growth.[56] In contrast, in a model with endogenous skill-intensive adoption, rising inequality and weak productivity growth go hand in hand.

Another important explanation for weak aggregate growth is declining population growth. Interestingly, a slowdown in population growth through the lens of a model where ideas are harder to find (Jones, 1995) predicts a declining share of labor devoted to innovation. The fact that ever-more resources are devoted to an activity that is getting harder and harder is puzzling. Taking the role of globalization seriously resolves this tension as technology adoption in emerging markets sustains innovation in advanced economies.

Conclusion. In my calibration, weak domestic technology adoption entirely erases gains from additional innovation in the aftermath of market integration between advanced economies and emerging markets. The mechanism can generate sizable real wage losses for unskilled workers in rich countries, and reconciles weak aggregate growth in advanced economies with rising innovative efforts.

Nonetheless, openness and globalization can play a powerful role in sustaining long-run growth due to the inherent non-rivalry of ideas when emerging markets converge fully and start to contribute to the technological frontier. Concerns about the adverse effects of the ability of emerging markets to compete with advanced economies in the research sector are misplaced through the lens of the model: innovation in emerging markets would push out the technological frontier, and simultaneously induce a reallocation of skilled labor toward adoption activity within advanced economies. This, in turn, would generate broad-based wage growth across worker groups.

References

Acemoglu, Daron (2002). "Directed technical change". In: *The review of economic studies* 69.4, pp. 781–809.

— (2003). "Patterns of skill premia". In: *The Review of Economic Studies* 70.2, pp. 199–230.

— (2009). "Introduction to Modern Economic Growth". In.

Acemoglu, Daron, Philippe Aghion, and Fabrizio Zilibotti (2006). "Distance to frontier, selection, and economic growth". In: *Journal of the European Economic association* 4.1, pp. 37–74.

Acemoglu, Daron and David Autor (2011). "Skills, tasks and technologies: Implications for employment and earnings". In: *Handbook of labor economics*. Vol. 4. Elsevier, pp. 1043–1171.

Acemoglu, Daron, Gino Gancia, and Fabrizio Zilibotti (2015). "Offshoring and directed technical change". In: *American Economic Journal: Macroeconomics* 7.3, pp. 84–122.

[56] See Aghion, Jones, and Jones (2017) for the powerful pro-growth effects of automation in an endogenous growth model.

Acemoglu, Daron and Pascual Restrepo (2020). “Robots and jobs: Evidence from US labor markets”. In: *Journal of Political Economy* 128.6, pp. 2188–2244.

— (2021). *Tasks, automation, and the rise in US wage inequality*. Tech. rep. National Bureau of Economic Research.

Acemoglu, Daron and Fabrizio Zilibotti (2001). “Productivity differences”. In: *The Quarterly Journal of Economics* 116.2, pp. 563–606.

Acemoglu, Daron et al. (2018). “Innovation, reallocation, and growth”. In: *American Economic Review* 108.11, pp. 3450–91.

Achdou, Yves et al. (2022). “Income and wealth distribution in macroeconomics: A continuous-time approach”. In: *The review of economic studies* 89.1, pp. 45–86.

Aghion, Philippe and Peter Howitt (1990). *A model of growth through creative destruction*. Tech. rep. National Bureau of Economic Research.

Aghion, Philippe, Benjamin F Jones, and Charles I Jones (2017). *Artificial intelligence and economic growth*. Vol. 23928. National Bureau of Economic Research Cambridge, MA.

Aghion, Philippe et al. (2019). *A theory of falling growth and rising rents*. Tech. rep. National Bureau of Economic Research.

Akcigit, Ufuk and Sina T Ates (2019). *What happened to US business dynamism?* Tech. rep. National Bureau of Economic Research.

Akcigit, Ufuk, Sina T Ates, and Giammario Impullitti (2018). *Innovation and trade policy in a globalized world*. Tech. rep. National Bureau of Economic Research.

Alesina, Alberto, Enrico Spolaore, and Romain Wacziarg (2005). “Trade, growth and the size of countries”. In: *Handbook of economic growth*. Vol. 1. Elsevier, pp. 1499–1542.

Allen, Treb and Costas Arkolakis (2014). “Trade and the Topography of the Spatial Economy”. In: *The Quarterly Journal of Economics* 129.3, pp. 1085–1140.

Alvaredo, Facundo et al. (2020). “Distributional national accounts guidelines, methods and concepts used in the world inequality database”. PhD thesis. PSE (Paris School of economics).

Alvarez, Fernando and Robert E Jr Lucas (2007). “General equilibrium analysis of the Eaton–Kortum model of international trade”. In: *Journal of monetary Economics* 54.6, pp. 1726–1768.

Andrews, Dan, Chiara Criscuolo, and Peter N Gal (2015). “Frontier firms, technology diffusion and public policy: Micro evidence from OECD countries”. In.

— (2016). “The best versus the rest: the global productivity slowdown, divergence across firms and the role of public policy”. In.

Anzoategui, Diego et al. (2019). “Endogenous technology adoption and R&D as sources of business cycle persistence”. In: *American Economic Journal: Macroeconomics* 11.3, pp. 67–110.

Arkolakis, Costas, Arnaud Costinot, and Andrés Rodríguez-Clare (2012). “New trade models, same old gains?” In: *American Economic Review* 102.1, pp. 94–130.

Arkolakis, Costas et al. (2018). “Innovation and production in the global economy”. In: *American Economic Review* 108.8, pp. 2128–73.

Atkeson, Andrew and Ariel Tomas Burstein (2010). “Innovation, firm dynamics, and international trade”. In: *Journal of political economy* 118.3, pp. 433–484.

Atkinson, Anthony B. and Joseph E. Stiglitz (1969). “A New View of Technological Change”. In: *The Economic Journal* 79.315, pp. 573–578.

Autor, David, David Dorn, and Gordon H Hanson (2013). “The China syndrome: Local labor market effects of import competition in the United States”. In: *American Economic Review* 103.6, pp. 2121–68.

Autor, David et al. (2020). “The fall of the labor share and the rise of superstar firms”. In: *The Quarterly Journal of Economics* 135.2, pp. 645–709.

Autor, David H (2019). “Work of the Past, Work of the Future”. In: *AEA Papers and Proceedings*. Vol. 109, pp. 1–32.

Autor, David H, Frank Levy, and Richard J Murnane (2003). “The skill content of recent technological change: An empirical exploration”. In: *The Quarterly journal of economics* 118.4, pp. 1279–1333.

Banerjee, Abhijit V and Esther Duflo (2005). “Growth theory through the lens of development economics”. In: *Handbook of economic growth* 1, pp. 473–552.

Barro, Robert J (1991). “Economic growth in a cross section of countries”. In: *The quarterly journal of economics* 106.2, pp. 407–443.

Barro, Robert J and Jong Wha Lee (2013). “A new data set of educational attainment in the world, 1950–2010”. In: *Journal of development economics* 104, pp. 184–198.

Barro, Robert J and Xavier Sala-i Martin (1997). “Technological Diffusion, Convergence, and Growth”. In: *Journal of Economic Growth* 2.1, pp. 1–26.

Basu, Susanto and David N. Weil (1998). “Appropriate Technology and Growth”. In: *The Quarterly Journal of Economics* 113.4, pp. 1025–1054.

Beaudry, Paul, Mark Doms, and Ethan Lewis (2010). “Should the personal computer be considered a technological revolution? Evidence from US metropolitan areas”. In: *Journal of political Economy* 118.5, pp. 988–1036.

Benassy, Jean-Pascal (1996). “Taste for variety and optimum production patterns in monopolistic competition”. In: *Economics Letters* 52.1, pp. 41–47.

Benhabib, Jess, Jesse Perla, and Christopher Tonetti (2021). “Reconciling models of diffusion and innovation: A theory of the productivity distribution and technology frontier”. In: *Econometrica* 89.5, pp. 2261–2301.

Berry, Christopher R and Edward L Glaeser (2005). “The divergence of human capital levels across cities”. In: *Papers in regional science* 84.3, pp. 407–444.

Bloom, Nicholas et al. (2020). “Are ideas getting harder to find?” In: *American Economic Review* 110.4, pp. 1104–44.

Bound, John and George Johnson (1992). “Changes in the structure of wages in the 1980’s: an evaluation of alternative explanations”. In: *The American Economic Review* 82, pp. 371–392.

Brainerd, Elizabeth (1998). “Winners and losers in Russia’s economic transition”. In: *American Economic Review*, pp. 1094–1116.

Broda, Christian and David E Weinstein (2006). “Globalization and the Gains from Variety”. In: *The Quarterly journal of economics* 121.2, pp. 541–585.

Buera, Francisco J and Ezra Oberfield (2020). “The global diffusion of ideas”. In: *Econometrica* 88.1, pp. 83–114.

Buera, Francisco J et al. (2022). “Skill-biased structural change”. In: *The Review of Economic Studies* 89.2, pp. 592–625.

Burstein, Ariel and Javier Cravino (2015). “Measured aggregate gains from international trade”. In: *American Economic Journal: Macroeconomics* 7.2, pp. 181–218.

Burstein, Ariel T and Alexander Monge-Naranjo (2009). “Foreign know-how, firm control, and the income of developing countries”. In: *The Quarterly Journal of Economics* 124.1, pp. 149–195.

Cai, Jie, Nan Li, and Ana Maria Santacreu (2022). “Knowledge diffusion, trade, and innovation across countries and sectors”. In: *American Economic Journal: Macroeconomics* 14.1, pp. 104–45.

Caselli, Francesco (1999). “Technological revolutions”. In: *American economic review* 89.1, pp. 78–102.

Caselli, Francesco and Wilbur John II Coleman (2006). “The world technology frontier”. In: *American Economic Review* 96.3, pp. 499–522.

Cette, Gilbert, John Fernald, and Benoit Mojon (2016). “The pre-Great Recession slowdown in productivity”. In: *European Economic Review* 88, pp. 3–20.

Coffano, Monica and Gianluca Tarasconi (2014). “CRIOS-Patstat database: sources, contents and access rules”. In: *Center for Research on Innovation, Organization and Strategy, CRIOS working paper* 1.

Comin, Diego and Mark Gertler (2006). “Medium-term business cycles”. In: *American Economic Review* 96.3, pp. 523–551.

Comin, Diego and Bart Hobijn (2004). “Cross-country technology adoption: making the theories face the facts”. In: *Journal of monetary Economics* 51.1, pp. 39–83.

— (2010). “An exploration of technology diffusion”. In: *American Economic Review* 100.5, pp. 2031–59.

Comin, Diego and Marti Mestieri (2014). “Technology diffusion: Measurement, causes, and consequences”. In: *Handbook of economic growth*. Vol. 2. Elsevier, pp. 565–622.

Comin, Diego A and Bart Hobijn (2007). *Implementing technology*.

De Ridder, Maarten (2019). “Market power and innovation in the intangible economy”. In.

Decker, Ryan A et al. (2017). “Declining dynamism, allocative efficiency, and the productivity slowdown”. In: *American Economic Review* 107.5, pp. 322–26.

Dixit, Avinash K and Joseph E Stiglitz (1977). “Monopolistic competition and optimum product diversity”. In: *The American economic review* 67.3, pp. 297–308.

Eaton, Jonathan and Samuel Kortum (1999). "International technology diffusion: Theory and measurement". In: *International Economic Review* 40.3, pp. 537–570.

— (2002). "Technology, geography, and trade". In: *Econometrica* 70.5, pp. 1741–1779.

Engbom, Niklas et al. (2019). *Firm and worker dynamics in an aging labor market*. Tech. rep. Federal Reserve Bank of Minneapolis Minneapolis, MN.

Fernald, John G (2015). "Productivity and Potential Output before, during, and after the Great Recession". In: *NBER macroeconomics annual* 29.1, pp. 1–51.

Foster, Andrew D and Mark R Rosenzweig (1996). "Technical change and human-capital returns and investments: evidence from the green revolution". In: *The American economic review*, pp. 931–953.

— (2010). "Microeconomics of technology adoption". In: *Annu. Rev. Econ.* 2.1, pp. 395–424.

Ganong, Peter and Daniel Shoag (2017). "Why has regional income convergence in the US declined?" In: *Journal of Urban Economics* 102, pp. 76–90.

Garcia-Macia, Daniel, Chang-Tai Hsieh, and Peter J Klenow (2019). "How destructive is innovation?" In: *Econometrica* 87.5, pp. 1507–1541.

Giannone, Elisa (2017). "Skilled-biased technical change and regional convergence". In: *University of Chicago. Unpublished manuscript*.

Golder, Peter N and Gerard J Tellis (1997). "Will it ever fly? Modeling the takeoff of really new consumer durables". In: *Marketing Science* 16.3, pp. 256–270.

Goldin, Claudia and Lawrence F Katz (2010). *The race between education and technology*. harvard university press.

Greenwood, Jeremy and Mehmet Yorukoglu (1997). "1974". In: *Carnegie-Rochester conference series on public policy*. Vol. 46. Elsevier, pp. 49–95.

Griliches, Zvi (1957). "Hybrid corn: An exploration in the economics of technological change". In: *Econometrica, Journal of the Econometric Society*, pp. 501–522.

Gross, Till and Paul Klein (2024). *Innovation-driven growth in a multi-country world*.

Grossman, Gene M and Elhanan Helpman (1991a). *Innovation and growth in the global economy*. MIT press.

— (1991b). "Quality ladders in the theory of growth". In: *The review of economic studies* 58.1, pp. 43–61.

Helpman, Elhanan (2016). *Globalization and wage inequality*. Tech. rep. National Bureau of Economic Research.

Hopenhayn, Hugo, Julian Neira, and Rish Singhania (2018). *From population growth to firm demographics: Implications for concentration, entrepreneurship and the labor share*. Tech. rep. National Bureau of Economic Research.

Hopenhayn, Hugo and Liyan Shi (2020). *Knowledge creation and diffusion with limited appropriation*. Tech. rep. working paper.

Hopenhayn, Hugo A (1992). "Entry, exit, and firm dynamics in long run equilibrium". In: *Econometrica: Journal of the Econometric Society*, pp. 1127–1150.

Hsieh, Chang-Tai and Peter J Klenow (2014). "The life cycle of plants in India and Mexico". In: *The Quarterly Journal of Economics* 129.3, pp. 1035–1084.

Imbert, Clement et al. (2022). "Migrants and firms: Evidence from china". In: *American Economic Review* 112.6, pp. 1885–1914.

Jones, Charles I (1995). "R & D-based models of economic growth". In: *Journal of political Economy* 103.4, pp. 759–784.

— (2020). *The End of Economic Growth? Unintended Consequences of a Declining Population*. Tech. rep. National Bureau of Economic Research.

Jovanovic, Boyan (1997). "Learning and growth". In: *Advances in Economics*.

Katz, Lawrence F and Kevin M Murphy (1992). "Changes in relative wages, 1963–1987: supply and demand factors". In: *The Quarterly journal of economics* 107.1, pp. 35–78.

Khan, Mosahid and Hélène Dernis (2006). "Global overview of innovative activities from the patent indicators perspective". In.

Klette, Tor Jakob and Samuel Kortum (2004). "Innovating firms and aggregate innovation". In: *Journal of political economy* 112.5, pp. 986–1018.

Konig, Michael et al. (2021). "From Imitation to Innovation: Where Is all that Chinese R&D Going?" In: *Econometrica*.

König, Michael D, Jan Lorenz, and Fabrizio Zilibotti (2016). "Innovation vs. imitation and the evolution of productivity distributions". In: *Theoretical Economics* 11.3, pp. 1053–1102.

Krugman, Paul (1980). "Scale economies, product differentiation, and the pattern of trade". In: *The American Economic Review* 70.5, pp. 950–959.

Lewis, Ethan (2011). "Immigration, skill mix, and capital skill complementarity". In: *The Quarterly Journal of Economics* 126.2, pp. 1029–1069.

Lind, Nelson and Natalia Ramondo (2022). *Global Innovation and Knowledge Diffusion*. Tech. rep. National Bureau of Economic Research.

Lucas, Robert E (2009a). "Trade and the Diffusion of the Industrial Revolution". In: *American Economic Journal: Macroeconomics* 1.1, pp. 1–25.

Lucas, Robert E Jr (1993). "Making a Miracle". In: *Econometrica: Journal of the Econometric Society*, pp. 251–272.

— (2009b). "Ideas and growth". In: *Economica* 76.301, pp. 1–19.

Lucas, Robert E Jr and Benjamin Moll (2014). "Knowledge growth and the allocation of time". In: *Journal of Political Economy* 122.1, pp. 1–51.

Luttmer, Erzo GJ (2007). "Selection, growth, and the size distribution of firms". In: *The Quarterly Journal of Economics* 122.3, pp. 1103–1144.

Mankiw, N Gregory, David Romer, and David N Weil (1992). "A contribution to the empirics of economic growth". In: *The quarterly journal of economics* 107.2, pp. 407–437.

McGrattan, Ellen R and Edward C Prescott (2010). "Technology capital and the US current account". In: *American Economic Review* 100.4, pp. 1493–1522.

Melitz, Marc J (2003). "The impact of trade on intra-industry reallocations and aggregate industry productivity". In: *econometrica* 71.6, pp. 1695–1725.

Milanovic, Branko (2016). *Global inequality: A new approach for the age of globalization*. Harvard University Press.

Nelson, Richard R and Edmund S Phelps (1966). "Investment in humans, technological diffusion, and economic growth". In: *The American economic review* 56.1/2, pp. 69–75.

Olmstead-Rumsey, Jane (2019). "Market Concentration and the Productivity Slowdown". In.

Parente, Stephen L and Edward C Prescott (1994). "Barriers to technology adoption and development". In: *Journal of political Economy* 102.2, pp. 298–321.

Perla, Jesse and Christopher Tonetti (2014). "Equilibrium imitation and growth". In: *Journal of Political Economy* 122.1, pp. 52–76.

Perla, Jesse, Christopher Tonetti, and Michael E Waugh (2021). "Equilibrium technology diffusion, trade, and growth". In: *American Economic Review* 111.1, pp. 73–128.

Peters, Michael and Conor Walsh (2019). "Declining dynamism, increasing markups and missing growth: The role of the labor force". In: *Increasing Markups and Missing Growth: The Role of the Labor Force (November 6, 2019)*.

— (2021). *Population growth and firm dynamics*. Tech. rep. National Bureau of Economic Research.

Ramondo, Natalia, Andrés Rodríguez-Clare, and Milagro Saborío-Rodríguez (2016). "Trade, domestic frictions, and scale effects". In: *American Economic Review* 106.10, pp. 3159–84.

Rempel, Mark (2021). "Innovation and Firm Dynamics with Patent Wars". In.

Rivera-Batiz, Luis A and Paul M Romer (1991). "Economic integration and endogenous growth". In: *The Quarterly Journal of Economics* 106.2, pp. 531–555.

Romer, Paul M (1990). "Endogenous technological change". In: *Journal of political Economy* 98.5, Part 2, S71–S102.

Rubinton, Hannah (2020). "The Geography of Business Dynamism and Skill Biased Technical Change". In: *FRB St. Louis Working Paper* 2020-020.

Salgado, Sergio (2020). "Technical change and entrepreneurship". In: *Available at SSRN 3616568*.

Sampson, Thomas (2016). "Dynamic selection: an idea flows theory of entry, trade, and growth". In: *The Quarterly Journal of Economics* 131.1, pp. 315–380.

— (2023). "Technology gaps, trade, and income". In: *American Economic Review* 113.2, pp. 472–513.

Schumpeter, Joseph A (1942). *Capitalism, socialism and democracy*. routledge.

Somale, Mariano (2021). "Comparative advantage in innovation and production". In: *American Economic Journal: Macroeconomics* 13.3, pp. 357–96.

Stolper, Wolfgang F and Paul A Samuelson (1941). "Protection and real wages". In: *The Review of Economic Studies* 9.1, pp. 58–73.

Trouvain, Florian (2023). “Technology adoption, innovation, and inequality in a global world”. In.
Trouvain, Florian and Giovanni L Violante (2025). “Investment Specific Technological Change and the Skill Premium”. In: *Work in progress*.
Young, Alwyn (1998). “Growth without scale effects”. In: *Journal of political economy* 106.1, pp. 41–63.

A Theory Appendix

A.1 Production Firm

A.1.1 Static minimization problem of firm in production sector

Optimality can be split into a number of steps, where first I begin by deriving the efficient demand for each capital good, x_z, holding A fixed. Without loss of generality, one can think of the capital goods x_j as contained in the interval $[0, A]$ where $\int_0^A dj = A$. Given total expenditure on capital goods $\int p_j x_j dj = p_j x$ where $\int x_j dj = x$, I can ask how much expenditure is spend on each particular variety. The problem reads

$$\begin{array}{ll} \max & \int_0^A \left(\frac{x_j}{\alpha}\right)^\alpha dj \\ s.t. & \int p_j x_j dj \leq I. \end{array} \tag{43}$$

This well-known problem (Dixit and Stiglitz, 1977) leads to the following first order condition

$$\frac{x_j}{x_z} = \left(\frac{p_j}{p_z}\right)^{-\frac{1}{1-\alpha}},$$

and since the capital goods are homogeneous it follows that $x_j = x_k \, \forall j, k$. As a consequence, the total quantity of each individual capital good variety must read $p_j x_j = \frac{p_x \tilde{x}}{A}$ where the last equality holds because of the symmetry assumption.

Now I can substitute this into the firm production function and find the minimal cost of producing one unity of output, given factor prices. This leads to the following cost minimization problem

$$\begin{array}{ll} \min & wl + p_x \tilde{x} \\ s.t. & \left(\int_0^A \left(\frac{\tilde{x}}{\alpha}\frac{1}{A}\right)^\alpha dj\right)\left(\frac{l}{1-\alpha}\right)^{1-\alpha} \geq 1 \end{array}$$

The problem further simplifies to

$$\begin{array}{ll} \min & wl + p_x \tilde{x} \\ s.t. & \left(\frac{\tilde{x}}{\alpha}\right)^\alpha \left(\frac{Al}{1-\alpha}\right)^{1-\alpha} \geq 1 \end{array}$$

which has the convenient Cobb-Douglas structure with labor-augmenting technological change. The first order conditions lead to the constant ratio of expenditure shares on labor and capital

$$\frac{p_x \tilde{x}}{wl} = \frac{\alpha}{1-\alpha}$$

Together with the binding constraint, $\left(\frac{\tilde{x}}{\alpha}\right)^\alpha \left(\frac{Al}{1-\alpha}\right)^{1-\alpha} = 1$, the cost-minimizing bundle of labor and capital leads to a marginal (and average) unit cost of

$$mc \quad = \quad (p_x)^{\alpha}\left(\tfrac{w}{A}\right)^{1-\alpha} \ .$$

Average and marginal cost coincide since the production function features constant returns in capital and labor, conditional on A.

This constant-marginal cost results is important as it simplifies the firm's price setting problem, taking aggregate variables as given. Formally, the problem reads

$$\max_p \quad Y p^{-\sigma}\left[p - mc\right]$$

which leads to the well-known constant markup over marginal cost,

$$p = \frac{\sigma}{\sigma - 1} mc.$$

This constitutes a solution to the static firm problem. Since profits are strictly decreasing in marginal cost, it is indeed optimal to achieve lowest cost and then charge a constant markup over marginal cost.

A.1.2 Dynamic Firm Problem and Adoption Gap

The firms' adoption problem is summarized in the following HJB equation

$$\begin{aligned} V\left(r_t + \delta_{\mathrm{X}}\right) &= \max_{h_{i,t}} \pi^0\left(A_{i,t}\right) - w_{H,t} h_{i,t} + \dot{A}_{i,t}\partial_A V + \dot{V} \\ \text{s.t.} & \\ \dot{A}_{i,t} &= \nu A_{\mathrm{F},t}^{1-\theta} A_{i,t}^{\theta} h_{i,t}^{\beta} - \delta_{\mathrm{I}} A_{i,t}, \end{aligned} \tag{44}$$

where δ_{X} is an exogenous firm death shock and $\partial_A V := \frac{\partial V}{\partial A}$ denotes the partial derivative

To solve the production firm's adoption problem, it is useful to rewrite the problem using a normalized value function $v = \frac{V}{w_L}$, as well as normalizing the state variable A by A_F, i.e. the state becomes z. With these assumptions, I obtain a system that is stationary in the steady state. In the log utility case with $r = \rho + g_{\mathrm{F}}$ along a balanced growth path, this leads to the following recursive formulation of the firm adoption problem,

$$\begin{aligned} v\left(\rho + \delta_{\mathrm{X}}\right) &= \max_h \tfrac{\pi_t^o(z)}{w_L} - sh + \dot{z}\partial_z v + \dot{v} \\ s.t. & \\ \dot{z} &= \nu z^{\theta} h^{\beta} - \left(g_{\mathrm{F}} + \delta_{\mathrm{I}}\right) z. \end{aligned} \tag{45}$$

A solution to the program (45) needs to satisfy the following first order condition

$$\left\{\frac{\beta\nu z^{\theta}\partial_z v}{s}\right\}^{\frac{1}{1-\beta}} = h\ . \tag{46}$$

Equation (46) captures the tradeoff of the effect on firm value of a marginal increase in h relative to its cost s. In anticipation of the solution, I derive the derivative of h with respect to z and t, which yields

$$\begin{aligned}\frac{\partial_{zz}v}{\partial_z v} + \frac{\theta}{z} &= (1-\beta)\,\frac{\partial_z h}{h}\\ \frac{\partial_z \dot{v}}{\partial_z v} - \frac{\dot{s}}{s} &= (1-\beta)\,\frac{\partial_t h}{h}.\end{aligned}$$

Next, derive the envelope condition of the HJB equation to get

$$\begin{aligned}(\partial_z v)\,(\rho+\delta_{\mathrm{X}}) &= \frac{\partial_z \pi}{w_L} + \dot{z}\,(\partial_{zz}v) + \left(\theta z^{\theta-1}\nu h^{\beta} - (g_{\mathrm{F}}+\delta_{\mathrm{I}})\right)(\partial_z v) + \partial_z \dot{v}\\ (\partial_z v)\,(\rho+\delta_{\mathrm{X}}+(1-\theta)\,(g_{\mathrm{F}}+\delta_{\mathrm{I}})) &= \frac{\partial_z \pi}{w_L} + \dot{z}\,(\partial_{zz}v) + (\partial_z v)\,\frac{\theta}{z}\left(z^{\theta}\nu h^{\beta} - z\,(g_{\mathrm{F}}+\delta_{\mathrm{I}})\right) + (\partial_z v)\left\{(1-\beta)\,\frac{\partial_t h}{h} + \frac{\dot{s}}{s}\right\}\\ (\partial_z v)\,(\rho+\delta_{\mathrm{X}}+(1-\theta)\,(g_{\mathrm{F}}+\delta_{\mathrm{I}})) &= \frac{\partial_z \pi}{w_L} + (\partial_z v)\left(\frac{\partial_{zz}v}{\partial_z v} + \frac{\theta}{z}\right)\dot{z} + (\partial_z v)\left\{(1-\beta)\,\frac{\partial_t h}{h} + \frac{\dot{s}}{s}\right\}\\ \rho+\delta_{\mathrm{X}}+(1-\theta)\,(g_{\mathrm{F}}+\delta_{\mathrm{I}}) &= \frac{\partial_z \pi}{w_L}\frac{1}{\partial_z v} + (1-\beta)\,\frac{\dot{h}}{h} + \frac{\dot{s}}{s}.\end{aligned}$$

Now I can substitute in the first order condition and use the fact that I know the derivative of the profit function to get

$$\begin{aligned}\frac{\dot{h}}{h} &= \frac{1}{1-\beta}\left\{(\rho+\delta_{\mathrm{X}}+(1-\theta)\,(g_{\mathrm{F}}+\delta_{\mathrm{I}})) - \frac{1}{\partial_z v}\left[\frac{\pi^{o}}{w_L}\,\frac{(1-\alpha)(\sigma-1)}{z}\right] - \frac{\dot{s}}{s}\right\}\\ \frac{\dot{h}}{h} &= \frac{1}{1-\beta}\left\{(\rho+\delta_{\mathrm{X}}+(1-\theta)\,(g_{\mathrm{F}}+\delta_{\mathrm{I}})) - \frac{\beta\nu z^{\theta}h^{\beta-1}}{s}\left[\frac{\pi^{o}}{w_L}\,\frac{(1-\alpha)(\sigma-1)}{z}\right] - \frac{\dot{s}}{s}\right\}\end{aligned}$$

Moreover, recall that the law of motion of relative technology reads

$$\frac{\dot{z}}{z} = \nu z^{\theta-1}h^{\beta} - (g_{\mathrm{F}}+\delta_{\mathrm{I}})\,.$$

In the steady state, we have that

$$h^{1-\beta} = \frac{1}{s}\,\frac{\beta(1-\alpha)(\sigma-1)(g_{\mathrm{F}}+\delta_{\mathrm{I}})}{\rho+\delta_{\mathrm{X}}+(1-\theta)(g_{\mathrm{F}}+\delta_{\mathrm{I}})}\left[\frac{\pi^{o}}{w_L}\right]\frac{\nu z^{\theta-1}}{(g_{\mathrm{F}}+\delta_{\mathrm{I}})} \tag{47}$$

$$z^{1-\theta} = \frac{\nu h^{\beta}}{g_{\mathrm{F}}+\delta_{\mathrm{I}}} \tag{48}$$

If we combine these two equations one can see that a constant spending on learning activity follows

$$hs = \frac{\beta(1-\alpha)(\sigma-1)(g_{\mathrm{F}}+\delta_{\mathrm{I}})}{\rho+\delta_{\mathrm{X}}+(1-\theta)(g_{\mathrm{F}}+\delta_{\mathrm{I}})}\left[\frac{\pi^{o}}{w_L}\right]\ .$$

This leads to an inequality that needs to be satisfied for the equilibrium to be well-defined, namely

$$\beta\left(1-\alpha\right)\left(\sigma-1\right) \quad < \quad \tfrac{\rho+\delta_{\mathrm{X}}}{g_{\mathrm{F}}+\delta_{\mathrm{I}}}+\left(1-\theta\right).$$

The left hand side represents the additional benefit of improving your productivity, which combines the diminishing returns in learning (β) with the elasticity of the profit function ($\left(\sigma-1\right)\left(1-\alpha\right)$). The right hand side consist of effective costs in steady state, which is related to effective discounting as well as the advantage of backwardness. The firm needs to take into account that as it climbs up the technological ladder, the pull force introduced through the advantage of backwardness diminishes. This gives rise to an endogenous adoption gap as a function of the relative price of skill. Moreover, climbing up the ladder is costly when discounting is high since the benefits only accrue in the future.[57]

A.1.3 Q-Theory of Investment

A complementary approach to derive the firm problem is to set up a standard current-value Hamiltonian. Written in these terms, an obvious connection to the large literature on investment dynamics in the presence of adjustment frictions emerges.

The current-value Hamiltonian set up in an environment with stationary aggregates ($\dot{s}=0$,constant growth rates and interest rates) reads

$$H=\tilde{\pi}^{o}-sh+q_{t}\left[\nu z^{\theta}h^{\beta}-\left(g_{\mathrm{F}}+\delta_{\mathrm{I}}\right)z\right]$$

with optimality conditions

$$\begin{aligned} H_{h} &= 0 \Rightarrow \\ s &= q_{t}\beta\nu z^{\theta}h^{\beta-1} \end{aligned}$$

$$\begin{aligned} H_{z} &= -\dot{q}_{t}+\left(r-g_{w_{L}}+\delta_{\mathrm{X}}\right)q_{t} \Rightarrow \\ \partial_{z}\tilde{\pi}^{o}+q_{t}\left[\nu\theta z^{\theta-1}h^{\beta}-\left(g_{\mathrm{F}}+\delta_{\mathrm{I}}\right)\right] &= -\dot{q}_{t}+\left(r-g_{w_{L}}+\delta_{\mathrm{X}}\right)q_{t} \Rightarrow \\ \partial_{z}\tilde{\pi}^{o}+q_{t}\left[\theta\frac{\dot{z}}{z}-\left(\left(r-g_{w_{L}}+\delta_{\mathrm{X}}\right)+\left(1-\theta\right)\left(g_{\mathrm{F}}+\delta_{\mathrm{I}}\right)\right)\right] &= -\dot{q}_{t}. \end{aligned}$$

Letting $q_{t}=\partial_{z}v$, it is immediate that this approach coincides with the envelope conditions used before.

Next, define the integrating factor $\zeta_{t}=\theta\log z_{t}-\left[r-g_{w_{L}}+\delta_{\mathrm{X}}+\left(1-\theta\right)\left(g_{\mathrm{F}}+\delta_{\mathrm{I}}\right)\right]t$ where the time

[57] Note that the size of the fixed cost of entering the production sector impacts technology adoption since h_{i} is proportional to firm profits. However, h_{D}, i.e., aggregate normalized demand for skilled labor in the production sector is unrelated to f_{E}. One could break the link between z and f_{E} by assuming that the parameter ν is proportional to the fixed cost $\nu=\nu_{0}\cdot f_{\mathrm{E}}^{-\beta}$.

derivative equals $\dot{\zeta}_t = \theta \frac{\dot{z}}{z} - \left((r - g_{w_L} + \delta_{\mathrm{X}}) + (1-\theta)(g_{\mathrm{F}} + \delta_{\mathrm{I}}) \right)$. Using this integrating factor, I get

$$\begin{aligned}
\dot{q}_t e^{\zeta_t} + e^{\zeta_t} q_t \dot{\zeta}_t &= e^{\zeta_t} \partial_z \tilde{\pi}^o \\
\partial_t q_t e^{\zeta_t} &= e^{\zeta_t} \partial_z \tilde{\pi}^o \Rightarrow \\
\int_{t_0}^{\infty} \partial_t q_t e^{\zeta_t} dt &= \int_{t_0}^{\infty} e^{\zeta_t} \partial_z \tilde{\pi}^o dt \Rightarrow \\
q_{t_0} e^{\zeta_{t_0}} - \underbrace{q_\infty e^{\zeta_\infty}}_{=0} &= \int_{t_0}^{\infty} e^{\zeta_t} \partial_z \tilde{\pi}^o dt
\end{aligned}$$

where the value $q_\infty e^{\zeta_\infty}$ converges to zero since $\lim_{t\to\infty} \zeta_t = -\infty$. Consequently, the proposition in the main body of the paper follows

$$\begin{aligned}
q_t &= \int_t^{\infty} e^{\zeta_s - \zeta_t} \partial_z \tilde{\pi}^o ds \\
&= \int_t^{\infty} e^{-\left[r - g_{w_L} + \delta_{\mathrm{X}} + (1-\theta)(g_{\mathrm{F}} + \delta_{\mathrm{I}})\right](s-t)} \left(\frac{z_s}{z_t} \right)^{\theta} \partial_z \tilde{\pi}^o ds.
\end{aligned}$$

A.1.4 Firm value function off and on the balanced growth path

Suppose that free entry into innovation and production holds. In that case, it must be that $f_{\mathrm{E}} = v(t, z)$. Now the value function solves the HJB

$$(r + \delta_{\mathrm{X}} - g_{w_L}) v = \max_h \tfrac{\pi^o}{w_L} - sh + \dot{z}(\partial_z v) + \dot{v}.$$

This dynamic HJB equation is tied to the free entry condition in a useful way, see the discussion in Peters and Walsh (2019), which I leverage in the next steps. Note that totally differentiating $f_{\mathrm{E}} = v(z,t)$ with respect to time t implies $\dot{z}(\partial_z v) = -\dot{v}$. I use this relationship to simplify the HJB equation where it must be understood that h solves the dynamic adoption problem. Rearranging yields

$$v = \frac{\frac{\pi^o}{w_L} - sh}{r_t + \delta_{\mathrm{X}} - g_{w_L}}$$

where I did not assume anything about the stationarity of any of the variables.

Care must be taken for the case when the free entry condition does not hold. In that case, I can compute the firm value by piecing together the part of the problem where no entry occurs (so I know exactly what the measure of firms is and hence can back out profits and the optimal adoption decision) plus the value when free entry is again binding. This is relevant because entry is going to be responsive to learning activity, which pushes down current profits and might thus command a smaller measure of firms in equilibrium.

A.2 Research Sector

Long-run growth rate. To see why the long-run growth rate equals $\frac{g_L}{1-\phi}$, rearrange the resource constraint in innovation

$$
\begin{aligned}
\dot{A}_{\mathrm{F}} &= \gamma A_{\mathrm{F}}^{\phi}\left(\frac{H_{\mathrm{F}}}{L}\right)^{\lambda-1} H_{\mathrm{F}} - \delta_{\mathrm{I}} A_{\mathrm{F}} \Rightarrow \\
(g_{\mathrm{F}} + \delta_{\mathrm{I}}) &= \gamma A_{\mathrm{F}}^{\phi-1} h_{\mathrm{F}}^{\lambda-1} H_{\mathrm{F}} \\
(g_{\mathrm{F}} + \delta_{\mathrm{I}}) &= \gamma \frac{L}{A_{\mathrm{F}}^{1-\phi}} h_{\mathrm{F}}^{\lambda-1} \frac{H_{\mathrm{F}}}{L} \Rightarrow \\
a_{\mathrm{F}} &= \frac{\gamma h_{\mathrm{F}}^{\lambda}}{g_{\mathrm{F}} + \delta_{\mathrm{I}}}
\end{aligned}
$$

where I used the definitions $a_{\mathrm{F}} = \frac{A_{\mathrm{F}}^{1-\phi}}{L}$ and $h_{\mathrm{F}} = \frac{H_{\mathrm{F}}}{L}$. It is easy to see now that along a balanced growth path with positive population growth and $\phi < 1$, it must be that $g_{\mathrm{F}} = \frac{g_L}{1-\phi}$.[58]

Similarly, the law of motion of normalized ideas follows from noting that by definition

$$
\frac{\dot{a}_{\mathrm{F}}}{a_{\mathrm{F}}} = (1-\phi)\, g_{\mathrm{F}} - g_L
$$

and after substituting out $g_{\mathrm{F}} = \gamma A_{\mathrm{F}}^{\phi-1}\left(\frac{H_{\mathrm{F}}}{L}\right)^{\lambda-1} H_{\mathrm{F}} - \delta_{\mathrm{I}}$,

$$
\begin{aligned}
\frac{\dot{a}_{\mathrm{F}}}{a_{\mathrm{F}}} &= (1-\phi)\left[\gamma A_{\mathrm{F}}^{\phi-1}\left(\frac{H_{\mathrm{F}}}{L}\right)^{\lambda-1} \frac{H_{\mathrm{F}}}{L} L - \delta_{\mathrm{I}}\right] - g_L \\
&= (1-\phi)\left[\frac{\gamma h_{\mathrm{F}}^{\lambda}}{a_{\mathrm{F}}} - \delta_{\mathrm{I}}\right] - g_L \\
&= (1-\phi)\left[\frac{\gamma h_{\mathrm{F}}^{\lambda}}{a_{\mathrm{F}}} - \left(\delta_{\mathrm{I}} + \frac{g_L}{1-\phi}\right)\right]
\end{aligned}
$$

which implies

$$
\dot{a}_{\mathrm{F}} = (1-\phi)\left[\gamma h_{\mathrm{F}}^{\lambda} - a_{\mathrm{F}}\left(\delta_{\mathrm{I}} + \frac{g_L}{1-\phi}\right)\right]
$$

as in the main text.

[58] Note that if I had chosen the congestion force using $f_{\mathrm{R}} = H_{\mathrm{F}}^{1-\lambda}$, the link between aggregate per capita growth and population growth would be $g_{\mathrm{F}} = \frac{\lambda}{1-\phi} g_L$. In an earlier version, I used this exact formulation, which works well in the closed economy. In the open economy, small countries would have a large incentive to specialize in innovation to the extent that this congestion externality is "local" as opposed to global.

Decentralized equilibrium. Assume free entry holds

$$V_{\mathrm{I}} \quad = \quad \frac{f_{\mathrm{R}} w_H}{A_{\mathrm{F}}^{\phi}} \quad . \tag{49}$$

Totally differentiating (49) with respect to time t yields

$$\dot{V}_{\mathrm{I}} = V_{\mathrm{I}} \left((1-\lambda) \left(g_{H_{\mathrm{F}}} - g_L \right) + g_w + g_s - \phi g_{\mathrm{F}} \right) \tag{50}$$

where I used the fact that $f_{\mathrm{R}} = \frac{\left(\frac{H_{\mathrm{F}}}{L}\right)^{1-\lambda}}{\gamma}$.

Recall the expression for the net present value of an idea

$$V_{\mathrm{I}} \quad = \quad \int_{t+\tau}^{\infty} \exp\left(-\int_t^u \left(r_v + \delta_{\mathrm{I}} \right) dv \right) \pi_{\mathrm{I},u} du \quad . \tag{51}$$

Totally differentiating (51) with respect to time yields

$$\dot{V}_{\mathrm{I}} = -\exp\left(-\int_t^{t+\tau} \left(r_v + \delta_{\mathrm{I}} \right) dv \right) \pi_{\mathrm{I},t+\tau} \cdot \left[1 + \dot{\tau}_t \right] + \left(r_t + \delta_{\mathrm{I}} \right) V_{\mathrm{I}}, \tag{52}$$

which expresses the value of an innovation in terms of its properly discounted flow profits, $\exp\left(-\int_t^{t+\tau} \left(r_v + \delta_I \right) dv \right) \pi_{I,t+}$ $[1 + \dot{\tau}_t]$, taking into account changes in the waiting time $\dot{\tau}$, as well as appreciation $\dot{V}_{\mathrm{I}}$. Combining (50) and (52) yields

$$V_{\mathrm{I}} = \frac{\exp\left(-\int_t^{t+\tau} \left(r_v + \delta_{\mathrm{I}} \right) dv \right) \pi_{\mathrm{I},t+\tau} \cdot \left[1 + \dot{\tau}_t \right]}{r_t - g_{w_L} - g_s + \delta_{\mathrm{I}} - (1-\lambda) \left(g_{H_{\mathrm{F}}} - g_L \right) + \phi g_{\mathrm{F}}}$$

which is identical to the expression in the paper.

In anticipation of solving for a balanced growth path equilibrium, I normalize the value function using $\frac{w_L}{A_{\mathrm{F}}^{\phi}}$ as normalizing factor so that $v_{\mathrm{I}} = \frac{V_{\mathrm{I}}}{w_L} A_{\mathrm{F}}^{\phi}$, and by free entry $v_{\mathrm{I}} = \frac{s}{\gamma} h_{\mathrm{F}}^{1-\lambda}$. Using this normalization

$$v_{\mathrm{I}} = \frac{A_{\mathrm{F},t}^{\phi}}{w_{L,t}} \frac{\exp\left(-\int_t^{t+\tau} \left(r_v + \delta_{\mathrm{I}} \right) dv \right)}{r_t - g_{w_L} - g_s + \delta_{\mathrm{I}} - (1-\lambda) \left(g_{H_{\mathrm{F}}} - g_L \right) + \phi g_{\mathrm{F}}} \pi_{\mathrm{I},t+\tau} \left(1 + \dot{\tau}_t \right)$$

$$v_{\mathrm{I}} = \frac{A_{\mathrm{F},t}^{\phi}}{w_{L,t}} \frac{\exp\left(-\int_t^{t+\tau} \left(r_v + \delta_{\mathrm{I}} \right) dv \right)}{r_t - g_{w_L} - g_s + \delta_{\mathrm{I}} - (1-\lambda) \left(g_{H_{\mathrm{F}}} - g_L \right) + \phi g_{\mathrm{F}}} \frac{\alpha w_{L,t+\tau} L_{\mathrm{P},t+\tau}}{A_{\mathrm{F},t+\tau} z_{t+\tau}} \left(1 + \dot{\tau}_t \right)$$

$$v_{\mathrm{I}} = \frac{\exp\left(-\int_t^{t+\tau} \left(r_v + \delta_{\mathrm{I}} - g_{w_L} + \phi g_{\mathrm{F}} \right) dv \right)}{r_t - g_{w_L} - g_s + \delta_{\mathrm{I}} - (1-\lambda) \left(g_{H_{\mathrm{F}}} - g_L \right) + \phi g_{\mathrm{F}}} \frac{\alpha L_{\mathrm{P},t+\tau}}{A_{\mathrm{F},t+\tau}^{1-\phi} z_{t+\tau}} \left(1 + \dot{\tau}_t \right)$$

$$v_{\mathrm{I}} = \frac{\exp\left(-\int_t^{t+\tau} \left(r_v + \delta_{\mathrm{I}} - g_{w_L} + \phi g_{\mathrm{F}} \right) dv \right)}{r_t - g_{w_L} - g_s + \delta_{\mathrm{I}} - (1-\lambda) \left(g_{H_{\mathrm{F}}} - g_L \right) + \phi g_{\mathrm{F}}} \frac{\alpha l_{\mathrm{P},t+\tau}}{a_{\mathrm{F},t+\tau} z_{t+\tau}} \left(1 + \dot{\tau}_t \right)$$

where $h_{\mathrm{F}} := \frac{H_{\mathrm{F}}}{L}$ and $a_{\mathrm{F}} = \frac{A_{\mathrm{F}}^{1-\phi}}{L}$ are normalized variables that are constant along a balanced growth path.

Given log utility, the real rate equals $r = \rho + g_{\mathrm{F}}$. Along a balanced growth path, v_I simplifies to

$$\begin{aligned} v_{\mathrm{I}} &= \frac{\exp\left(-\int_t^{t+\tau}\left(\tilde{\rho}+g_{\mathrm{F}}+\delta_{\mathrm{I}}\right)dv\right)}{\tilde{\rho}+g_{\mathrm{F}}+\delta_{\mathrm{I}}}\frac{\alpha l_{\mathrm{P}}}{a_{\mathrm{F}}z} \\ &= \frac{\exp\left(\frac{\tilde{\rho}+g_{\mathrm{F}}+\delta_{\mathrm{I}}}{g_{\mathrm{F}}+\delta_{\mathrm{I}}}\log z\right)}{\tilde{\rho}+g_{\mathrm{F}}+\delta_{\mathrm{I}}}\frac{\alpha l_{\mathrm{P}}}{a_{\mathrm{F}}z} \\ &= \frac{1}{\tilde{\rho}+g_{\mathrm{F}}+\delta_{\mathrm{I}}}\frac{\alpha l_{\mathrm{P}} z^{\frac{\tilde{\rho}}{g_{\mathrm{F}}+\delta_{\mathrm{I}}}}}{a_{\mathrm{F}}} \end{aligned}$$

where I used $\dot{s} = \dot{\tau} = 0$, and $g_{W_L} = g_A = g_{\mathrm{F}} = \frac{g_L}{1-\phi}$, and the definition $\tilde{\rho} = \rho - g_L$.

Using free entry and in particular $\frac{s}{\gamma}h_{\mathrm{F}}^{1-\lambda} = v_{\mathrm{I}}$ I can derive normalized equilibrium demand for skilled labor

$$h_{\mathrm{F}} = \left\{\frac{\gamma}{s}\frac{1}{\tilde{\rho}+g_{\mathrm{F}}+\delta_{\mathrm{I}}}\frac{\alpha l_{\mathrm{P}} z^{\frac{\tilde{\rho}}{g_{\mathrm{F}}+\delta_{\mathrm{I}}}}}{a_{\mathrm{F}}}\right\}^{\frac{1}{1-\lambda}}.$$

Combining this with the resource constraint for ideas in the steady state, $a_{\mathrm{F}} = \frac{\gamma h_{\mathrm{F}}^{\lambda}}{g_{\mathrm{F}}+\delta_{\mathrm{I}}}$, leads to

$$h_{\mathrm{F}} = \left\{\frac{1}{s}\frac{g_{\mathrm{F}}+\delta_{\mathrm{I}}}{\tilde{\rho}+g_{\mathrm{F}}+\delta_{\mathrm{I}}}\alpha l_{\mathrm{P}} z^{\frac{\tilde{\rho}}{g_{\mathrm{F}}+\delta_{\mathrm{I}}}}\right\},$$

which is the same expression as in the main part of the paper.

A.2.1 Waiting time for innovator

The waiting time τ for an innovator's idea to be adopted can be derived as follows. Recall equation (15). Use an integrating factor and note that on the balanced growth path with a constant adoption gap, $g_A = g_F$. Normalizing the time of entry to zero ($t_{\mathrm{E}} = 0$) so that calendar time t coincides with waiting time, I can derive the waiting time as a solution to the following differential equation

$$\dot{W}_t = -\delta_{\mathrm{I}} W_t - A_t\left(g_A + \delta_{\mathrm{I}}\right).$$

Using an integrating factor $e^{\delta_{\mathrm{I}}t}$

$$\begin{aligned}
\int_0^t \frac{\partial \exp(\delta_{\mathrm{I}}u)W_u}{\partial u} &= -\int_0^t \exp\left(\delta_{\mathrm{I}}t\right) A_u \left(\delta_{\mathrm{I}} + g_A\right) du \\
\exp\left(\delta_{\mathrm{I}}t\right) W_t - W_0 &= -A_0 \int_0^t e^{\int_0^u g_A(x)+\delta_{\mathrm{I}}dx} \left(\delta_{\mathrm{I}} + g_A\right) du \\
\exp\left(\delta_{\mathrm{I}}t\right) W_t - W_0 &= -A_0 \int_0^t e^{\int_0^u \zeta_x dx} \left(\zeta_u\right) du \\
\exp\left(\delta_{\mathrm{I}}t\right) W_t - W_0 &= -A_0 \left(e^{\int_0^t g_A(x)+\delta_{\mathrm{I}}dx} - 1\right) \\
W_t - W_0 \exp\left(-\delta_{\mathrm{I}}t\right) &= A_0 \exp\left(-\delta_{\mathrm{I}}t\right) - A_t \\
W_t - \left(A_{\mathrm{F},0} - A_0\right) \exp\left(-\delta_{\mathrm{I}}t\right) &= A_0 \exp\left(-\delta_{\mathrm{I}}t\right) - A_t \\
W_t - \left(A_{\mathrm{F},0}\right) \exp\left(-\delta_{\mathrm{I}}t\right) &= -A_t \\
\left(A_{\mathrm{F},0}\right) \exp\left(-\delta_{\mathrm{I}}t\right) - A_0 e^{\int_0^t g_A(u)du} &= W_t \\
\left(A_{\mathrm{F},0}\right) \exp\left(-\delta_{\mathrm{I}}t\right) \left(1 - z_0 e^{\int_0^t g_A(u)+\delta_{\mathrm{I}}du}\right) &= W_t.
\end{aligned}$$

Now set $W(0,t) = 0$, which is the point in time when the waiting time is zero. From the previous derivations it is clear that this happens when

$$1 = z_0 e^{\int_0^t g_A(u)+\delta_{\mathrm{I}}du}, \tag{53}$$

i.e., the waiting time is implicitly defined by (53). Taking logs and rearranging yields

$$-\frac{\log z_0}{\frac{\int_0^t g_A(u)+\delta_{\mathrm{I}}du}{t}} = t.$$

In general, instead of starting at time zero, entering cohorts start in calendar time t and hit the market at $t + \tau_t$, so the expression generalizes to

$$\tau = -\frac{\log z_t}{\frac{\int_t^{t+\tau} g_A(u)+\delta_{\mathrm{I}}du}{\tau}}, \tag{54}$$

in line with the claim in the main text.

This waiting time is an endogenous object that depends on inventors' and adopters' choices, not just in t but also in periods going forward. In the steady state, however, the expression collapses to a simple statistic

$$\tau = -\frac{\log z}{g_A + \delta_{\mathrm{I}}}.$$

Next, I derive the time derivative $\dot{\tau}$ which is important to compute transition dynamics. Note that (54) implies

$$\tau\delta_{\mathrm{I}} = \log A_{\mathrm{F},t} - \log A_{t+\tau}$$

I totally differentiate this expression to obtain

$$\begin{aligned}
d\tau\delta_{\mathrm{I}} &= g_{\mathrm{F}}(t)\,dt - g_A(t+\tau)(dt+d\tau) \Rightarrow \\
\frac{d\tau}{dt} &= \frac{g_{\mathrm{F}}(t) - g_A(t+\tau)}{g_A(t+\tau)+\delta_{\mathrm{I}}}.
\end{aligned}$$

One might be concerned whether $1+\tau' > 0$, i.e., if $\tau' > -1$. To see that this concern is immaterial take account of the following fact. Because $\beta < 1$, the marginal benefit of adopting technology are infinite so that I can safely assume $g_A > -\delta_{\mathrm{I}}$ (if there was no adoption the two would be equal). Then, I prove by contradiction that $\tau' < -1$ cannot be the case. Suppose $\tau' < -1$. Then,

$$\begin{aligned}
\frac{g_{\mathrm{F}} - g_A(t+\tau)}{\delta_{\mathrm{I}} + g_A(t+\tau)} &< -1 && \Rightarrow \\
g_{\mathrm{F}} - g_A(t+\tau) &< -\delta_{\mathrm{I}} - g_A(t+\tau) && \Rightarrow \\
g_{\mathrm{F}} &< -\delta_{\mathrm{I}}.
\end{aligned}$$

However, even if there is no research effort whatsoever, the worst frontier growth rate must be at least as high as $-\delta_{\mathrm{I}}$. Thus by contradiction it must be that $\dot{\tau} > -1$.

A.3 Constrained Planner Problem

The allocation of skilled labor will be generically inefficient, and the inefficiency is central to understand the response of the economy to globalization. I derive the socially optimal allocation of skilled labor where the only margin the social planner decides on is the allocation of skilled labor between innovation and adoption taking endogenous firm entry in the production sector as given.[59]

The planner solves

$$\max_{\{H_{\mathrm{F},t},c_t\}_{t\geq 0}} \int_0^\infty e^{-\tilde{\rho}t}\log(c_t)\,dt,$$

subject to the following constraints

$$\begin{aligned}
Y &= \left(\frac{AL_{\mathrm{P}}}{1-\alpha}\right)^{1-\alpha}\left(\frac{K}{\alpha}\right)^{\alpha} \\
\dot{K} &= Y - C - \delta_k K \\
\dot{A}_{\mathrm{F}} &= \gamma A_{\mathrm{F}}^{\phi} L^{1-\lambda} H_{\mathrm{F}}^{\lambda} \\
\dot{A} &= \nu A^{\theta} A_{\mathrm{F}}^{1-\theta}\left(\frac{H-H_{\mathrm{F}}}{M}\right)^{\beta},
\end{aligned}$$

[59]The model is not well-suited to study this margin. For one, the fact that adoption has to happen on the firm level means there is a strange incentive to curtail firm entry. Moreover, I took out scale effects in the production sector using the productivity shifter $M_t^{-\frac{1}{\sigma-1}}$, which introduces inefficiencies unrelated to the key tradeoff of allocating skilled labor to innovation or adoption.

and $\frac{\dot{L}}{L} = \frac{\dot{H}}{H} = g_L$.

The constrained efficient allocation of skilled labor across innovation and adoption equals

$$\left(\frac{h_{\mathrm{D}}}{h_{\mathrm{F}}}\right)^{\mathrm{SP}} = \frac{\beta}{1-\theta}\frac{1}{\lambda}\left[\frac{\tilde{\rho}}{g_{\mathrm{F}}} + (1-\phi)\right], \tag{55}$$

where SP stands for social planner. Recall the decentralized allocation (DC)

$$\left(\frac{h_{\mathrm{D}}}{h_{\mathrm{F}}}\right)^{\mathrm{DC}} = \frac{\beta}{1-\theta}\frac{1}{\alpha}\left[\frac{\tilde{\rho}}{g_{\mathrm{F}}} + 1\right]\frac{z^{-\frac{\tilde{\rho}}{g_A}}}{1+\frac{\tilde{\rho}+\delta_{\mathrm{X}}+g_L}{(1-\theta)g_{\mathrm{F}}}}, \tag{56}$$

which generically differs from the planner solution implying that the allocation of skilled labor is generically inefficient.

I provide a comprehensive discussion of the discrepancy between planner and decentralized allocation next where I isolate the distortions induced by the innovation and adoption margin separately. To do so, I need to generalize the baseline model to allow for an additional use of skilled labor in production, beyond its role in technology adoption. If there is only innovation and adoption, shutting down one margin makes the allocation trivially efficient. If, on the other hand, there is skilled labor in innovation, adoption, and production, I can shut down the innovation and adoption margin in turn, and study under what conditions the decentralized equilibrium coincides with the planner solution.

A.3.1 Solve Planner Problem

The planner takes M and L_{P} as given, and I can solve for the optimal allocation using a Hamiltonian.

$$H = \max_{H_{\mathrm{F}},C} e^{-(\rho-g_L)t}\log(c) + \mu_K\left[(AL_{\mathrm{P}})^{1-\alpha}K^{\alpha} - C - \delta_k K\right] + \mu_A\left[\nu A^{\theta}A_{\mathrm{F}}^{1-\theta}\left(\frac{h_{\mathrm{D}}}{m}\right)^{\beta} - \delta_{\mathrm{I}}A\right] + \mu_{A_{\mathrm{F}}}\left[\gamma A_{\mathrm{F}}^{\phi}Lh_{\mathrm{F}}^{\lambda} - \delta_{\mathrm{I}}A_{\mathrm{F}}\right]$$

$$
\begin{aligned}
dC: &\qquad \mu_K L = e^{-(\rho-g_L)t}\frac{1}{c} \\
dk: &\qquad -\frac{\dot{\mu}_K}{\mu_K} = \alpha\left(\frac{AL_{\mathrm{P}}}{K}\right)^{1-\alpha} - \delta_k \\
dh_{\mathrm{F}}: &\qquad \frac{\mu_A}{\mu_{A_{\mathrm{F}}}} = \frac{\lambda\gamma A_{\mathrm{F}}^{\phi} L h_{\mathrm{F}}^{\lambda}}{\beta\nu A^{\theta} A_{\mathrm{F}}^{1-\theta}\left(\frac{h_{\mathrm{D}}}{m}\right)^{\beta}\frac{h_{\mathrm{F}}}{h_{\mathrm{D}}}} \\
dA: &\qquad -\frac{\dot{\mu}_A}{\mu_A} = \frac{\mu_K}{\mu_A}\left(\frac{1-\alpha}{A}\right)(AL_{\mathrm{P}})^{1-\alpha}K^{\alpha} + \left[\theta\nu A^{\theta-1}A_{\mathrm{F}}^{1-\theta}\left(\frac{h_{\mathrm{D}}}{m}\right)^{\beta} - \delta_{\mathrm{I}}\right] \\
dA_{\mathrm{F}}: &\qquad -\frac{\dot{\mu}_{A_{\mathrm{F}}}}{\mu_{A_{\mathrm{F}}}} = \frac{\mu_A}{\mu_{A_{\mathrm{F}}}}\left[(1-\theta)\,\nu A^{\theta}A_{\mathrm{F}}^{-\theta}\left(\frac{h_{\mathrm{D}}}{m}\right)^{\beta}\right] + \left[\phi\gamma A_{\mathrm{F}}^{\phi-1}Lh_{\mathrm{F}}^{\lambda} - \delta_{\mathrm{I}}\right]
\end{aligned}
$$

where the equation (dA) implies a link between the marginal utility of an extra unit of capital (μ_K) and the multiplier (μ_A), and after rearranging

$$
-\frac{\dot{\mu}_A}{\mu_A} = \frac{\mu_K L_{\mathrm{P}}}{\mu_A}(1-\alpha)\left(\frac{K}{AL_{\mathrm{P}}}\right)^{\alpha} + \left[\theta\nu A^{\theta-1}A_{\mathrm{F}}^{1-\theta}\left(\frac{h_{\mathrm{D}}}{m}\right)^{\beta} - \delta_{\mathrm{I}}\right]
$$

it becomes clear that in a steady state $\frac{\dot{\mu}_A}{\mu_A} = \frac{\dot{\mu}_K}{\mu_K} + g_L$ since $\left(\frac{K}{AL_{\mathrm{P}}}\right)^{\alpha} + \left[\theta\nu A^{\theta-1}A_{\mathrm{F}}^{1-\theta}\left(\frac{h_{\mathrm{D}}}{m}\right)^{\beta} - \delta_{\mathrm{I}}\right]$ is stationary, and $\frac{\dot{L}_{\mathrm{P}}}{L_{\mathrm{P}}} = g_L$ as the relative share of labor devoted to production is constant.

The usual link between marginal product of capital, discounting, and per capita consumption growth emerges, which pins down the ratio of physical capital to effective labor used for production (as opposed to entry)

$$
\frac{K}{AL_{\mathrm{P}}} = \left(\frac{\alpha}{\rho + g_{\mathrm{F}} + \delta_k}\right)^{\frac{1}{1-\alpha}}.
$$

Note that the planner accumulates more capital than households in the decentralized equilibrium due to the markup externality. Next, note that $\frac{\dot{\mu}_A}{\mu_A} = \frac{\dot{\mu}_{A_{\mathrm{F}}}}{\mu_{A_{\mathrm{F}}}}$. Combining ($dA_F$) with ($dh_{\mathrm{F}}$), and recall

$\tilde{\rho} = \rho - g_L$, yields

$$-\frac{\dot{\mu}_{A_\mathrm{F}}}{\mu_{A_\mathrm{F}}} = \frac{\mu_A}{\mu_{A_\mathrm{F}}}\left[(1-\theta)\,\nu A^\theta A_\mathrm{F}^{-\theta}\left(\frac{h_\mathrm{D}}{m}\right)^\beta\right] + \left[\phi\gamma A_\mathrm{F}^{\phi-1} L h_\mathrm{F}^\lambda - \delta_\mathrm{I}\right]$$

$$\rho - g_L + g_\mathrm{F} = \frac{\mu_A}{\mu_{A_\mathrm{F}}}\left[(1-\theta)\,\nu A^\theta A_\mathrm{F}^{-\theta}\left(\frac{h_\mathrm{D}}{m}\right)^\beta\right] + \left[\phi\gamma A_\mathrm{F}^{\phi-1} L h_\mathrm{F}^\lambda - \delta_\mathrm{I}\right]$$

$$\tilde{\rho} + g_\mathrm{F} = \frac{\lambda\gamma A_\mathrm{F}^\phi L h_\mathrm{F}^\lambda}{\beta\nu A^\theta A_\mathrm{F}^{1-\theta}\left(\frac{h_\mathrm{D}}{m}\right)^\beta \frac{h_\mathrm{F}}{h_\mathrm{D}}}\left[(1-\theta)\,\nu A^\theta A_\mathrm{F}^{-\theta}\left(\frac{h_\mathrm{D}}{m}\right)^\beta\right] + \left[\phi\gamma A_\mathrm{F}^{\phi-1} L h_\mathrm{F}^\lambda - \delta_\mathrm{I}\right]$$

$$\tilde{\rho} + g_\mathrm{F} = \frac{\lambda\gamma A_\mathrm{F}^\phi L h_\mathrm{F}^\lambda}{\beta\nu A^\theta A_\mathrm{F}^{1-\theta}\left(\frac{h_\mathrm{D}}{m}\right)^\beta \frac{h_\mathrm{F}}{h_\mathrm{D}}}\left[(1-\theta)\,\nu A^\theta A_\mathrm{F}^{-\theta}\left(\frac{h_\mathrm{D}}{m}\right)^\beta\right] + \left[\phi\gamma A_\mathrm{F}^{\phi-1} L h_\mathrm{F}^\lambda - \delta_\mathrm{I}\right]$$

$$\tilde{\rho} + g_\mathrm{F} = \frac{\lambda\gamma A_\mathrm{F}^{\phi-1} L h_\mathrm{F}^\lambda}{\beta\frac{h_\mathrm{F}}{h_\mathrm{D}}}(1-\theta) + \left[\phi\gamma A_\mathrm{F}^{\phi-1} L h_\mathrm{F}^\lambda - \delta_\mathrm{I}\right].$$

Using $a_\mathrm{F} = \frac{A_\mathrm{F}^{1-\phi}}{L}$ and in the steady state $a_\mathrm{F} = \frac{\gamma h_\mathrm{F}^\lambda}{g_\mathrm{F}+\delta_\mathrm{I}}$, and $g_\mathrm{F} = \frac{1}{1-\phi} g_L$,

$$\tilde{\rho} + g_\mathrm{F} = \frac{\lambda\left(g_\mathrm{F}+\delta_\mathrm{I}\right)}{\beta\frac{h_\mathrm{F}}{h_\mathrm{D}}}(1-\theta) + \left[\phi\left(g_\mathrm{F}+\delta_\mathrm{I}\right) - \delta_\mathrm{I}\right].$$

Rearranging yields the final result

$$\frac{h_\mathrm{D}}{h_\mathrm{F}} = \frac{\beta}{1-\theta}\cdot\frac{1}{\lambda}\left[\frac{\tilde{\rho}}{g_\mathrm{F}+\delta_\mathrm{I}} + (1-\phi)\right].$$

A.3.2 Planner Problem with Skilled Labor in Production

To study the distinct inefficiencies induced by externalities and markups in each sector, I generalize the baseline mode and assume that the production of intermediate goods also requires skilled labor as a production factor

$$y_i = \left(\int_{j\in\Omega_{A_i}}\left(\frac{x_{ij}}{\alpha}\right)^\alpha dj\right)\left(\frac{l_i^{1-\eta} h_{\mathrm{p},i}^\eta}{1-\alpha}\right)^{1-\alpha}$$

with $\eta \in [0,1)$, and $h_\mathrm{P} := \frac{\int h_{\mathrm{p},i} di}{L}$ analogous to the definition of h_F and h_D.

The planner takes M and L_P as given, and chooses the allocation of skilled labor across sectors, and consumption to take account of the inter-temporal dimension of the problem. The difference to the previous setup is that skilled labor is also used in production so setting the skilled labor share in production equal to zero restores the original. I set up the present value Hamiltonian and solve the following program

$$
\begin{aligned}
H \quad = \max_{H_{\mathrm{F}},C} e^{-\tilde{\rho}t}\log(c) \quad & +\mu_K\left[\left(\frac{l_{\mathrm{P}}^{1-\eta}h_{\mathrm{P}}^{\eta}}{1-\alpha}\right)^{1-\alpha}\left(\frac{K}{AL}\right)^{\alpha}AL - c\cdot L-\delta_k K\right] \\
& +\mu_A\left[\nu A^{\theta}A_{\mathrm{F}}^{1-\theta}\left(\frac{h_{\mathrm{D}}}{m}\right)^{\beta}-\delta_{\mathrm{I}}A\right] \\
& +\mu_{A_{\mathrm{F}}}\left[\gamma A_{\mathrm{F}}^{\phi}Lh_{\mathrm{F}}^{\lambda}-\delta_{\mathrm{I}}A_{\mathrm{F}}\right].
\end{aligned}
$$

The first order conditions read

$$
\begin{aligned}
dc: &\quad \mu_K L = e^{-\tilde{\rho}t}\frac{1}{c} \\
dh: &\quad \mu_k\frac{\eta(1-\alpha)}{h_{\mathrm{p}}}\left(\frac{l_{\mathrm{P}}^{1-\eta}h_{\mathrm{P}}^{\eta}}{1-\alpha}\right)^{1-\alpha}\left(\frac{K}{AL}\right)^{\alpha}AL = \mu_A\frac{\beta\nu A^{\theta}A_{\mathrm{F}}^{1-\theta}}{h_{\mathrm{D}}}\left(\frac{h_{\mathrm{D}}}{m}\right)^{\beta} = \mu_{A_{\mathrm{F}}}\frac{\lambda\gamma A_{\mathrm{F}}^{\phi}Lh_{\mathrm{F}}^{\lambda}}{h_{\mathrm{F}}}
\end{aligned}
$$

and the optimality conditions associated with the co-state variables read

$$
\begin{aligned}
dC: &\quad \mu_K L = e^{-\tilde{\rho}t}\frac{1}{c} \\
dh: &\quad \mu_K\frac{\eta(1-\alpha)}{h_{\mathrm{p}}}\left(\frac{l_{\mathrm{P}}^{1-\eta}h_{\mathrm{P}}^{\eta}}{1-\alpha}\right)^{1-\alpha}\left(\frac{K}{AL}\right)^{\alpha}AL = \mu_A\frac{\beta\nu A^{\theta}A_{\mathrm{F}}^{1-\theta}}{h_{\mathrm{D}}}\left(\frac{h_{\mathrm{D}}}{m}\right)^{\beta} = \mu_{A_{\mathrm{F}}}\frac{\lambda\gamma A_{\mathrm{F}}^{\phi}Lh_{\mathrm{F}}^{\lambda}}{h_{\mathrm{F}}} \\
dk: &\quad -\dot{\mu}_K = \mu_K\left[\alpha\left(\frac{l_{\mathrm{P}}^{1-\eta}h_{\mathrm{P}}^{\eta}}{1-\alpha}\right)^{1-\alpha}k^{\alpha-1}-\delta_k\right] \text{ (with } k:=\frac{K}{AL}\text{)} \\
dA: &\quad -\dot{\mu}_A = \mu_K\left(\frac{1-\alpha}{A}\right)\left(\frac{l_{\mathrm{P}}^{1-\eta}h_{\mathrm{P}}^{\eta}}{1-\alpha}\right)^{1-\alpha}\left(\frac{K}{AL}\right)^{\alpha}AL+\mu_A\left[\theta\nu A^{\theta-1}A_{\mathrm{F}}^{1-\theta}\left(\frac{h_{\mathrm{D}}}{m}\right)^{\beta}-\delta_{\mathrm{I}}\right] \\
dA_{\mathrm{F}}: &\quad -\dot{\mu}_{A_{\mathrm{F}}} = \mu_A\left[(1-\theta)\nu A^{\theta}A_{\mathrm{F}}^{-\theta}\left(\frac{h_{\mathrm{D}}}{m}\right)^{\beta}\right]+\mu_{A_{\mathrm{F}}}\left[\phi\gamma A_{\mathrm{F}}^{\phi-1}Lh_{\mathrm{F}}^{\lambda}-\delta_{\mathrm{I}}\right].
\end{aligned}
$$

Equation (dA) implies a link between the marginal utility of an extra unit of capital (μ_K) and the multiplier (μ_A), and after rearranging

$$-\frac{\dot{\mu}_K}{\mu_K} = \alpha \left(\frac{l_{\mathrm{P}}^{1-\eta} h_{\mathrm{P}}^{\eta}}{1-\alpha} \right)^{1-\alpha} k^{\alpha-1} - \delta_k \tag{57}$$

$$-\frac{\dot{\mu}_A}{\mu_A} = \frac{\mu_K}{\mu_A} (1-\alpha) \left(\frac{l_{\mathrm{P}}^{1-\eta} h_{\mathrm{P}}^{\eta}}{1-\alpha} \right)^{1-\alpha} \left(\frac{K}{AL} \right)^{\alpha} L + \left[\theta \nu A^{\theta-1} A_{\mathrm{F}}^{1-\theta} \left(\frac{h_{\mathrm{D}}}{m} \right)^{\beta} - \delta_{\mathrm{I}} \right] \tag{58}$$

$$-\frac{\dot{\mu}_{A_{\mathrm{F}}}}{\mu_{A_{\mathrm{F}}}} = \frac{\mu_A}{\mu_{A_{\mathrm{F}}}} \left[(1-\theta) \nu A^{\theta} A_{\mathrm{F}}^{-\theta} \left(\frac{h_{\mathrm{D}}}{m} \right)^{\beta} \right] + \left[\phi \gamma A_{\mathrm{F}}^{\phi-1} L h_{\mathrm{F}}^{\lambda} - \delta_{\mathrm{I}} \right]. \tag{59}$$

Inefficient Adoption. Focus first on the case with exogenous innovation, so I can ignore all terms that involve the co-state variable $\mu_{A_{\mathrm{F}}}$. Note that from (dc) I have $-\frac{\dot{\mu}_K}{\mu_K} = \rho + g_{\mathrm{F}}$ along a balanced growth path. Next, note that the only way for $\frac{\dot{\mu}_A}{\mu_A}$ in (58) to be constant is for $\frac{\dot{\mu}_A}{\mu_A} = \frac{\dot{\mu}_K}{\mu_K} + g_L$ to be true. This implies

$$\tilde{\rho} + g_{\mathrm{F}} + \delta_{\mathrm{I}} = \frac{\mu_K}{\mu_A} (1-\alpha) \left(\frac{l_{\mathrm{P}}^{1-\eta} h_{\mathrm{P}}^{\eta}}{1-\alpha} \right)^{1-\alpha} \left(\frac{K}{AL} \right)^{\alpha} L + \theta \nu z^{\theta-1} \left(\frac{h_{\mathrm{D}}}{m} \right)^{\beta}$$

$$\tilde{\rho} + g_{\mathrm{F}} + \delta_{\mathrm{I}} = \frac{\mu_K}{\mu_A} (1-\alpha) \left(\frac{l_{\mathrm{P}}^{1-\eta} h_{\mathrm{P}}^{\eta}}{1-\alpha} \right)^{1-\alpha} \left(\frac{K}{AL} \right)^{\alpha} L + \theta (g_{\mathrm{F}} + \delta_{\mathrm{I}})$$

$$\tilde{\rho} + g_{\mathrm{F}} + \delta_{\mathrm{I}} = \frac{\mu_K}{\mu_A} (1-\alpha) \left(\frac{l_{\mathrm{P}}^{1-\eta} h_{\mathrm{P}}^{\eta}}{1-\alpha} \right)^{1-\alpha} \left(\frac{K}{AL} \right)^{\alpha} L + \theta (g_{\mathrm{F}} + \delta_{\mathrm{I}}) \tag{60}$$

where the second lines uses the link between h_{D} and z in the steady state. Next, note

$$\frac{\mu_k}{\mu_A} \frac{\eta (1-\alpha)}{h_{\mathrm{p}}} \left(\frac{l_{\mathrm{P}}^{1-\eta} h_{\mathrm{P}}^{\eta}}{1-\alpha} \right)^{1-\alpha} \left(\frac{K}{AL} \right)^{\alpha} L = \frac{\beta \nu z^{\theta-1}}{h_{\mathrm{D}}} \left(\frac{h_{\mathrm{D}}}{m} \right)^{\beta}$$

$$\frac{\mu_k}{\mu_A} = \frac{h_{\mathrm{P}}}{h_{\mathrm{D}}} \frac{\beta (g_{\mathrm{F}} + \delta_{\mathrm{I}})}{\eta (1-\alpha) \left(\frac{l_{\mathrm{P}}^{1-\eta} h_{\mathrm{P}}^{\eta}}{1-\alpha} \right)^{1-\alpha} \left(\frac{K}{AL} \right)^{\alpha} L}$$

which can be used in (60) to get

$$\frac{h_{\mathrm{D}}}{h_{\mathrm{P}}} = \frac{\beta}{1-\theta} \frac{1}{\eta} \left\{ \frac{1}{1 + \frac{\tilde{\rho}}{(1-\theta)(g_{\mathrm{F}} + \delta_{\mathrm{I}})}} \right\}$$

which characterizes the efficient allocation of skilled labor in the case without innovation but with skilled labor as a factor of production in the intermediates good sector.

To obtain the decentralized solution, note that a fraction of total output is spend on adoption and a constant fraction is spent on skilled labor in production

$$\begin{aligned} H_{\mathrm{D}} w_H &= \frac{Y}{\sigma} \frac{\beta(\sigma-1)(1-\alpha)(g_{\mathrm{F}}+\delta_{\mathrm{I}})}{\rho+\delta_{\mathrm{X}}+(1-\theta)(g_{\mathrm{F}}+\delta_{\mathrm{I}})} \\ H_{\mathrm{P}} w_H &= Y \frac{\sigma-1}{\sigma}(1-\alpha)\eta \end{aligned}$$

and the ratio equals

$$\left(\frac{h_{\mathrm{D}}}{h_{\mathrm{P}}}\right)^{\mathrm{DC}} = \frac{\beta}{1-\theta}\frac{1}{\eta}\left\{\frac{1}{1+\frac{\tilde{\rho}+\delta_{\mathrm{X}}+g_L}{(1-\theta)g_{\mathrm{F}}}}\right\}. \tag{61}$$

This setup is very close to the one in Parente and Prescott (1994). This expression contrasts with the solution to the constrained planner problem in (62)

$$\left(\frac{h_{\mathrm{D}}}{h_{\mathrm{P}}}\right)^{\mathrm{SP}} = \frac{\beta}{1-\theta}\frac{1}{\eta}\left\{\frac{1}{1+\frac{\tilde{\rho}}{(1-\theta)g_{\mathrm{F}}}}\right\}. \tag{62}$$

Since $\delta_{\mathrm{X}}+g_L>0$, the decentralized allocation suffers from underinvestment in technology adoption. As argued before, this inefficiency is a generic feature of models where incumbent firms make costly adoption choices, and entrants learn from incumbents. Note that I abstract away from any learning spillovers among incumbents, which would amplify this inefficiency.[60] Moreover, the result highlights how the efficiency result in Parente and Prescott (1994) constitutes a knife-edge case hinging on zero population growth ($g_L=0$) and no churn among incumbents ($\delta_{\mathrm{X}}=0$), in which case (61) and (62) agree. If there was firm entry in Parente and Prescott (1994), one would have to assume some form of spillover to tie together the productivity of incumbents and entrants so that a stationary firm size distribution emerges.

Inefficient Innovation. Next, suppose firms are automatically at the technological frontier, so let's set μ_A to zero while $z=1$. This scenario is almost identical to Jones (1995). In that case,

$$\begin{aligned} dh: \quad & \mu_{A_{\mathrm{F}}}\frac{\lambda\gamma A_{\mathrm{F}}^{\phi} L h_{\mathrm{F}}^{\lambda}}{h_{\mathrm{F}}} = \mu_k \frac{\eta(1-\alpha)}{h_{\mathrm{p}}}\left(\frac{l_{\mathrm{P}}^{1-\eta}h_{\mathrm{P}}^{\eta}}{1-\alpha}\right)^{1-\alpha}\left(\frac{K}{AL}\right)^{\alpha} AL \\ dA_F: \quad & -\frac{\dot{\mu}_{A_{\mathrm{F}}}}{\mu_{A_{\mathrm{F}}}} = \frac{\mu_K}{\mu_{A_{\mathrm{F}}}}(1-\alpha)\left(\frac{l_{\mathrm{P}}^{1-\eta}h_{\mathrm{P}}^{\eta}}{1-\alpha}\right)^{1-\alpha}\left(\frac{K}{A_{\mathrm{F}}L}\right)^{\alpha} L + \left[\phi\gamma A_{\mathrm{F}}^{\phi-1} L h_{\mathrm{F}}^{\lambda} - \delta_{\mathrm{I}}\right]. \end{aligned}$$

[60] There is a large literature suggesting that technology adoption effort may be inefficiently low, see for example Foster and Rosenzweig (2010) in the context of agricultural production in developing economies.

The same logic as before implies $\frac{\dot{\mu}_{A_\mathrm{F}}}{\mu_{A_\mathrm{F}}} = \frac{\dot{\mu}_K}{\mu_k} + g_L$ so

$$\tilde{\rho} + g_\mathrm{F} + \delta_\mathrm{I} = \frac{\mu_K}{\mu_{A_\mathrm{F}}}(1-\alpha)\left(\frac{l_\mathrm{P}^{1-\eta}h_\mathrm{P}^{\eta}}{1-\alpha}\right)^{1-\alpha}\left(\frac{K}{A_\mathrm{F}L}\right)^{\alpha} L + \left[\phi\gamma A_\mathrm{F}^{\phi-1}Lh_\mathrm{F}^{\lambda} - \delta_\mathrm{I}\right]$$
$$\tilde{\rho} + g_\mathrm{F} + \delta_\mathrm{I} = \frac{\gamma L}{A_\mathrm{F}^{1-\phi}}h_\mathrm{F}^{\lambda}\left(\lambda\frac{h_\mathrm{P}}{h_\mathrm{F}}\frac{1}{\eta} + \phi\right).$$

Now using $a_\mathrm{F} = \frac{A_\mathrm{F}^{1-\phi}}{L}$ and $\gamma h_\mathrm{F}^{\lambda} = a_\mathrm{F}(g_\mathrm{F} + \delta_\mathrm{I})$ I get the ratio of skilled labor in production vis-a-vis innovation.

$$\frac{h_\mathrm{P}}{h_\mathrm{F}} = \frac{\eta}{\lambda}\left\{\frac{\tilde{\rho}}{g_\mathrm{F} + \delta_\mathrm{I}} + 1 - \phi\right\}.$$

In contrast, the decentralized allocation in the model without adoption equals

$$\frac{h_\mathrm{P}}{h_\mathrm{F}} = \frac{\eta}{\alpha}\left(\frac{\tilde{\rho}}{g_\mathrm{F} + \delta_\mathrm{I}} + 1\right), \tag{63}$$

which can be derived following the step in Jones (1995), and, trivially, features all the same externalities already present in Jones (1995): markup distortions in α leading to too little innovation, negative dynamic knowledge externalities for the case of $\phi < 0$ leading to too much innovation, and a congestion externality for $\lambda < 1$ that also leads to too much innovation, all else equal. For the innovation-production tradeoff considered in (63), Jones concludes that for reasonable parameter values, the economy is characterized by insufficient innovation even when there are negative research externalities with $\phi < 0$ and $\lambda < 1$.

Jones' conclusion is partly driven by the assumption that the production sector is efficient. In a model with technology adoption, as I have shown before, this tradeoff becomes more interesting since technology adoption features externalities as well. It is ex ante unclear weather externalities in innovation or technology adoption dominate.

A final remark before I conclude this welfare analysis relates to the fact that the full model with both technology adoption and innovation features externalities that are absent when considering the simplified model where I isolated each margin separately. To see this, assume $\alpha = \lambda$, $\phi = 0$, and $\delta_\mathrm{X} = -g_L$, which renders innovation and adoption efficient when considered in isolation. In that case, it is still true that planner and decentralized allocation disagree, which can be seen by computing the ratio of the planner and the decentralized allocation

$$\frac{\left(\frac{h_\mathrm{D}}{h_\mathrm{F}}\right)^{\mathrm{SP}}}{\left(\frac{h_\mathrm{D}}{h_\mathrm{F}}\right)^{\mathrm{DC}}} = z^{\frac{\tilde{\rho}}{g_A + \delta_\mathrm{I}}}\left(1 + \frac{\tilde{\rho}}{(1-\theta)(g_\mathrm{F} + \delta_\mathrm{I})}\right), \tag{64}$$

and since $z < 1$ and $1 + \frac{\tilde{\rho}}{(1-\theta)(g_\mathrm{F}+\delta_\mathrm{I})} > 1$ it is unclear if the decentralized allocation features too little or too much adoption. My interpretation of this result is that the advantage of backwardness on the one hand, and the wait-in-line assumption and congestion in bringing ideas to market on the other, induce externalities similar to the search literature. Innovators don't take into account that they make adoption easier. And adopters don't internalize their impact on the waiting time.

A.4 Open Economy

Derivation of the share of ideas originating from country k.

You start with the free entry condition into research, and because the value of innovation is the same no matter where you innovate (because ideas are sold to the same world market frictionlessly), and assuming all countries innovate (which is always true for $\lambda < 1$), I have

$$\begin{aligned} \frac{(h_{\mathrm{F},n})^{1-\lambda} w_{H,n}}{\gamma_n \left(A_{\mathrm{F}}^{\mathrm{W}}\right)^{\phi}} &= \frac{(h_{\mathrm{F},k})^{1-\lambda} w_{H,k}}{\gamma_k \left(A_{\mathrm{F}}^{\mathrm{W}}\right)^{\phi}} \Rightarrow \\ \frac{(h_{\mathrm{F},n})^{1-\lambda} w_{H,n}}{\gamma_n} &= \frac{(h_{\mathrm{F},k})^{1-\lambda} w_{H,k}}{\gamma_k}. \end{aligned} \tag{65}$$

Next, use the fact that by the resource constraint, a link between share of ideas and skilled labor devoted to innovation emerges

$$(g_{\mathrm{F}} + \delta_{\mathrm{I}}) \chi_n a_{\mathrm{F}}^{\mathrm{W}} = \gamma_n h_{\mathrm{F},n}^{\lambda}.$$

Substituting this in (65) and rearranging yields

$$\begin{aligned} \frac{\left(\frac{\chi_n}{\gamma_n}\right)^{\frac{1-\lambda}{\lambda}} w_{H,n}}{\gamma_n} &= \frac{\left(\frac{\chi_k}{\gamma_k}\right)^{\frac{1-\lambda}{\lambda}} w_{H,k}}{\gamma_k} \Rightarrow \\ \frac{\chi_n}{\chi_k} &= \left(\frac{\gamma_n}{\gamma_k}\right)^{\frac{1}{1-\lambda}} \left(\frac{w_{H,n}}{w_{H,k}}\right)^{-\frac{\lambda}{1-\lambda}}. \end{aligned}$$

Next, note that the high-skilled wage ratio is a function of the skill premium alone $\frac{w_{H,k}}{w_{H,n}} = \left(\frac{s_k}{s_n}\right)^{\frac{1-(\beta+\theta)}{1-\theta}}$. Summing over all countries implies

$$\sum_n \frac{\chi_n}{\chi_k} = \sum_n \left(\frac{\gamma_n}{\gamma_k}\right)^{\frac{1}{1-\lambda}} \left(\frac{w_{H,n}}{w_{H,k}}\right)^{-\frac{\lambda}{1-\lambda}}$$

$$\sum_n \frac{\chi_n}{\chi_k} = \sum_n \left(\frac{\gamma_n}{\gamma_k}\right)^{\frac{1}{1-\lambda}} \left(\left(\frac{s_n}{s_k}\right)^{\frac{1-(\beta+\theta)}{1-\theta}}\right)^{-\frac{\lambda}{1-\lambda}}$$

$$\frac{1}{\chi_k} = \frac{\sum_n \gamma_n^{\frac{1}{1-\lambda}} \left(s_n^{\frac{1-(\beta+\theta)}{1-\theta}}\right)^{-\frac{\lambda}{1-\lambda}}}{\gamma_k^{\frac{1}{1-\lambda}} \left(s_k^{\frac{1-(\beta+\theta)}{1-\theta}}\right)^{-\frac{\lambda}{1-\lambda}}} \Rightarrow$$

$$\chi_k = \frac{\gamma_k^{\frac{1}{1-\lambda}} \left(s_k^{\frac{1-(\beta+\theta)}{1-\theta}}\right)^{-\frac{\lambda}{1-\lambda}}}{\sum_n \gamma_n^{\frac{1}{1-\lambda}} \left(s_n^{\frac{1-(\beta+\theta)}{1-\theta}}\right)^{-\frac{\lambda}{1-\lambda}}}.$$

This establishes the result in the main part of the paper.

Derivation of the key market clearing condition in the open economy.

$$\frac{w_{H,k}}{\gamma_k} h_{\mathrm{F},k}^{1-\lambda} = \frac{1}{\tilde{\rho} + g_{\mathrm{F}} + \delta_{\mathrm{I}}} \frac{\alpha L_{\mathrm{P}}}{\left(A_{\mathrm{F}}^{\mathrm{W}}\right)^{1-\phi}} \sum_n w_n z_n^{\frac{\tilde{\rho}}{g_A+\delta_{\mathrm{I}}}}$$

$$\frac{w_{H,k}}{\gamma_k} h_{\mathrm{F},k}^{1-\lambda} = \frac{1}{\tilde{\rho} + g_{\mathrm{F}} + \delta_{\mathrm{I}}} \frac{\alpha l_{\mathrm{P}}}{\left(a_{\mathrm{F}}^{\mathrm{W}}\right)} \sum_n w_n z_n^{\frac{\tilde{\rho}}{g_A+\delta_{\mathrm{I}}}}$$

$$\frac{w_{H,k}}{\gamma_k} h_{\mathrm{F},k}^{1-\lambda} = \frac{g_{\mathrm{F}} + \delta_{\mathrm{I}}}{\tilde{\rho} + g_{\mathrm{F}} + \delta_{\mathrm{I}}} \frac{\alpha l_{\mathrm{P}}}{\gamma_k h_{\mathrm{F},k}^{\lambda}} \chi_k \sum_n w_n z_n^{\frac{\tilde{\rho}}{g_A+\delta_{\mathrm{I}}}}$$

$$z_k s_k h_{\mathrm{F}} = \frac{g_{\mathrm{F}} + \delta_{\mathrm{I}}}{\tilde{\rho} + g_{\mathrm{F}} + \delta_{\mathrm{I}}} \alpha l_{\mathrm{P}} \chi_k \sum_n z_n^{1+\frac{\tilde{\rho}}{g_A+\delta_{\mathrm{I}}}}$$

$$z_k s_k h_{\mathrm{F},k} = \frac{g_{\mathrm{F}} + \delta_{\mathrm{I}}}{\tilde{\rho} + g_{\mathrm{F}} + \delta_{\mathrm{I}}} \frac{\gamma_k^{\frac{1}{1-\lambda}} s_k^{-\frac{1-\beta-\theta}{1-\theta}\frac{\lambda}{1-\lambda}}}{\sum_n \gamma_n^{\frac{1}{1-\lambda}} s_n^{-\frac{1-\beta-\theta}{1-\theta}\frac{\lambda}{1-\lambda}}} \alpha l_{\mathrm{P}} \sum_n z_n^{1+\frac{\tilde{\rho}}{g_A+\delta_{\mathrm{I}}}}.$$

Note that $z = s^{-\frac{\beta}{1-\theta}} \kappa_z$, same as in the closed economy, and collect terms that are constant along a balanced growth path so

$$h_{\mathrm{F},k} = \Lambda_{\mathrm{FO}} \frac{\gamma_k^{\frac{1}{1-\lambda}} s_k^{-\frac{1-\beta-\theta}{1-\theta}\frac{1}{1-\lambda}}}{\sum_n \gamma_n^{\frac{1}{1-\lambda}} s_n^{-\frac{1-\beta-\theta}{1-\theta}\frac{\lambda}{1-\lambda}}} \sum_n z_n^{1+\frac{\tilde{\rho}}{g_A+\delta_{\mathrm{I}}}}.$$

Combine this with the market clearing condition

$$h_{\text{tot},k} = \frac{\Lambda_{\text{D}}}{s_k} + \Lambda_{\text{FO}} \frac{\gamma_k^{\frac{1}{1-\lambda}} s_k^{-\frac{1-\beta-\theta}{1-\theta}\frac{1}{1-\lambda}}}{\sum_n \gamma_n^{\frac{1}{1-\lambda}} s_n^{-\frac{1-\beta-\theta}{1-\theta}\frac{\lambda}{1-\lambda}}} \sum_n z_n^{1+\frac{\tilde{\rho}}{g_A+\delta_{\text{I}}}}.$$

Generalized ACR Formula. Recall the expression for real wages in the open economy relative to the closed economy reads

$$\frac{w_k^{open}}{w_k^{closed}} = \left(\frac{h_{\text{F},k}^{open}}{h_{\text{F},k}^{closed}}\right)^{\frac{\lambda}{1-\phi}} \left(\frac{1}{\chi_k}\right)^{\frac{1}{1-\phi}} \cdot \left(\frac{s_k^{open}}{s_k^{closed}}\right)^{-\frac{\beta}{1-\theta}}$$

and

$$\frac{w_H^{open}}{w_H^{closed}} = \left(\frac{h_{\text{F},k}^{open}}{h_{\text{F},k}^{closed}}\right)^{\frac{\lambda}{1-\phi}} \left(\frac{1}{\chi_k}\right)^{\frac{1}{1-\phi}} \cdot \left(\frac{s_k^{open}}{s_k^{closed}}\right)^{\frac{1-(\beta+\theta)}{1-\theta}}.$$

To derive this expression, start with the resource constraint which implies a link between equilibrium research effort and the share of ideas invented by country k

$$a_{\text{F}}^{\text{W}} = \frac{\gamma_k h_{\text{F},k}^{\lambda}}{(g_{\text{F}} + \delta_{\text{I}})\,\chi_k}.$$

Now the ratio of frontier technology in the open and closed economy is given by $\frac{A_{\text{F}}^{\text{W},open}}{A_{\text{F}}^{\text{W},closed}} = \left(\frac{a_{\text{F}}^{\text{W},open}}{a_{\text{F}}^{\text{W},closed}}\right)^{\frac{1}{1-\phi}} = \left(\frac{h_{\text{F},k}^{open}}{h_{\text{F},k}^{closed}}\right)^{\frac{\lambda}{1-\phi}} \left(\frac{1}{\chi_k}\right)^{\frac{1}{1-\phi}}$ where I used the fact that $\chi_k^{closed} = 1$.

To study the real wage effects I have to account for the adoption margin since $\frac{w_k^{open}}{w_k^{closed}} = \frac{A_{\text{F}}^{\text{W},open}}{A_{\text{F}}^{\text{W},closed}} \frac{z_k^{open}}{z_k^{closed}}$. Note that $\frac{z_k^{open}}{z_k^{closed}} = \left(\frac{s_k^{open}}{s_k^{closed}}\right)^{-\frac{\beta}{1-\theta}}$, which delivers the result. For skilled wages the skill premium needs to be added, $\frac{w_H^{open}}{w_H^{closed}} = \frac{A_F^{W,open}}{A_F^{W,closed}} \frac{z^{open}}{z^{closed}} \frac{s^{open}}{s^{closed}}$.

These wage ratios reflect the long-run differences in wages after all temporary adjustments have taken place. In particular, since the long-run supply of capital is perfectly elastic, the capital-effective labor ratio is the same in the open and closed economy and is thus netted out in the ratio. This concludes the derivation.

Open Economy with country size differences. I next generalize the framework to allow for countries of heterogeneous sizes, which helps me to take the framework to the data. I use the "home economy" aka the West as baseline for the normalizations. This means that the definition of the normalized world technological frontier is unchanged. Define $b_n = \frac{L_n}{L} \in R^+$ as a weight attached to country c relative to the home economy. You could as well pick $L_{\text{W}} = \sum_n L_n$ or any other normalizing factor that grows as the same rate as the labor force.

The growth of the technological frontier now reads

$$\begin{aligned}
\frac{\dot{A}_{\mathrm{F}}^{\mathrm{W}}}{A_{\mathrm{F}}^{\mathrm{W}}} &= \sum_n \frac{\dot{A}_{\mathrm{F},n}}{A_{\mathrm{F}}^{\mathrm{W}}} \\
&= \sum_n \chi_n \frac{\dot{A}_{\mathrm{F},n}}{A_{\mathrm{F},n}} \\
&= \sum_n \chi_n \left(\frac{\gamma_n}{\chi_n} \left(A_{\mathrm{F}}^{\mathrm{W}}\right)^{\phi-1} L_n h_{\mathrm{F},n}^{\lambda} - \delta_{\mathrm{I}} \right) \\
&= \sum_n \chi_n \left(\frac{\gamma_n}{\chi_n} \frac{1}{a_{\mathrm{F}}^{\mathrm{W}}} h_{\mathrm{F},n}^{\lambda} b_n - \delta_{\mathrm{I}} \right) \\
&= \sum_n \chi_n \left(\frac{\gamma_n}{\chi_n} \frac{h_{\mathrm{F},n}^{\lambda} b_n}{a_{\mathrm{F}}^{\mathrm{W}}} - \delta_{\mathrm{I}} \right) \\
&= \sum_n \frac{\gamma_n h_{\mathrm{F},n}^{\lambda} b_n}{a_{\mathrm{F}}^{\mathrm{W}}} - \delta_{\mathrm{I}}.
\end{aligned}$$

This result of course coincides with the previous result when countries are equal-sized and $b_n = 1$. Consequently,

$$g_{a_{\mathrm{F}}^{\mathrm{W}}} = (1-\phi) \left\{ \sum_n \frac{\gamma_n h_{\mathrm{F},n}^{\lambda} b_n}{a_{\mathrm{F}}^{\mathrm{W}}} - \left(\delta_{\mathrm{I}} + \frac{1}{1-\phi} g_L \right) \right\}$$

and in the steady state

$$a_{\mathrm{F}}^{\mathrm{W}} = \frac{\sum_n \gamma_n h_{\mathrm{F},n}^{\lambda} b_n}{g_{\mathrm{F}} + \delta_{\mathrm{I}}}.$$

By free entry, I have

$$\frac{h_{\mathrm{F},n}^{1-\lambda} w_{H,n}}{\gamma_n} = \frac{h_{\mathrm{F},k}^{1-\lambda} w_{H,k}}{\gamma_k}$$

and by the resource constraint I have

$$\frac{\chi_n}{\chi_k} = \frac{\gamma_n h_{\mathrm{F},n}^{\lambda} b_n}{\gamma_k h_{\mathrm{F},k}^{\lambda} b_k}.$$

Combining free entry and resource constraint delivers

$$\chi_k = \frac{\gamma_k^{\frac{1}{1-\lambda}} s_k^{-\frac{\lambda}{1-\lambda}\frac{1-\theta-\beta}{1-\theta}} b_k}{\sum_n \gamma_n^{\frac{1}{1-\lambda}} s_n^{-\frac{\lambda}{1-\lambda}\frac{1-\theta-\beta}{1-\theta}} b_n}$$

where I used $w_{H,k} \propto A_{\mathrm{F}}^{\mathrm{W}} s^{\frac{1-\theta-\beta}{1-\theta}}$.

I next derive the demand for skilled labor as before, using

$$\frac{s_k}{\gamma_k} h_{\mathrm{F},k}^{1-\lambda} = \frac{1}{w_k} \frac{1}{\tilde{\rho} + g_{\mathrm{F}} + \delta_{\mathrm{I}}} \frac{\alpha}{\left(A_{\mathrm{F}}^{\mathrm{W}}\right)^{1-\phi}} \sum_n L_{\mathrm{P},n} w_n z_n^{\frac{\tilde{\rho}}{g_A + \delta_{\mathrm{I}}}}$$

$$\frac{s_k}{\gamma_k} h_{\mathrm{F},k}^{1-\lambda} = \frac{1}{w_k} \frac{1}{\tilde{\rho} + g_{\mathrm{F}} + \delta_{\mathrm{I}}} \frac{\alpha l_{\mathrm{P}}}{a_{\mathrm{F}}^{\mathrm{W}}} \sum_n w_n z_n^{\frac{\tilde{\rho}}{g_A + \delta_{\mathrm{I}}}} \frac{L_n}{L}$$

$$\frac{s_k z_k}{\gamma_k} h_{\mathrm{F},k}^{1-\lambda} = \frac{1}{\tilde{\rho} + g_{\mathrm{F}} + \delta_{\mathrm{I}}} \frac{\alpha l_{\mathrm{P}}}{a_{\mathrm{F}}^{\mathrm{W}}} \sum_n z_n^{1+\frac{\tilde{\rho}}{g_A + \delta_{\mathrm{I}}}} b_n$$

$$\frac{s_k z_k}{\gamma_k} h_{\mathrm{F},k}^{1-\lambda} = \frac{1}{\tilde{\rho} + g_{\mathrm{F}} + \delta_{\mathrm{I}}} \frac{\alpha l_{\mathrm{P}}}{a_{\mathrm{F}}^{\mathrm{W}}} \sum_n z_n^{1+\frac{\tilde{\rho}}{g_A + \delta_{\mathrm{I}}}} b_n.$$

Next, use $\chi_k = \frac{\gamma_k h_{\mathrm{F},k}^{\lambda} b_k}{a_{\mathrm{F}}^{\mathrm{W}}(g_{\mathrm{F}} + \delta_{\mathrm{I}})}$ to obtain

$$\frac{s_k z_k}{\gamma_k} h_{\mathrm{F},k}^{1-\lambda} = \frac{1}{\tilde{\rho} + g_{\mathrm{F}} + \delta_{\mathrm{I}}} \frac{\alpha l_{\mathrm{P}}}{a_{\mathrm{F}}^{\mathrm{W}}} \sum_n z_n^{1+\frac{\tilde{\rho}}{g_A + \delta_{\mathrm{I}}}} b_n$$

$$\frac{s_k z_k}{\gamma_k} h_{\mathrm{F},k}^{1-\lambda} = \frac{g_{\mathrm{F}} + \delta_{\mathrm{I}}}{\tilde{\rho} + g_{\mathrm{F}} + \delta_{\mathrm{I}}} \frac{\chi_k \alpha l_{\mathrm{P}}}{\gamma_k h_{\mathrm{F},k}^{\lambda} b_k} \sum_n z_n^{1+\frac{\tilde{\rho}}{g_A + \delta_{\mathrm{I}}}} b_n$$

$$s_k z_k h_{\mathrm{F},k}^{1-\lambda} = \frac{g_{\mathrm{F}} + \delta_{\mathrm{I}}}{\tilde{\rho} + g_{\mathrm{F}} + \delta_{\mathrm{I}}} \frac{\chi_k \alpha l_{\mathrm{P}}}{h_{\mathrm{F},k}^{\lambda} b_k} \sum_n z_n^{1+\frac{\tilde{\rho}}{g_A + \delta_{\mathrm{I}}}} b_n$$

$$s_k z_k h_{\mathrm{F},k} = \frac{g_{\mathrm{F}} + \delta_{\mathrm{I}}}{\tilde{\rho} + g_{\mathrm{F}} + \delta_{\mathrm{I}}} \frac{\gamma_k^{\frac{1}{1-\lambda}} s_k^{-\frac{\lambda}{1-\lambda}\frac{1-\theta-\beta}{1-\theta}} b_k}{\sum_n \gamma_n^{\frac{1}{1-\lambda}} s_n^{-\frac{\lambda}{1-\lambda}\frac{1-\theta-\beta}{1-\theta}} b_n} \alpha l_{\mathrm{P}} \frac{1}{b_k} \sum_n z_n^{1+\frac{\tilde{\rho}}{g_A + \delta_{\mathrm{I}}}} b_n$$

$$s_k z_k h_{\mathrm{F},k} = \frac{g_{\mathrm{F}} + \delta_{\mathrm{I}}}{\tilde{\rho} + g_{\mathrm{F}} + \delta_{\mathrm{I}}} \frac{\gamma_k^{\frac{1}{1-\lambda}} s_k^{-\frac{\lambda}{1-\lambda}\frac{1-\theta-\beta}{1-\theta}}}{\sum_n \gamma_n^{\frac{1}{1-\lambda}} s_n^{-\frac{\lambda}{1-\lambda}\frac{1-\theta-\beta}{1-\theta}} b_n} \alpha l_{\mathrm{P}} \sum_n z_n^{1+\frac{\tilde{\rho}}{g_A + \delta_{\mathrm{I}}}} b_n$$

$$\kappa_z s_k^{\frac{1-\theta-\beta}{1-\theta}} h_{\mathrm{F},k} = \frac{g_{\mathrm{F}} + \delta_{\mathrm{I}}}{\tilde{\rho} + g_{\mathrm{F}} + \delta_{\mathrm{I}}} \frac{\gamma_k^{\frac{1}{1-\lambda}} s_k^{-\frac{\lambda}{1-\lambda}\frac{1-\theta-\beta}{1-\theta}}}{\sum_n \gamma_n^{\frac{1}{1-\lambda}} s_n^{-\frac{\lambda}{1-\lambda}\frac{1-\theta-\beta}{1-\theta}} b_n} \alpha l_{\mathrm{P}} \sum_n z_n^{1+\frac{\tilde{\rho}}{g_A + \delta_{\mathrm{I}}}} b_n$$

$$h_{\mathrm{F},k} = \Lambda_{\mathrm{FO}} \frac{g_{\mathrm{F}} + \delta_{\mathrm{I}}}{\tilde{\rho} + g_{\mathrm{F}} + \delta_{\mathrm{I}}} \frac{\gamma_k^{\frac{1}{1-\lambda}} s_k^{-\frac{1}{1-\lambda}\frac{1-\theta-\beta}{1-\theta}}}{\sum_n \gamma_n^{\frac{1}{1-\lambda}} s_n^{-\frac{\lambda}{1-\lambda}\frac{1-\theta-\beta}{1-\theta}} b_n} \alpha l_{\mathrm{P}} \sum_n z_n^{1+\frac{\tilde{\rho}}{g_A + \delta_{\mathrm{I}}}} b_n.$$

Note that for the symmetric country case with all endogenous variables and exogenous productivity shifters identical across countries, the expression would coincide with the solution to the closed economy setup since the term $\sum_n b_n$ appears in both nominator and denominator.

Inserting this expression into the market clearing condition and iterating over a set of skill premia across countries delivers the long-run steady state allocation

$$\frac{\Lambda_{\mathrm{D}}}{s_k} + \Lambda_{\mathrm{FO}} \frac{g_{\mathrm{F}} + \delta_{\mathrm{I}}}{\tilde{\rho} + g_{\mathrm{F}} + \delta_{\mathrm{I}}} \frac{\gamma_k^{\frac{1}{1-\lambda}} s_k^{-\frac{1}{1-\lambda}\frac{1-\theta-\beta}{1-\theta}}}{\sum_n \gamma_n^{\frac{1}{1-\lambda}} s_n^{-\frac{\lambda}{1-\lambda}\frac{1-\theta-\beta}{1-\theta}} b_n} \alpha l_{\mathrm{P}} \sum_n z_n^{1+\frac{\tilde{\rho}}{g_A+\delta_{\mathrm{I}}}} b_n = h_{\mathrm{tot},k}.$$

A.4.1 Trade cost and bargaining in open economy

I will next show how to extend the open economy model to include trade cost and bargaining. These extensions are potentially important in taming the impact of the East on innovation in the West. In a model where innovation happens in the West, and the East is a relatively large country that adopts Western technology, the pull force can become so strong that the skill premium in the West shoots up so much that Western workers are left impoverished. Two reasonable assumptions to avoid this outcome relate to the role of trade cost, and bargaining over innovator rents, which I consider in turn.

Trade cost. Recall the flow profits of innovation in the steady state are proportional to

$$\pi_{\mathrm{I}} \propto \sum_n \frac{\alpha L_n^P w_n z^{\frac{\tilde{\rho}}{g_A+\delta_{\mathrm{I}}}}}{A_F z_n}.$$

One could easily introduce ice berg trade cost from innovator i to technology user c

$$\pi_{\mathrm{I}} \propto \sum_n \frac{1}{\tau_{in}} \frac{\alpha L_n^P w_n z^{\frac{\tilde{\rho}}{g_A+\delta_{\mathrm{I}}}}}{A_F z_n}$$

where $\tau_{in} \geq 1$ and $\tau_{ii} = 1$.

The trade cost is somewhat non standard because the model differs from, for instance, from Krugman (1980) where the trade elasticity $\sigma - 1$ shapes the importance of trade cost. The reason is that the only good that is traded is the royalty associated with innovator profits. The locally optimal markup is applied onto domestic cost, which maximizes total domestic revenues, and there is no way to further raise profits. However, the profits need to be shipped back to the holder of the patents, and trade costs apply so only a fraction $\frac{1}{\tau_{ij}}$ of the value of the idea arrive at home.

One could envision that financial markets can help overcome this trade cost when both countries hold ideas. That is, instead of shipping the royalty back one could simply trade it for the royalty that accrues to foreign firms in the domestic economy. The extent to which this is possible depends on the degree of specialization and would not work if all ideas are held in the West.

Bargaining over royalties. An alternative interpretation for the wedge parameter τ could be microfounded by considering explicit bargaining over innovator profits. For example, one could imagine that the Chinese government could try to extract rents from foreign innovators' ideas. This would help reduce the impact of foreign adoption on innovation in the West. If the (corrupt) government wastes this income, the model coincides exactly with the previous extension. Alternatively, one can rebate the profits to households.

A.4.2 Total Asset Stock in Open Economy

I derive the total value of assets in the open economy, which depends on the value of domestic capital, domestic production sector firms, and global intellectual property. Note that ideas are adopted at different points in time in different countries, which I need to keep track off. I focus on the two-country case, and I focus on the value of an idea, which is the complicated object when constructing the total value of assets held by the home economy. I focus on the case where all intellectual property is held in the home economy (West).

It is easy to see that the value of an innovation $V_{\mathrm{I}}^{\mathrm{W}}$ can be split into home and foreign component

$$V_{\mathrm{I}}^{\mathrm{W}} = V_{\mathrm{I}} + V_{\mathrm{I}}^{*}.$$

Next, note that along a balanced growth path a simple expression for the value of an innovation of ideas that are already adopted, denoted by V_A, obtains

$$\begin{aligned} V_A &= \frac{1}{\rho - g_L + g_{\mathrm{F}} + \delta_{\mathrm{I}}} \frac{\alpha L_{\mathrm{P}} w_t}{A} \\ V_A^{*} &= \frac{1}{\rho - g_L + g_{\mathrm{F}} + \delta_{\mathrm{I}}} \frac{\alpha L_{\mathrm{P}}^{*} w_t^{*}}{A^{*}}. \end{aligned}$$

Conveniently, changes in the real wage for a fixed long-run interest rate have no bearing on the value of an innovation since they directly cancel with the total numbers of adopted ideas in the denominator (recall $w \propto A$).

The forces that do matter are i) changes in the waiting time, and ii) an extensive margin effect in the form of another country using ideas. I focus on the first issue first, and introduce the following notation. Let $V_{\mathrm{I},t}\left(k\right)$ denote the value of an innovation at time t that is k years away from adoption. Clearly, this asset has value, but its value ought to be below ideas that are already in use at time t. Using the HJB equation $\left(r + \delta_{\mathrm{I}}\right) V_{\mathrm{I}} = 0 + \dot{V}_{\mathrm{I}}$, and noting $V_{\mathrm{I},t+k}\left(0\right) = V_{A,t+k}$, I have

$$\begin{aligned} V_{\mathrm{I},t}\left(k\right) &= V_{\mathrm{I},t+k}\left(0\right) e^{-\left(r+\delta_{\mathrm{I}}\right)k} \\ &= V_{A,t+k} e^{-\left(r+\delta_{\mathrm{I}}\right)k} \\ &= V_{A,t} e^{-\left(r+\delta_{\mathrm{I}}-g_L\right)k} \end{aligned}$$

where the last line follows from using $\frac{V_{A,t+k}}{V_{A,t}} = e^{g_L k}$.

Next, note that if we knew the distribution of k, i.e., the waiting time of non-adopted ideas, we could already compute the total value of intellectual property using

$$A_t \cdot V_{A,t} + (A_{\mathrm{F},t} - A_t) \cdot \int_0^{\tau} V_{\mathrm{I},t}(k)\, dF(k) + A_t^* \cdot V_{A,t}^* + (A_{\mathrm{F},t} - A_t^*) \cdot \int_0^{\tau^*} V_{\mathrm{I},t}(k^*)\, dF(k^*)$$

where $dF(k)$ is the appropriate marginal distribution over wait times ranging from zero to τ.

The distribution of k can be derived as follows. Note that the share of ideas close to adoption among all unadopted ideas depends on how much entry there is, which is directly proportional to the gross entry rate $g_{\mathrm{F}} + \delta_{\mathrm{I}}$. In fact, the distribution k follows a truncated exponential distribution

$$f(k) = \frac{(g_{\mathrm{F}} + \delta_{\mathrm{I}})\, e^{(g_{\mathrm{F}}+\delta_{\mathrm{I}})k}}{e^{(g_{\mathrm{F}}+\delta_{\mathrm{I}})\tau} - 1}, k \in [0, \tau],$$

where the distribution in the foreign economy features a different τ. Putting the pieces together, and focusing on the value of innovation in the advanced economy, yields

$$\begin{aligned}
A_t \cdot V_{A,t} + (A_{\mathrm{F},t} - A_t) \cdot \int_0^{\tau} V_{\mathrm{I},t}(k)\, dF(k) &= A_t \cdot V_{A,t}\left(1 + \left(\frac{A_{\mathrm{F},t}}{A_t} - 1\right) \cdot \int_0^{\tau} \frac{V_{\mathrm{I},t}(k)}{V_{A,t}} dF(k)\right) \\
&= A_t \cdot V_{A,t}\left(1 + \left(\frac{1-z}{z}\right) \cdot \int_0^{\tau} \frac{V_{\mathrm{I},t}(k)}{V_{A,t}} dF(k)\right) \\
&= A_t \cdot V_{A,t}\left(1 + \left(\frac{1-z}{z}\right) \cdot \int_0^{\tau} e^{-(r+\delta_{\mathrm{I}}-g_L)k} dF(k)\right) \\
&= A_t \cdot V_{A,t}\left(1 + \left(\frac{1-z}{z}\right) \cdot \int_0^{\tau} e^{-(r+\delta_{\mathrm{I}}-g_L)k} \frac{(g_{\mathrm{F}} + \delta_{\mathrm{I}})\, e^{(g_{\mathrm{F}}+\delta_{\mathrm{I}})k}}{e^{(g_{\mathrm{F}}+\delta_{\mathrm{I}})\tau} - 1} dk\right) \\
&= A_t \cdot V_{A,t}\left(1 + \left(\frac{1-z}{z}\right) \frac{g_{\mathrm{F}} + \delta_{\mathrm{I}}}{e^{(g_{\mathrm{F}}+\delta_{\mathrm{I}})\tau} - 1} \cdot \int_0^{\tau} e^{-(r-g_{\mathrm{F}}-g_L)k} dk\right) \\
&= A_t \cdot V_{A,t}\left(1 + \left(\frac{1-z}{z}\right) \frac{g_{\mathrm{F}} + \delta_{\mathrm{I}}}{e^{(g_{\mathrm{F}}+\delta_{\mathrm{I}})\tau} - 1} \cdot \int_0^{\tau} e^{-(\rho-g_L)k} dk\right) \\
&= A_t \cdot V_{A,t}\left(1 + \left(\frac{1-z}{z}\right) \frac{g_{\mathrm{F}} + \delta_{\mathrm{I}}}{e^{(g_{\mathrm{F}}+\delta_{\mathrm{I}})\tau} - 1} \cdot (\rho - g_L)\left[e^{-(\rho-g_L)k}\right]_{\tau}^{0}\right) \\
&= A_t \cdot V_{A,t}\left(1 + \left(\frac{1-z}{z}\right) \frac{1 - e^{-(\rho-g_L)\tau}}{e^{(g_{\mathrm{F}}+\delta_{\mathrm{I}})\tau} - 1} \cdot (g_{\mathrm{F}} + \delta_{\mathrm{I}})(\rho - g_L)\right) \\
&= A_t \cdot V_{A,t}\left(1 + \left(\frac{1-z}{z}\right) \frac{1 - z^{\frac{\rho-g_L}{g_{\mathrm{F}}+\delta_{\mathrm{I}}}}}{\frac{1-z}{z}} \cdot (g_{\mathrm{F}} + \delta_{\mathrm{I}})(\rho - g_L)\right) \\
&= A_t \cdot V_{A,t}\left(1 + \left(1 - z^{\frac{\rho-g_L}{g_{\mathrm{F}}+\delta_{\mathrm{I}}}}\right) \cdot (g_{\mathrm{F}} + \delta_{\mathrm{I}})(\rho - g_L)\right)
\end{aligned}$$

where I substituted out $\tau = -\frac{\log z}{g_{\mathrm{F}}+\delta_{\mathrm{I}}}$. Now adding the value of intellectual property across both countries yields

$$Tot\_pat\_value = \frac{\alpha L_{\mathrm{P}} w}{\tilde{\rho} + g_{\mathrm{F}} + \delta_{\mathrm{I}}} \left\{ \left(1 + \left(1 - z^{\frac{\rho - g_L}{g_{\mathrm{F}}+\delta_{\mathrm{I}}}}\right) \cdot (g_{\mathrm{F}} + \delta_{\mathrm{I}})(\rho - g_L)\right) + \underbrace{\frac{z^* b^*}{z} \left(1 + \left(1 - z^{*\frac{\rho - g_L}{g_{\mathrm{F}}+\delta_{\mathrm{I}}}}\right) \cdot (g_{\mathrm{F}} + \delta_{\mathrm{I}})(\rho - g_L)\right)}_{\text{value of patents abroad}} \right\}$$

where b^* keeps track of the relative country size $b^* = \frac{L_{\mathrm{P}}^*}{L_{\mathrm{P}}}$. Using this expression it is easy to compare the total value of ideas to the counterfactual autarky equilibrium. In the counterfactual autarky equilibrium, the value of patents abroad would be gone, and I evaluate adoption gap and wages at the counterfactual autarky level.

Next, I add up all assets in the economy and normalize by L to get

$$\begin{aligned} Assets &= \frac{Tot\_pat\_value}{L} + \frac{K}{L} + mV_M \\ &= w\left(\frac{Tot\_pat\_value}{Lw} + \frac{\alpha^2}{1-\alpha}\frac{l_P}{\rho + g_{\mathrm{F}} + \delta_K} + m \cdot f_{\mathrm{E}}\right). \end{aligned}$$

I can construct this expression and compare it conveniently across different counterfactual scenarios where all elements inside the parentheses are stationary. Lastly, note that one can simply apply the factor χ to the value of patents when both countries are innovation, and one can use the wage of the emerging market as normalizing factor to perform the same analysis.

A.5 Hopenhayn Version

A.5.1 Static Problem

I solve the static problem of the firm first. I ignore the overhead adoption cost for now but it will show up in the dynamic problem, of course. Recall the production function

$$y_i = \left(\left(\int_{j\in\Omega_{A_i}} \left(\frac{x_j}{\alpha}\right)^{\alpha}\right)\left(\frac{l_i^{\zeta}}{\zeta(1-\alpha)}\right)^{1-\alpha}\right)$$

where the reader should take account of the curvature induced by $\zeta < 1$. Note that $\left(\frac{1}{\alpha}\right)^{\alpha}\left(\frac{1}{\zeta(1-\alpha)}\right)^{1-\alpha}$ is a normalizing constant for beautification purposes only. [61]

The firm solves the static problem –abstracting away from the technology adoption margin for

[61] One could easily put the curvature on the outer brackets as well without changing the economics of the paper, i.e., $y_i = \left(\frac{1}{\zeta}\left(\int_{j\in\Omega_{A_i}}\left(\frac{x_{ij}}{\alpha}\right)^{\alpha} dj\right)\left(\frac{l_i}{1-\alpha}\right)^{1-\alpha}\right)^{\zeta}$. This formulation, however, leads to slightly less elegant solutions.

now– where the price of the final good is normalized to unity, $P_y := 1$.

$$\max_{x_j, l_i} y_i - wl_i + \int p_{x,j} x_j dj \tag{66}$$

I solve this problem in two stages. First, consider a given expenditure on intermediate goods $p_x \int x_j dj := R_X$ and solve for the optimal demand for each intermediate good

$$x_j^D = \frac{R_X}{P_X^{\frac{\alpha}{\alpha-1}}} p_{x,j}^{\frac{1}{\alpha-1}}$$

where the ideal price index reads $P_X = \left(\int p_{x,j}^{-\frac{\alpha}{1-\alpha}} dj\right)^{-\frac{1-\alpha}{\alpha}}$. Given symmetry, all prices are identical so the firms optimally spreads total capital expenditure evenly across intermediates

$$x_j = \frac{X}{A_i} \; \forall j.$$

Now I plug this solution back into (66), and solve the following static maximization problem

$$\max_{X_i, l_i} \left(\left(\frac{X_i}{\alpha}\right)^{\alpha} \left(A_i \left(\frac{l_i}{\zeta(1-\alpha)}\right)^{\zeta} \right)^{1-\alpha} \right) - wl_i - p_x X_i$$

which leads to the following first-order conditions

$$y_i \alpha = X_i p_x$$
$$y_i \zeta (1-\alpha) = l_i w.$$

Factor payments are constant shares of revenue, due to the Cobb-Douglas assumption, so firm operating profits equal $\pi^o = y - l_i w - X_i p_x = (1-\zeta)(1-\alpha) y$. Moreover, after inverting the first order conditions and plugging them back into the production function, output appears as a function of of input prices, and importantly, of the productivity of the firm A_i

$$
\begin{aligned}
y_i &= \left(\frac{X_i}{\alpha}\right)^{\alpha}\left(A_i\left(\frac{l_i}{\zeta(1-\alpha)}\right)^{\zeta}\right)^{1-\alpha} \\
&= \left(\frac{y_i\alpha}{p_x\alpha}\right)^{\alpha}\left(A_i\left(\frac{\left(\frac{y_i\zeta(1-\alpha)}{w}\right)}{\zeta(1-\alpha)}\right)^{\zeta}\right)^{1-\alpha} \\
&= \left(\frac{y_i}{p_x}\right)^{\alpha}\left(A_i\left(\frac{y_i}{w}\right)^{\zeta}\right)^{1-\alpha} \\
&= y_i^{\alpha+\zeta(1-\alpha)}(p_x)^{-\alpha}\left(\frac{w^{\zeta}}{A_i}\right)^{-(1-\alpha)} \\
&= y_i^{\alpha+\zeta(1-\alpha)}(p_x)^{-\alpha}\left(\frac{w^{\zeta}}{A_i}\right)^{-(1-\alpha)} \\
&\Leftrightarrow \\
y_i &= \left\{(p_x)^{-\alpha}\left(\frac{w^{\zeta}}{A_i}\right)^{-(1-\alpha)}\right\}^{\frac{1}{(1-\alpha)(1-\zeta)}}.
\end{aligned}
$$

Take account of the partial derivative of output with respect to productivity

$$
\frac{\partial \log y}{\partial \log A} = \frac{1}{1-\zeta}, \tag{67}
$$

which will become relevant for the dynamic problem considered next. Since $\pi = (1-\zeta)(1-\alpha)y$, it follows that $\frac{\partial \pi}{\partial A} = (1-\alpha)\frac{y}{A} = \frac{1}{1-\zeta}\frac{\pi}{A}$.

A.5.2 Dynamic Adoption Problem

I first solve the problem of an incumbent firm, i.e., the technology adoption problem, before I turn attention to the free entry condition. The dynamic problem reads

$$
\begin{aligned}
(r+\delta_X)V &= \pi^o + \dot{V} + V_A\dot{A} - w_H h \\
&\text{s.t.} \\
\dot{A} &= \nu A^{\theta}A_F^{1-\theta}h^{\beta} - \delta_I A.
\end{aligned}
$$

To make the problem stationary, I normalize the value function by the real wage of unskilled labor, and I focus on the state variable $z := \frac{A}{A_F}$, which will be constant in equilibrium while A and A_F will be continuously growing. Define $v := \frac{V}{w}$, and $z := \frac{A}{A_F}$, and note that this change of variable has no effect on the underlying economics, which can now be studied in the stationary system

$$\left(r-g_w+\delta_X\right)v=\frac{\pi^o}{w}+\dot{v}+v_z\dot{z}-sh \tag{68}$$

$$\text{s.t.}$$

$$\dot{z}=\nu z^\theta h^\beta-\left(\delta_I+g_F\right)z.$$

An interior solution to the system (68) satisfies the first-order condition

$$h=\left\{\frac{v_z\beta\nu z^\theta}{s}\right\}^{\frac{1}{1-\beta}}, \tag{69}$$

which captures the tradeoff between the cost of adoption in terms of the price of skilled labor, and the benefit of improving the firm's productivity.

To derive the differential equation that governs optimal skilled labor investment choices, I differentiate (68) with respect to z, and after using the Envelope theorem, I arrive at

$$\left(r-g_w+\delta_X\right)=\frac{\pi_z^o}{wv_z}+\frac{\dot{v}_z}{v_z}+\frac{v_{zz}\dot{z}}{v_z}+\nu\theta z^{\theta-1}h^\beta-\left(\delta_I+g_F\right). \tag{70}$$

Now taking logs of (69) and differentiating with respect to time yields

$$\frac{\dot{h}}{h}=\frac{1}{1-\beta}\left\{\frac{\dot{v}_z}{v_z}+\frac{v_{zz}}{v_z}\dot{z}+\theta\frac{\dot{z}}{z}-\frac{\dot{s}}{s}\right\}. \tag{71}$$

Now plug (70) into (71) to get

$$\frac{\dot{h}}{h}=\frac{1}{1-\beta}\left\{\left(r-g_w+\delta_X\right)-\frac{\pi_z^o}{wv_z}-\theta\nu z^{\theta-1}h^\beta+\left(\delta_I+g_F\right)+\theta\frac{\dot{z}}{z}-\frac{\dot{s}}{s}\right\}$$
$$\frac{\dot{h}}{h}=\frac{1}{1-\beta}\left\{\left(r-g_w+\delta_X\right)-\frac{\pi_z^o}{wv_z}-\theta\nu z^{\theta-1}h^\beta+\left(\delta_I+g_F\right)+\theta\left[\nu z^{\theta-1}h^\beta-\left(\delta_A+g_F\right)\right]-\frac{\dot{s}}{s}\right\}$$
$$\frac{\dot{h}}{h}=\frac{1}{1-\beta}\left\{r-g_w+\delta_X+\left(1-\theta\right)\left(\delta_I+g_F\right)-\frac{\dot{s}}{s}-\frac{\pi_z^o}{wv_z}\right\}$$
$$\frac{\dot{h}}{h}=\frac{1}{1-\beta}\left\{r-g_w+\delta_X+\left(1-\theta\right)\left(\delta_I+g_F\right)-\frac{\dot{s}}{s}-\frac{\beta h^{\beta-1}\nu z^{\theta-1}}{s}\frac{\pi_z^o z}{\pi^o}\frac{\pi^o}{w}\right\}.$$

The final piece to derive the dynamic investment equation involves deriving the partial elasticity of profits with respect to productivity, which is $\frac{1}{1-\zeta}$ as can be seen in equation (67), which yields

$$\frac{\dot{h}}{h} = \frac{1}{1-\beta}\left\{r - g_w + \delta_X + (1-\theta)(\delta_I + g_F) - \frac{\dot{s}}{s} - \frac{\beta \nu z^{\theta} h^{\beta-1}}{s}\left[\frac{1}{1-\zeta}\frac{\pi^o}{w}\frac{1}{z}\right]\right\}. \tag{72}$$

For now, assume that a well-defined steady state exists, which I will prove later, together with stability and uniqueness properties, and assume that in such a steady state the adoption gap is constant and firms higher a constant number of skilled workers to adopt technology, i.e., $\dot{z} = 0$ and $\dot{h} = 0$ so that

$$z_* = \left(\frac{\nu}{\delta_I + g_F}\right)^{\frac{1}{1-\theta}} (h_*)^{\frac{\beta}{1-\theta}} \quad \text{(from resource constraint } \dot{z}\text{)}$$

Using this link, one can further simplify (72) to derive h_* in the steady state, which is inversely related to the skill premium,

$$r - g_w + \delta_X + (1-\theta)(\delta_I + g_F) = \frac{(\delta_I + g_F)}{h^* s}\frac{\beta}{1-\zeta}\frac{\pi}{w}$$

$$\Leftrightarrow$$

$$h_* = \frac{1}{s}\frac{(\delta_I + g_F)\frac{\beta}{1-\zeta}}{r - g_w + \delta_X + (1-\theta)(\delta_I + g_F)}\frac{\pi}{w}. \tag{73}$$

One can verify that the following inequality is necessary for a well-defined balanced growth path

$$\frac{\beta}{1-\zeta} < \frac{r - g_w + \delta_X}{\delta_I + g_F} + (1-\theta). \tag{74}$$

Proof. By contradiction, suppose that $\frac{\beta}{1-\zeta} \geq \frac{r-g_w+\delta_X}{\delta_I+g_F} + (1-\theta)$. In that case, spending on learning in the steady state, $h_* s$, is weakly larger than total operating profits $\frac{\pi^o}{w}$, so that the net profits $\frac{\pi}{w} = \left(\frac{\pi^o}{w} - s h_*\right)$ are non-positive. However, firms need to make positive profits to cover the initial fixed cost of entry. Thus, any well-defined balanced growth path requires (74) to hold. □

A.5.3 Free Entry

To close the model, Hopenhayn (1992) introduces a free entry condition such that the cost of entry, paid in labor, equals the present discounted value of the firm

$$f_{\mathrm{E}} w_{L,t} = \int V dF_t(A|\mathrm{E}). \tag{75}$$

In the main part of the paper I assume that $F(A|\mathrm{E}) = F(A)$ where $F(A)$ refers to the incumbent productivity distribution (expressing this in terms of total productivity A or relative productivity z is inconsequential since the firm takes the frontier A_{F} as given). The remaining steps are extremely similar to the baseline model.

A.6 Heterogeneous Firms with Partial Knowledge Spillovers

A simplifying assumption in the paper is the complete knowledge spillover from incumbents to entrants. An alternative specification is one where the entrant only obtains a fraction $\lambda_{\mathrm{E}}\mathbb{E}\left[z\right]$ where $\lambda_{\mathrm{E}} < 1$. This tweak turns the setting into a heterogeneous firm model where entrants learns from the most sophisticated incumbent, but imperfectly so. A well-defined equilibrium is characterized by a distribution $f\left(z\right)$ with support $z \in \left[\lambda_{\mathrm{E}}\mathbb{E}\left[z\right], z_{\max}\right]$. This leads to a normalized free entry condition

$$f_{\mathrm{E}} = v\left(\lambda_{\mathrm{E}}\mathbb{E}\left[z\right]\right).$$

Building on Hopenhayn (1992) and Melitz (2003), the profit ratio of any two firms can be expressed as $\frac{\pi_i}{\pi_j} = \left(\frac{z_i}{z_j}\right)^{(1-\alpha)(\sigma-1)}$, and normalized profits for firm i are given by $\frac{\pi(z_i)}{w} = \frac{(z_i)^{(1-\alpha)(\sigma-1)}}{\mathbb{E}\left[z^{(1-\alpha)(\sigma-1)}\right]} \frac{l_P}{m(\sigma-1)(1-\alpha)}$. Instead of using i subscript, I now index firms by z.

Next, consider the problem of some firm with productivity z in the steady state (so $\dot{v} = 0$)

$$\left(\rho + \delta_{\mathrm{X}}\right) v\left(z\right) = \max_{h} \pi\left(z\right) - sh + \left(\partial_z v\right) \cdot \left[\nu z^{\theta} h^{\beta} - \left(g_{\mathrm{F}} + \delta_{\mathrm{I}}\right) z\right].$$

The first order condition still reads

$$h\left(z\right) = \left\{\frac{\left(\partial_z v\right)\beta\nu z^{\theta}}{s}\right\}^{\frac{1}{1-\beta}}.$$

I assume for simplicity that $\dot{s} = 0$ and derive a similar dynamic investment equation as for the homogeneous firm case

$$\frac{\dot{h}_i}{h_i}\left(1-\beta\right) = \left(\rho + \delta_{\mathrm{X}} + \left(1-\theta\right)\left(g_{\mathrm{F}} + \delta_{\mathrm{I}}\right)\right) - \frac{\beta\nu z^{\theta} h^{\beta-1}}{s}\left[\frac{\pi}{w}\frac{\left(1-\alpha\right)\left(\sigma-1\right)}{z}\right].$$

It is useful to rewrite this expression relative to the firms with the maximum productivity

$$\begin{aligned}\frac{\dot{h}_i}{h_i}\left(1-\beta\right) = {} & \left(\rho + \delta_{\mathrm{X}} + \left(1-\theta\right)\left(g_{\mathrm{F}} + \delta_{\mathrm{I}}\right)\right) \\ & - \frac{\beta\left(1-\alpha\right)\left(\sigma-1\right)\left(z_{\max}\right)^{\theta-1} h_{\max}^{\beta-1}\nu}{s}\frac{\pi_{\max}}{w}\left(\frac{z_i}{z_{\max}}\right)^{\theta-1+(1-\alpha)(\sigma-1)}\left(\frac{h_{\max}}{h_i}\right)^{1-\beta}\end{aligned}$$

which helps to pin down the equilibrium dynamics. By construction, the most productive firm hires a constant amount of skilled labor with the only difference to the homogenous firm model being that the steady state profits are larger. This is a direct consequence of starting out with an initially lower productivity. Higher long-run profits have to make up for low profits after the firm just entered, since the entry cost are the same in both cases.

For this equilibrium to be well-defined, I need it to be true that entering firms improve their productivity so that they converge to the more profitable incumbents in the long-run. This is not trivial. One way to see this is to sue a phase-diagram in the h-z space assuming a stationary equilibrium exist for some $z_{\max}$. Since there is no risk the differential equation I have derived beforehand applies. Slightly rewriting and ignoring a a scalar $1-\beta$ (which does not matter for the issue of existence but surely matters for the speed of convergence) leads to

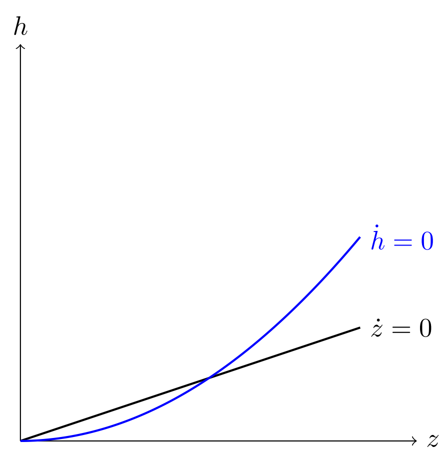
$$\frac{\dot{h}}{h} = \kappa_1 \left(1 - \left(\frac{z}{z_{\max}}\right)^{\theta-1+(1-\alpha)(\sigma-1)} \left(\frac{h_{\max}}{h}\right)^{1-\beta}\right).$$

Assuming $\dot{h} = 0$ implies $h = \left(\frac{z}{z_{\max}}\right)^{\frac{\theta-1+(1-\alpha)(\sigma-1)}{1-\beta}} h_{\max}$. Moreover, assume $\frac{\theta-1+(1-\alpha)(\sigma-1)}{1-\beta} > 0$ so the $\dot{h}$ locus is upward-sloping in z. Moreover, note that the $\dot{z}$ locus implies a positive link between h and z as well, $h = \left(\frac{g_F+\delta_I}{\nu} z^{1-\theta}\right)^{\frac{1}{\beta}}$. When

$$\frac{\theta - 1 + (1-\alpha)(\sigma-1)}{1-\beta} < \frac{1-\theta}{\beta},$$

one can show that the z-locus cuts the h-locus once from below. The associated stability analysis suggests convergence from below, see figure 13.

Figure 13. Phase Diagram



This structure gives rise to a meaningful stationary distribution whereby firms start out small and improve their productivity over time. Several features are noteworthy. First, the firm size distribution is independent of the relative price of skill s. What this suggests is that a new stationary equilibrium with a higher price of skill produces an identical wave but shifted to the left, i.e., a permanently lower level of adoption across all firms. This traveling wave property is not surprising in light of recent work on heterogeneous firms, see Luttmer (2007), König, Lorenz, and Zilibotti (2016), Sampson (2016), Benhabib, Perla, and Tonetti (2021) and Perla and Tonetti (2014). Crucially, the partial equilibrium elasticity $\frac{\partial \log \mathbb{E}[z]}{\partial \log s} = -\frac{\beta}{1-\theta}$ computed in the main text still applies. Second, demand for skilled labor in the production sector can be derived by integrating over all productivity levels

$$h_{\text{D}} = m \int_{\lambda_{\text{E}}\mathbb{E}[z]}^{z_{\max}} f(z)\, h(z)\, dz.$$

Third, the innovator problem in the steady state needs to be updated as follows

$$V = \mathbb{E}_z \left[V(z) \right] \tag{76}$$

where $V(z) = \int_{t+\tau(z)}^{\infty} \exp\left(-\int_t^s [r_u + \delta_{\text{X}}]\, du\right) \pi_{\text{I}}(z,s)\, ds$ is a function of the firm-specific z-level which matters both in terms of firm size and how long it takes for an idea to be adopted by a firm of type z. The problem is conceptually the same as before except now one needs to keep track of the distribution of firm-specific adoption gaps. Of course, equation (76) is also conceptually very close to the value function of an innovator in the open economy in the main text, which is a discretized version of (76) where z's pertain to different countries.

One desirable feature of the heterogeneous firm model is that it can speak to the distinction between the extensive and intensive margin of technology adoption, i.e., whether a specific technology is used in a country, and whether all firms in a country use this specific technology, see Comin and Hobijn (2004) and Banerjee and Duflo (2005) for empirical evidence on the importance of the intensive margin, i.e., differences across firms within a country.

A.7 Other Model Extensions

Stochastic Adoption. Since asset markets are complete and there are no stochastic shocks, risk plays no role when potential innovators consider entry into innovation. It is thus not surprising that stochastic adoption does not change any of the results qualitatively.

For example, a different version that I have experimented with is to let un-adopted ideas to be uniformly sampled at Poisson rate $\frac{A(g_A+\delta_I)dt}{A_F - A} = \frac{z}{1-z}(g_A + \delta_I)$ where $\frac{1}{A_F - A}$ is the uniform density and $A(g_A + \delta_I)\, dt$ is the flow of new ideas that are adopted at each instant. The probability density is then simply the product of the two, given statistical independence. On a balanced growth path with constant relative technology level z, it is again true that a z close to unity makes the adoption friction vanish. In contrast, as z approaches zero, the net present value of an innovation falls to zero as well since the adoption probability converges to zero as well.

Using this alternative functional form, one can follow the same steps as in the main text and compute the expected present discounted value of a patent. The insight that adoption and innovation are complementary on the market for ideas are robust to this alternative functional form. While stochastic adoption is more realistic in the sense that most innovators do not know when, if ever, their idea becomes profitable, this version of the model would be slightly less tractable regarding the market clearing condition for skilled labor.

The value function of an innovator would now look as follows

$$(r - g_V + \delta_{\mathrm{I}}) V_{\mathrm{I}} = 0 + \lambda(z) \left(V_{\mathrm{I}}^{\text{adopted}} - V_{\mathrm{I}} \right) \tag{77}$$

where $g_V = \frac{\dot{V}}{V}$ is pinned down by the free entry condition, $\lambda(z) = \frac{z}{1-z}(g_A + \delta_I)$, and $V_{\mathrm{I}}^{\text{adopted}} = \int_t^{\infty} e^{-\int_t^u r_v + \delta_{\mathrm{I}} dv} \pi_{\mathrm{I}}(u)\, du$. The zero in (77) is meant to highlight that there are no flow profits up until the idea is adopted.

Government & Complementary Infrastructure. A classic theme in the growth literature is government capacity, and the important role of complementary public infrastructure investment. For example, the adoption of motorized vehicles will not occur unless the government builds roads. An easy way to incorporate this aspect into the model is to generalize the adoption technology with a notion of public infrastructure $A_{\mathrm{G},t}$ so that the net flow of ideas reads $\dot{A}_i = A_{\mathrm{F}}^{1-\theta} A_{\mathrm{G}}^{(1-\phi)\theta} A_i^{\phi\theta} - \delta_{\mathrm{I}} A_i$. A normalized version simplifies to $\dot{z}_i = \nu z_{\mathrm{G}}^{(1-\phi)\theta} z_i^{\phi\theta} h_i^{\beta} - (\delta_{\mathrm{I}} + g_{\mathrm{F}}) z_i$ where the firm takes z_{G} as given. Without complementary government investment z_{G} will be zero and the returns to adoption are zero as well. If the government invests in $A_{\mathrm{G}}^{(1-\phi)\theta}$ at a rate consistent with a balanced growth path, then all of the previous derivations go through after updating the term θ by a factor of ϕ. Note that capable government employees are likely to be instrumental in building up complementary public infrastructure. The role of skilled labor remains pivotal.

Unskilled Labor in Innovation and Upward-Sloping Skilled Labor Supply. Another natural extension is to allow unskilled labor to be used in the research sector as well, i.e., suppose that the entry cost combine each labor type according to a Cobb-Douglas production function with skilled labor share κ, so the entry cost becomes $f_{\mathrm{R}} \left(\frac{w_H}{\kappa}\right)^{\kappa} \left(\frac{w}{1-\kappa}\right)^{1-\kappa}$. In that case, an expansion of the research sector will simultaneously raise demand for unskilled labor.[62] It will still be true, though, that the impact on unskilled worker wages can be characterized by the evolution of frontier technology A_{F} and the relative technology level z alone. Instead of making additional assumptions on the factor intensity in each sector, one can capture these considerations by allowing the supply of skilled labor to be upward sloping in the skill premium. To this end, suppose that an exogenously growing population

[62] See Helpman (2016) for a summary of the large literature on factor-biased trade and inequality.

of workers split between unskilled and skilled labor $N = L + H$ where $\psi(s)$ is the share of workers that are skilled, which is now an upward sloping function of the skill premium. The market clearing condition for skilled labor now reads

$$\frac{1}{s} z^{-\frac{\tilde{\rho}}{g_A + \delta_I}} \Lambda_F + \frac{1}{s} \Lambda_D = \frac{\psi(s)}{1 - \psi(s)}.$$

I assume $\frac{\psi(s)}{1-\psi(s)} = \psi_0 (s - \kappa)^\eta$ where the baseline case refers to $\eta \to 0$. As $\eta \to \infty$ and skilled labor supply becomes perfectly elastic,[63] the skill premium is exogenously fixed at $s = \kappa > 1$. In that scenario, the adoption margin is mute, and the implications of the theory with respect to a market integration shock are extremely similar to the simpler model of Jones (1995) without adoption margin and one type of labor. I highlight the role of skill scarcity in driving weak technology adoption. Surely skilled labor supply is somewhat elastic, but not nearly enough to counteract the substantial increase in the skill premium observed over the past couple of decades.

Skill Biased Technological Change. A common explanation for rising inequality is based on theories of skill-biased technological change, see Katz and Murphy (1992). Goldin and Katz (2010) present compelling empirical evidence from a number of studies covering almost two centuries that show how skill-biased technological change has shaped labor market outcomes. It is thus useful to consider how my model relates to this large literature. I generalize the model to include allow for changing task-content of work by modeling intermediate goods production as $y = \left((Ax)^\alpha l^{1-\alpha}\right)^{1-\tilde{\beta}} h^{\tilde{\beta}}$ so that both unskilled and skilled labor enters the production function ($\tilde{\beta} = 0$ is the baseline case in the paper).[64] The model remains mostly unchanged except for an additional term $\tilde{\Lambda}_{\tilde{\beta}}$ in the labor market clearing condition,

$$\frac{1}{s} \left(\tilde{\Lambda}_F z^{\frac{\tilde{\rho}}{g_F + \delta_I}} + \tilde{\Lambda}_D + \tilde{\Lambda}_{\tilde{\beta}} \right) = h_{tot}. \tag{78}$$

A changing task content is captured in an increase in $\tilde{\beta}$ (or $\Lambda_{\tilde{\beta}}$) and would raise the overall price of skill. This would push down aggregate growth as less skilled labor is available for innovation and adoption. As production requires more skill, less is available to invest in innovation and adoption.

As pointed out in Acemoglu and Autor (2011), skill-biased technological change generates wage growth *for all workers.* The reason is the strong complementarity between high and low skilled workers which ensures that technological change benefits everyone, even if it is biased. The theory proposed here is complementary to this literature by pointing out that a reallocation of skill across space or sectors can create real wage losses whenever skill is an important input to technology adoption. If so, the skill premium takes on a new role where an increase in the relative price of skilled labor reduces equilibrium adoption effort and thus hampers economic growth.

Note, however, that an increase in the relative price of skill driven by a changing task content of

[63] I do not model the actual investment cost, see for instance Acemoglu (2009) for a model with explicit schooling choice.
[64] Acemoglu and Restrepo (2020) show how to micro-found this Cobb-Douglas production function in a model of automation.

work will hit the innovation sector the hardest due to the second round effects through a rising adoption gap as $z^{\frac{\tilde{\rho}}{g_F+\delta_I}}$ falls. A changing task content of work is thus consistent with sluggish growth and rising inequality in this model, but it will not allow innovative activity to take off. The effect of globalization on the returns to innovation will resolve this tension and help make sense of rising innovative activity in advanced economies.

Skilled Labor in Production Sector Entry Technology. Note that in the baseline model the downward sloping relationship between the skill premium and the demand for skilled labor for adoption purposes is directly related to the fact that long-run firm profits are proportional to the cost of entry, which in turn is proportional to unskilled worker wages. A more general version utilizes a Cobb-Douglas bundle of high-skill and unskilled labor, $f_{\mathrm{E}} w_L^{\mu} w_H^{1-\mu}$. In that case the effect of an increase in the price of skill on technology adoption changes to

$$-\frac{\partial \log z}{\partial \log s} = \frac{\beta}{1-\theta} \cdot \mu, \tag{79}$$

while an additional effect on firm entry emerges, i.e., $\frac{\partial \log m}{\partial \log s} < 0$ for $\mu < 1$. The parameter μ governs to what extent a high skill premium translates into weak technology adoption relative to firm entry. In the open economy cross-country context it will be desirable to load the effect of skill-scarcity entirely on the technology adoption margin, in part because the firm-to-population ratio is no lower in poorer skill-scarce countries. The more general entry technology and effects on firm entry may be useful when thinking about declining business dynamics in advanced economies, see Decker et al. (2017), which is beyond the scope of the paper.[65] .

Immigration. A fully integrated equilibrium with free migration behaves differently from the baseline model. Unskilled migration in the baseline model reduces unskilled worker wages in the advanced economy in two ways. First, a direct negative effect on the unskilled wage as their factor becomes relatively more abundant. Second, since there are more unskilled workers, more skilled labor needs to be devoted to technology adoption, further hurting innovation and reducing frontier innovation.

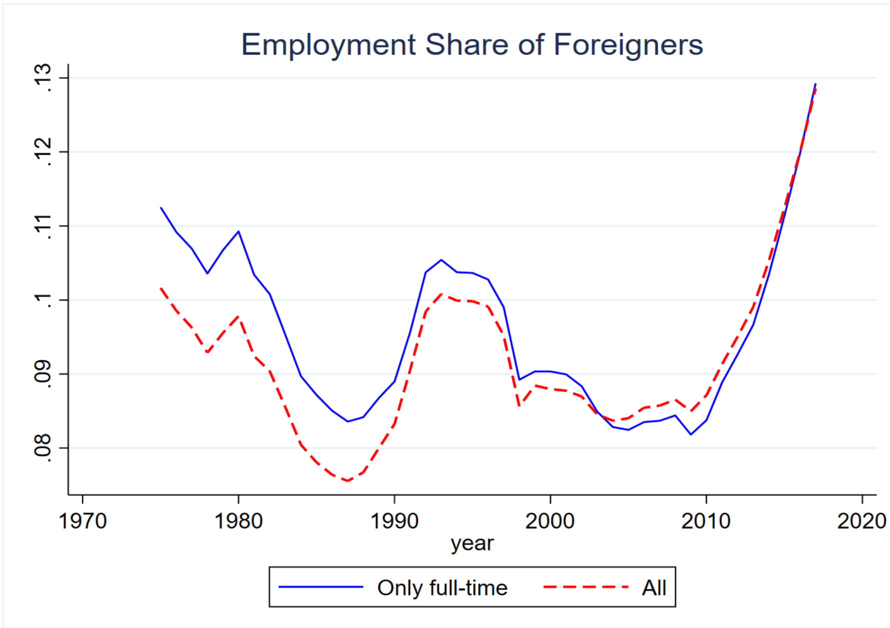
Skilled labor inflows, on the other hand, are extremely beneficial for unskilled workers as they lead to both heightened adoption and innovation. Interestingly, the effect on skilled labor is ambiguous and likely positive as well. Even though their relative wage has to fall, greater adoption and rising innovation can lift up the high skilled workers' wages.

Figure 14 helps us assess the relevance of immigration into Germany as a potential reason for weak wage growth. It turns out that the foreign employment share is falling since the mid 1990s, leading to an all-time low in the 2000s. This figure suggests that trends in immigration are unlikely to account for wage stagnation in Germany at the turn of the century. This is not to say that immigration doesn't matter – it seems hard to think about technological frontier growth, especially in the US, without the

[65] This point is related to Salgado (2020) where skill-biased technological change leads to less entry into entrepreneurship.

input of foreign researchers and entrepreneurs. The model can in principal be used to explore this issue, and consider the effect of brain drain and innovation gains in a global integrated equilibrium with realistic scale effects. I do not want to take up this task in this paper.

Figure 14. Immigration



The figure plots the share of foreign workers in West Germany, using the BHP of the IAB.

B Computational Appendix

Here I show how to compute the solution and solve for transition dynamics based on the finite difference method in Achdou et al. (2022).

B.1 Closed Economy

B.1.1 Production Sector Firms' Problem

The normalized firm problem reads

$$(r_t + \delta_{\mathrm{X}} - g_w)\, v\,(z, X) = \frac{\pi\,(z)}{w} - sh + \partial_z v \cdot \left[\nu z^{\theta} h^{\beta} - (\delta_{\mathrm{I}} + g_{\mathrm{F}})\, z\right] + \mathcal{A}v$$

subject to the law of motion for z_i and the evolution of aggregate state variables captured in X, including the equilibrium measure of firms, among other things. The framework can be readily extended to include stochastic productivity shocks in $\mathcal{A}$, but I drop this term here.

The first-order condition reads

$$\left\{\frac{(\partial_z v)\, \beta \nu z^{\theta}}{s}\right\}^{\frac{1}{1-\beta}} = h,$$

and I use the following boundary conditions

$$\left(\partial_z v^{\mathrm{f}}\right) = \frac{s}{\beta \nu \overline{z}^{\theta}} \left\{\frac{(\delta_{\mathrm{I}} + g_{\mathrm{F}})\, \overline{z}^{1-\theta}}{\nu}\right\}^{\frac{1-\beta}{\beta}}$$
$$\left(\partial_z v^{\mathrm{b}}\right) = \frac{s}{\beta \nu \underline{z}^{\theta}} \left\{\frac{(\delta_{\mathrm{I}} + g_{\mathrm{F}})\, \underline{z}^{1-\theta}}{\nu}\right\}^{\frac{1-\beta}{\beta}}$$

for forward and backward difference, respectively, where $\overline{z}$ and $\underline{z}$ represent the highest and lowest value of z over the grid space.

Rewrite the problem as follows

$$(r_t + \delta_{\mathrm{X}} - g_w)\, v\,(z) = B_t z^{(1-\alpha)(\sigma-1)} - sh + \partial_z v \cdot \left[\nu z^{\theta} h^{\beta} - (\delta_{\mathrm{I}} + g_{\mathrm{F}})\, z\right] \tag{80}$$

and note that B_t captures, among other things, the equilibrium measure of firms as well as the properly weighted average productivity level $\tilde{z} = \left\{\mathbb{E}\left[z^{(1-\alpha)(\sigma-1)}\right]\right\}^{\frac{1}{(1-\alpha)(\sigma-1)}}$, where the heterogeneity arises only for the case of imperfect spillover at entry, $\lambda_{\mathrm{E}} < 1$.

Note that $B_t M \tilde{z}^{(1-\alpha)(\sigma-1)} = \frac{Y}{\sigma}$, which can be rearranged to

$$B_t = \frac{l_{\mathrm{P}}}{m\tilde{z}^{(1-\alpha)(\sigma-1)}} \frac{1}{(1-\alpha)(\sigma-1)}.$$

Moreover, in the steady state $m = \frac{1-l_{\mathrm{P}}}{(g_L+\delta_{\mathrm{X}})f_{\mathrm{E}}}$. If I had a value for B_t, and a solution to (80), I could infer all endogenous variables. To find B_t, I solve (80) for a guess of B_t and check if the free entry condition holds

$$f_{\mathrm{E}} = \mathbb{E}_{z_{\mathrm{E}}}\left[v\left(z_{\mathrm{E}}\right)\right].$$

I raise the value of B_t if the cost of entry is higher than the present discounted value using a bisection method.

The density of incumbents g_z and entrants $g_{z_{\mathrm{E}}}$ follows from the KFE equation

$$\dot{g}_z = \mathcal{A}^T\left(g_z\right) g_z$$

where I make clear that the infinitesimal generator $\mathcal{A}^T\left(g_z\right)$ is itself a function of the stationary density since entrants' productivity is assumed to be a function of the average $z_{\mathrm{E}} = \lambda_{\mathrm{E}}\mathbb{E}\left[z\right]$. Note that a flow of $g_L + \delta_{\mathrm{X}}$ firms enters at productivity z_{E}, accounting for both net entry rate (which is the same as long run population growth) and exogenous firm death. In matrix form, this means I have to add $-\left(g_L + \delta_{\mathrm{X}}\right)$ on the diagonal, and $\left(g_L + \delta_{\mathrm{X}}\right)\mathbb{1}_{z=z_{\mathrm{E}}}$in each column of $\mathcal{A}^T$.[66]

Having solved for B_t and g_z, values for l_{P} and m follow as a function of the skill premium

$$l_{\mathrm{P}} = \frac{B_t \tilde{z}^{(1-\alpha)(\sigma-1)}}{\frac{(g_L+\delta_{\mathrm{X}})f_{\mathrm{E}}}{(1-\alpha)(\sigma-1)} + B_t \tilde{z}^{(1-\alpha)(\sigma-1)}}$$
$$m = \frac{1-l_{\mathrm{P}}}{\left(g_L+\delta_{\mathrm{X}}\right) f_{\mathrm{E}}}$$

which in turn imply aggregate demand for skilled labor from the production sector

$$h_{\mathrm{D}} = m \cdot \int h\left(z\right) g_z dz.$$

B.1.2 Research Sector Problem

An alternative approach uses the following normalization

$$v_{\mathrm{I}} := \frac{V_{\mathrm{I}}}{w_L A_{\mathrm{F}}^{-\phi}},$$

[66]If z_{E} falls between two grid points, I turn the entry draw into a probabilistic one according to how close z_{E} is to either grid point.

which leads to the following recursion

$$v_{\mathrm{I}}\left(r_t - g_{w_L} + \delta_{\mathrm{I}} + \phi g_{\mathrm{F}}\right) = e^{-\int_t^{t+\tau}\left(r_t - g_{w_L} + \delta_{\mathrm{I}} + \phi g_{\mathrm{F}}\right)dv} \frac{\alpha l_{\mathrm{P},t+\tau}}{z_{t+\tau} a_{\mathrm{F},t+\tau}} \cdot \left[1 + \dot{\tau}_t\right] + \dot{v}_{\mathrm{I}},$$

and in discretized form

$$v_{\mathrm{I},t}\left(r_{t+\Delta} - g_{w_L,t+\Delta} + \delta_{\mathrm{I}} + \phi g_{\mathrm{F},t+\Delta}\right) = e^{-\int_t^{t+\tau}\left(r_t - g_{w_L} + \delta_{\mathrm{I}} + \phi g_{\mathrm{F}}\right)dv} \frac{\alpha l_{\mathrm{P},t+\tau}}{z_{t+\tau} a_{\mathrm{F},t+\tau}} \cdot \left[1 + \dot{\tau}_t\right] + \frac{v_{\mathrm{I},t+\Delta} - v_{\mathrm{I},t}}{\Delta}.$$

For a free entry equilibrium, I want to make sure at all times that

$$v_{\mathrm{I}} \leq \frac{1}{\gamma} h_{\mathrm{F}}^{1-\lambda} s$$

holds. This recursive representation can be used to compute equilibrium in the research sector off the balanced growth path.

B.1.3 Transition Dynamics

Production sector. Taking the sequence $\{s_t, g_{\mathrm{F},t}, B_t\}_{t\in[0,T]}$ as well as initial and terminal conditions as given, I solve the production firms' problem along the transition path. The first order condition now reads

$$\left\{\frac{\left(\partial_z v_{n+1}\right)\beta\nu z_{n+1}^{\theta}}{s_{n+1}}\right\}^{\frac{1}{1-\beta}} = h,$$

where each n is a step in time. Given a solution, the value function in v_n follows recursively.
Firm entry in production. In principal, the entry margin responds to technology adoption along the transition path. In particular, when the knowledge spillover at entry is strong high technology adoption effort coincides with lower firm entry so as to respect the free entry condition.

Given a solution for $\{v_{\mathrm{E}}\}$, I can asses whether the value of entry equals the entry cost f_{E}, and update labor devoted to firm entry accordingly. Computationally, this is inconvenient because the linear entry technology leads to bang-bang solutions, which are avoided in general equilibrium but cumbersome to solve for during the transition.

I simplify the model by assuming that $l_{\mathrm{P},t} = l_{\mathrm{P,ss}}$, i.e., a constant amount of labor is devoted to entry consistent with the long run steady state level, which itself is only a function of constant exogenous parameters. We will see that this simplifying assumption, which also makes capital accumulation easier since the effective number of production workers won't jump around during the transition, is quantitatively innocuous in that the value of entry into the production sector deviates less than 1% from its long-run steady state.

Technology Adoption and Waiting time. The model is hard, in part, because innovators need to keep track of changes in τ, and multiple τ's in the open economy version. Note that τ_0 is a jump variable that is implicitly defined by

$$\tau_0 = -\frac{\log z_0}{\frac{\int_0^{\tau_0} g_A(x)+\delta_{\mathrm{I}} dx}{\tau_0}},$$

and for a given evolution of $A(t)$ one can use a simple bisection to get τ_0. Once the initial wait time is determined, one can use the law of motion of the waiting time

$$\dot{\tau}_t = \frac{g_{\mathrm{F},t} - g_{A,t+\tau}}{\delta_{\mathrm{I}} + g_{A,t+\tau}},$$

which can be inverted, and using the discrete time step approximation, yields next periods wait time as a function of frontier growth, productivity growth, and the previous wait time

$$\tau_{t+\Delta} = \tau_t + \Delta \cdot \frac{g_{\mathrm{F},t} - g_{A,t+\tau}}{\delta_{\mathrm{I}} + g_{A,t+\tau}}.$$

Research Sector. Note that for a guess $\{s_t\}$, and adoption choices in the production sector, the recursion derived beforehand implies a sequence $\{v_{\mathrm{I}}\}$. Together with the entry technology into innovation, this sequence implies a sequence of skilled labor devoted to innovation so as to ensure that free entry holds.

General Equilibrium. Key inputs in solving the research and production sector's forward-looking problems are wage growth and interest rate. Note that wages equal

$$w_{L,t} = \left(\frac{1-\alpha}{\alpha}\right)^{\alpha} \frac{\sigma-1}{\sigma} A_t k_t^{\alpha}$$

where $k_t = \frac{K_t}{A_t L_{\mathrm{P},t}}$ is effective capital per unskilled worker. Wage growth equals

$$g_{w_L} = g_{\mathrm{F}} + g_z + \alpha g_k,$$

where g_z and g_k are zero in the steady state.

Similarly, the interest rate equals

$$r_t = \left(\frac{\alpha}{1-\alpha}\right)^{1-\alpha} \frac{\sigma-1}{\sigma} \alpha k_t^{\alpha-1} - \delta_k.$$

To proceed, I need to pin down the evolution of the normalized capital stock k_t, which follows from the capitalists' problem.

Simplified Physical Capital Accumulation. I simplify by assuming that a constant fraction χ_s of domestic final output is reinvested in physical capital. I pick χ_s such that the autarky steady state is consistent with the solution to the representative household problem in the steady state in Autarky. Given this assumption, the law of motion of physical normalized capital $k = \frac{K}{AL_{\text{P}}}$ reads

$$\frac{\dot{k}}{k} = \chi_s \alpha^{-\alpha} (1-\alpha)^{-(1-\alpha)} k^{\alpha-1} - (g_L + g_{l_{\text{P}}} + g_{\text{F}} + g_z + \delta_k).$$

In discretized form, I have

$$k_{t+\Delta} = k_t \cdot e^{\Delta\left[\chi_s \alpha^{-\alpha}(1-\alpha)^{-(1-\alpha)} k^{\alpha-1} - \left(g_L + g_{l_{\text{P}}} + g_{\text{F}} + g_z + \delta_k\right)\right]}$$

where initial capital is given, and growth rates are computed as $\frac{\log\left(\frac{A_{t+\Delta}}{A_t}\right)}{\Delta}$. Based on the previous assumption, l_{E} is assumed to be constant. But note that even in the case of endogenous entry in the production sector, l_{E} only jumps at the very first moment, and is a smooth function of time for any $t > t_0$.This initial jump should be incorporated in the normalized starting value k_0. In the simplified case, k_0 is fully predetermined, and initial and long-run value coincide.

Given this law of motion of capital, together with optimal firm entry, the equilibrium interest rate follows

$$r_t = \left(\frac{\sigma-1}{\sigma}\alpha\right) \alpha^{-\alpha} (1-\alpha)^{-(1-\alpha)} \alpha k_t^{\alpha-1} - \delta_k$$

where the normalized capital stock in steady state reads

$$k_{\text{ss}} = \left\{\frac{\chi_s}{g_L + g_{\text{F}} + \delta_k}\right\}^{\frac{1}{1-\alpha}} \alpha^{-\frac{\alpha}{1-\alpha}} (1-\alpha)^{-1}.$$

When the saving rate, which is kept constant along the transition path, equals $\chi_s = \left(\frac{\sigma-1}{\sigma}\alpha\right) \cdot \alpha \frac{g_L + g_{\text{F}} + \delta_k}{\rho + g_{\text{F}} + \delta_k}$, I get $k_{\text{ss}} = \left\{\left(\frac{\sigma-1}{\sigma}\alpha\right) \cdot \frac{\alpha}{\rho + g_{\text{F}} + \delta_k}\right\}^{\frac{1}{1-\alpha}} \alpha^{-\frac{\alpha}{1-\alpha}} (1-\alpha)^{-1}$, which coincides with the normalized capital stock in the steady state when the saving rate is endogenous.

Algorithm. Note that I have simplified the problem so that the key endogenous object we are after is the sequence of skill premia $\{s_t\}$, which clear the market for skilled labor, and are consistent with forward-looking technology adoption and innovation choices in both production and research sector.

An algorithm that I have found to work well proceeds as follows.

1. Guess $\{s\}$.

2. Guess $\{a_{\text{F}}\}$, solve optimal adoption choice backward, and $\{\tilde{z}_{t,}\}$, $\{r_t\}$ forward. Iterate until convergence. Compute $\{\tau\}$, which depends on $\{z\}$ and $\{a_{\text{F}}\}$.

3. Solve innovator's problem holding $\{z, r\}$ fixed

(a) Holding τ fixed, solve the innovator problem backwards (there is not much to solve here other than computing the value of an idea recursively). Get a sequence of $\{v_\text{I}\}$.

   i. Given a sequence of $\{s\}$ and $\{v_\text{I}\}$, derive skilled labor devoted to entry $h_{\text{F},t} = \left\{\frac{\gamma \cdot v_{\text{I},t}}{s_t}\right\}^{\frac{1}{1-\lambda}}$

   ii. Use $\{h_\text{F}\}$ to derive new sequence $\{a_\text{F}^\text{new}\}$ using the resource constraint related to the creation of new ideas

   iii. Update $a_F^\text{next} = relax * \{a_\text{F}\} + (1 - relax) * \{a_\text{F}^\text{new}\}$

   iv. Go back to a) until convergence occurs, $a_\text{F} \approx a_\text{F}^\text{new}$.

(b) Update $\{\tau\}$ (I am still holding $\{z\}$ fixed but $\{\tau\}$ changes nonetheless because the frontier a_F is moving) and go back to a), stop when $\tau^{next} \approx \tau$, i.e., τ has converged to previous guess.

4. Go back to 2. using a new updated guess $\{a_\text{F}\}$. Keep iterating from 2. – 4. until convergence.

5. Now you have two sequences $\{h_\text{D}, h_\text{F}\}$ that are consistent with optimization and resource constraints in each sector. The only thing that remains to be checked is whether the sequences are consistent with market clearing for skilled labor. Most likely they are not and so we compute excess demand functions and update the sequence of skill prices gently in the right direction

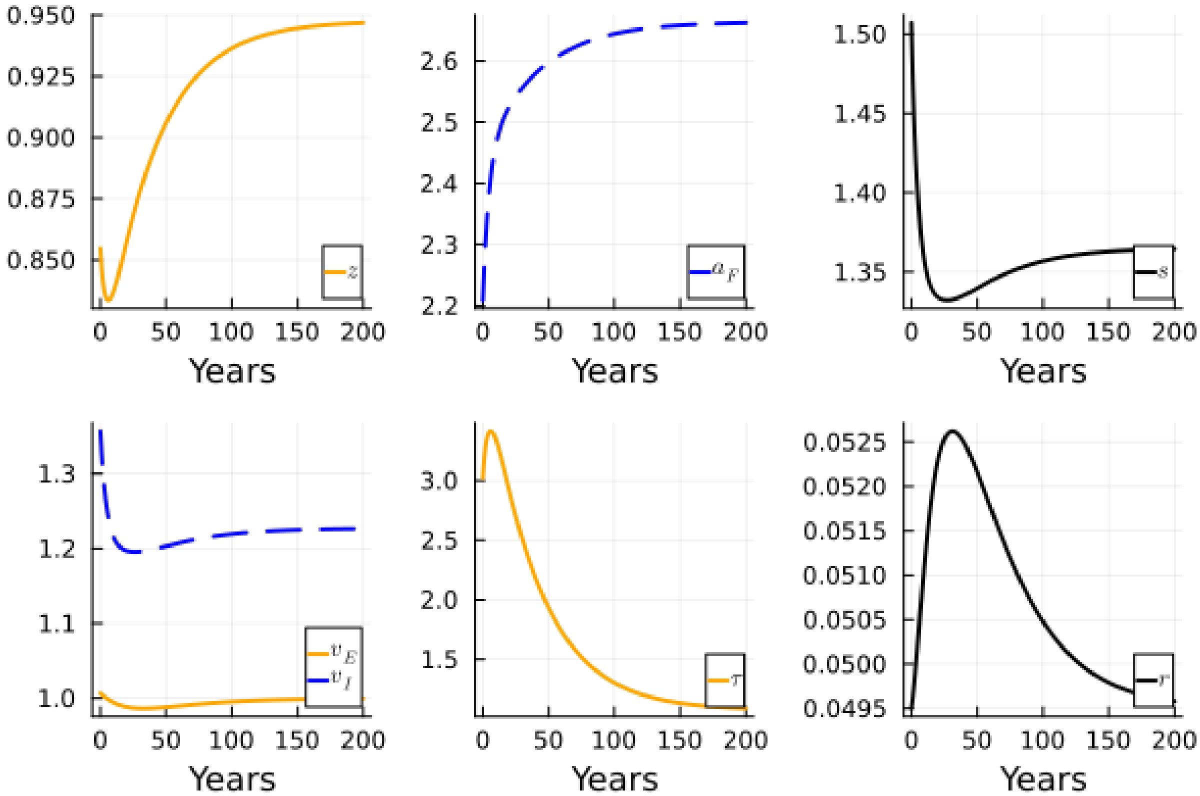
$$s^\text{next} = relax * s + (1 - relax) * s * e^{\theta_s[(h_\text{D}+h_\text{F})-h_\text{tot}]}$$

6. Go back to step 1. and start all over again for the new skill price guess up until convergence, i.e., $s^{next} \approx s$, which coincides with skilled labor market clearing $h_\text{tot} = h_\text{D} + h_\text{F}$

The figure below computes transition dynamics for a closed economy that is calibrated as in the main text. I shock the relative skilled labor supply from .15 to .18 by raising the skilled labor supply in the form of an unanticipated persistent shock at time zero. The transitions are then computed using the previous algorithm under perfect foresight. Figure 15 plots the results.

Three points are noteworthy. First, note that innovation responds quicker than technology adoption in the sense that convergence to its long-run steady state is faster. This is explained by the advantage of backwardness, which naturally leads to a lagged response of technology adoption. Second, the dynamics of the skill premium are non-monotone, and a fast initial fall in the price of skill is followed by slow but small rise in the price of skill later on. The increase in the price of skill, which ultimately is a function of the demand for skilled labor in this general equilibrium model, is directly related to increasing and lagged technology adoption effort. Third, the value of entry in the production sector is hardly moving, and deviates from its long-run value by less than 1% highlighting that the simplification is unlikely to matter quantitatively for the transition dynamics. The evolution of the value of entry into the research sector roughly follows the skill premium, and the difference in the two series is due to congestion whenever $\lambda < 1$.

Figure 15. Skilled Labor Expansion in Autarky



In this exercise I increase the relative skilled labor supply from .15 to .18. All of this increase occurs at time zero.

B.2 Open Economy

To solve for transition dynamics in the open economy, I follow similar steps. Note that the problem of production sector firms is unchanged. Firms in the research sector now need to take into account both opportunities abroad and competition at home when they consider whether to pay the fixed cost of producing an idea.

B.2.1 Innovation

The value of an innovation generalizes as follows in the open economy where I focus on the case of differences in size captured by country-specific unskilled labor endowments L_c. I consider a version where I use L from the home economy as the normalizing factor so $b_k = \frac{L_k}{L}$ relative to the home economy. The value of an idea reads

$$\begin{aligned}
V_{\mathrm{I}}\left(r+\delta_{\mathrm{I}}\right) &= \sum_c e^{-\int_t^{t+\tau_c}(r_v+\delta_{\mathrm{I}})dv}\pi_{\mathrm{I},c,t+\tau}\cdot\left[1+\dot{\tau}_{c,t}\right]+\dot{V}_{\mathrm{I}} \\
&= \sum_c e^{-\int_t^{t+\tau_c}(r_v+\delta_{\mathrm{I}})dv}\frac{\alpha L_{\mathrm{P},c,t+\tau_c}w_{L,c,t+\tau_c}}{A^{\mathrm{W}}_{\mathrm{F},t+\tau_c}z_{c,t+\tau_c}}\cdot\left[1+\dot{\tau}_{c,t}\right]+\dot{V}_{\mathrm{I}} \\
&= \sum_c e^{-\int_t^{t+\tau_c}(r_v+\delta_{\mathrm{I}})dv}\frac{\alpha l_{\mathrm{P}}w_{L,c,t+\tau_c}b_c}{A^{\mathrm{W}}_{\mathrm{F},t+\tau_c}z_{c,t+\tau_c}}L\cdot\left[1+\dot{\tau}_{c,t}\right]+\dot{V}_{\mathrm{I}},
\end{aligned}$$

where I used the fact that the allocation of unskilled labor in production vis-a-vis entry is kept constant.

Next, use the normalization $v_{\mathrm{I},k,t} := \frac{V_{\mathrm{I}}}{w_{L,k}\left(A^{\mathrm{W}}_{\mathrm{F}}\right)^{-\phi}}$. Note that the normalization is country-specific because wages are. The normalized value of an idea reads

$$\begin{aligned}
v_{\mathrm{I},k,t}\left(r_k-g_{w_{L,k}}+\delta_{\mathrm{I}}+\phi g_{\mathrm{F}}\right) &= \sum_c \frac{e^{-\int_t^{t+\tau_c}(r_k+\delta_{\mathrm{I}})dv}}{w_{L,k,t}}\frac{\left(A^{\mathrm{W}}_{\mathrm{F},t}\right)^{\phi}}{A^{\mathrm{W}}_{\mathrm{F},t+\tau_c}z_{c,t+\tau_c}}\alpha L_{\mathrm{P},c,t+\tau_c}w_{L,c,t+\tau_c}\cdot\left[1+\dot{\tau}_{c,t}\right]+\dot{\tilde{v}}_{\mathrm{I}} \\
&= \sum_c \frac{e^{-\int_t^{t+\tau_c}(r_k+\delta_{\mathrm{I}})dv}}{w_{L,k,t}}\left(\frac{A^{\mathrm{W}}_{\mathrm{F},t}}{A^{\mathrm{W}}_{\mathrm{F},t+\tau_c}}\right)^{\phi}\frac{L_{t+\tau_c}\alpha l_{\mathrm{P}}w_{L,c,t+\tau_c}}{\left(A^{\mathrm{W}}_{\mathrm{F},t+\tau_c}\right)^{1-\phi}z_{c,t+\tau_c}}b_c\cdot\left[1+\dot{\tau}_{c,t}\right]+\dot{\tilde{v}}_{\mathrm{I}} \\
&= \sum_c e^{-\int_t^{t+\tau_c}\left(r_k-g_{w_L}+\delta_{\mathrm{I}}+\phi g_{\mathrm{F}}\right)dv}\left(\frac{\alpha l_{\mathrm{P}}}{a^{\mathrm{W}}_{\mathrm{F},t+\tau_c}z_{c,t+\tau_c}}\frac{w_{L,c,t+\tau_c}}{w_{L,k,t+\tau_c}}\right)\cdot b_c\cdot\left[1+\dot{\tau}_{c,t}\right]+\dot{\tilde{v}}_{\mathrm{I}}
\end{aligned}$$

where the second line uses the fact that I assumed that firm entry in production stays at its long-run steady state value, i.e., $l_{\mathrm{P},t}=l_{\mathrm{P}}$.

The reader will note that the expression is almost identical to the closed economy one, except I have to sum over all countries and take into account that wages w_c, waiting times τ_c, adoption gaps z_c, and weights b_c are country specific.

A country-specific free entry condition holds

$$\tilde{v}_{\mathrm{I},k,t} = \frac{1}{\gamma_k}h^{1-\lambda}_{\mathrm{F},k,t}s_{k,t},$$

where a binding inequality is guaranteed as long as $\lambda < 1$.

The evolution of the technological frontier follows from the free entry equilibrium and the entry technology

$$\begin{aligned}\frac{\dot{A}_{\mathrm{F}}^{\mathrm{W}}}{A_{\mathrm{F}}^{\mathrm{W}}} &= \sum_c \frac{A_{\mathrm{F},c}}{A_{\mathrm{F}}^{\mathrm{W}}} \frac{\dot{A}_{\mathrm{F},c}}{A_{\mathrm{F},c}} \\ &= \sum_c \chi_c \left\{ \frac{\gamma_c \left(A_{\mathrm{F}}^{\mathrm{W}}\right)^{\phi} L h_{\mathrm{F},c}^{\lambda} b_c}{A_{\mathrm{F},c}} - \delta_{\mathrm{I}} \right\} \\ g_{\mathrm{F}}^{\mathrm{W}} &= \sum_c \frac{\gamma_c h_{\mathrm{F},c}^{\lambda} b_c}{a_{\mathrm{F}}^{\mathrm{W}}} - \delta_{\mathrm{I}}.\end{aligned}$$

The normalized technological frontier $a_{\mathrm{F}}^{\mathrm{W}} := \left(\frac{A_{\mathrm{F}}^{\mathrm{W}}}{L^{\frac{1}{1-\phi}}}\right)^{1-\phi}$ thus evolves according to

$$\begin{aligned}g_{a_{\mathrm{F}}^{\mathrm{W}}} &= (1-\phi)\, g_{\mathrm{F}}^{\mathrm{W}} - g_L \\ &= (1-\phi) \left\{ \frac{\sum_c \gamma_c h_{\mathrm{F},c}^{\lambda} b_c}{a_{\mathrm{F}}^{\mathrm{W}}} - \left(\delta_{\mathrm{I}} + \frac{g_L}{1-\phi}\right) \right\}.\end{aligned}$$

Note that this law of motion in the open economy is independent of the split of ownership of patents assumed at time zero, i.e., the share ζ does not show up.

Initial Conditions. In contrast, the evolution of country-specific normalized technology depends on ζ as follows. Note that at time $t \to 0$ where 0 demarcates the time when markets become integrated, the normalized frontier technology level in the West reads

$$\begin{aligned}\left(\frac{A_{\mathrm{F},0+\Delta}}{L^{\frac{1}{1-\phi}}}\right)^{1-\phi} &= \left(\frac{\zeta A_{\mathrm{F},0-\Delta}}{L^{\frac{1}{1-\phi}}}\right)^{1-\phi} \\ &= (\zeta)^{1-\phi}\, a_{\mathrm{F},0-\Delta},\end{aligned}$$

and similarly the normalized level of frontier technology in the foreign country reads $(1-\zeta)^{1-\phi}\, a_{\mathrm{F},0-\Delta}$ where $\Delta \to 0$. Note that both countries lose some of their ideas due to duplication.

Recall that I define the autarky frontier level relative to an initial real wage gap $\omega := \frac{w_{0-\Delta}}{w_{0-\Delta}^{*}}$, I use the following relationship to pin down the implied level of frontier technology in the East

$$A_{\mathrm{F},0-\Delta}^{*} = \frac{z_{0-\Delta}^{*}}{z_{0-\Delta}} A_{\mathrm{F},0-\Delta}$$

where $z_{0-\Delta}^{*}$ follows from solving the autarky equilibrium in the emerging market. Because I assume the emerging market has such a small research productivity in the integrated equilibrium, a counterfactually low level of income is implied. To fix this, I simply assume that copying works exactly like innovation except the fundamental productivity is higher than in innovation $\gamma^{*,copy} > \gamma^{*}$. It is easy to see that the allocation and the skill premium are unrelated to γ^{*}, which means I can simply solve the model for the low research productivity, and scale the frontier technology and real wage by the appropriate factor to obtain a wage gap ω. Some high level of $\gamma^{*,copy}$ will be exactly consistent with this wage gap.

Next, note that while the adoption gap in the advanced economy does not jump, the adoption gap in the emerging market changes discretely since it is defined relative to the frontier, which shifts out from the point of view of the emerging market. Formally,

$$z^*_{0+\Delta} = \frac{z^*_{0-\Delta} A^*_{\mathrm{F},0-\Delta}}{A^{\mathrm{W}}_{\mathrm{F},0+\Delta}}.$$

The evolution of this normalized technology level for either country over time changes slightly due to the global research externality

$$\frac{\dot{A}_{\mathrm{F},c}}{A_{\mathrm{F},c}} = \frac{\gamma_c h^{\lambda}_{\mathrm{F},c} b_c}{\chi_c a^{\mathrm{W}}_{\mathrm{F}}} - \delta_{\mathrm{I}},$$

where $\frac{A_{\mathrm{F},c}}{A^{\mathrm{W}}_{\mathrm{F}}} = \chi_c = \left(\frac{a_{\mathrm{F},c}}{a^{\mathrm{W}}_{\mathrm{F}}}\right)^{\frac{1}{1-\phi}}$ so $\chi_c a^{\mathrm{W}}_{\mathrm{F}} = (a_{\mathrm{F},c})^{\frac{1}{1-\phi}} \left(a^{\mathrm{W}}_{\mathrm{F}}\right)^{-\frac{\phi}{1-\phi}}$. In terms of normalized frontier growth rates I have

$$g_{a_{\mathrm{F},c}} = (1-\phi)\left[\frac{\gamma_c}{\chi_c}\frac{h^{\lambda}_{\mathrm{F},c} b_c}{a^{\mathrm{W}}_{\mathrm{F}}} - \left(\delta_{\mathrm{I}} + \frac{g_L}{1-\phi}\right)\right].$$

Capital Accumulation. Finally, note that the problem is simplified by assuming that a share of domestic final goods production is re-invested at a constant rate χ_s, which is the same across countries and consistent with the long-run equilibrium interest rate. This means that $1-\chi_s$ is the share of final output consumed.

Note that this share is sensitive to the initial allocation of patent ownership in the integrated equilibrium, ζ. At time zero, a share $Y^*\frac{\sigma-1}{\sigma}\alpha(1-\alpha)\zeta(1-\chi_s)$ of output in the East is consumed by the West, and vice versa a share $Y\frac{\sigma-1}{\sigma}\alpha(1-\alpha)\zeta(1-\chi_s)$ of output in the West is consumed by the East. This also means that some domestically used capital is now held by other countries. However, for the equilibrium dynamics of innovation and technology adoption this split is irrelevant. After understanding this simplification, the following algorithm applies.

Algorithm.

1. Guess $\{s_{c,t}\}$ for each country, holding $\{a^{\mathrm{W}}_{\mathrm{F},t}\}$ fixed.
2. Solve optimal adoption choice backward, and $\{\{\tilde{z}_{c,t}\},\{r_{c,t}\}\}$ forward. Iterate until convergence for both countries.

(a) Compute $\{\tau, \tau^*\}$ and $\{w, w^*\}$ taking $\{z, z^*, k, k^*\}$ as given for each country.

(b) Solve innovator's problem

i. Compute the value of innovation in each country recursively $\{v_{\mathrm{I}}, v_{\mathrm{I}}^*\}$

ii. Implies equilibrium research effort based on free entry into research in each country $\{h_{\mathrm{F}}, h_{\mathrm{F}}^*\}$

iii. Compute evolution of aggregate technological frontier using $\{h_{\mathrm{F}}, h_{\mathrm{F}}^*\}$.

iv. Go back to i) and iterate up until $\left(a_{\mathrm{F}}^{\mathrm{W}}\right)^{n+1} \approx \left(a_{\mathrm{F}}^{\mathrm{W}}\right)^{n}$ where n stands for n^{th} iteration. Update the new guess gently $\left(a_{\mathrm{F}}^{\mathrm{W}}\right)^{next} = relax * \left(a_{\mathrm{F}}^{\mathrm{W}}\right)^{old} + (1 - relax) * \left(a_{\mathrm{F}}^{\mathrm{W}}\right)^{new}$.

(c) Go back to a) and update $\{\tau, \tau^*\}$ and $\{w, w^*\}$ gently. Note that while I treat $\{z, z^*, k, k^*\}$ as fixed for now, waiting times and wages change since they also depend on the evolution of the technological frontier. Iterate till convergence. This always involves the inner loop i) – iv) and the outer loop a) – c).

3. Given a new solution for the evolution of the technological frontier, return to 2. and iterate on all previous loops up until convergence in $\{z, z^*, k, k^*, a_{\mathrm{F}}^{\mathrm{W}}\}$.

4. Finally check skilled labor market clearing in each country, and use the excess demand function to update the skill premium in the right direction. Iterate over all loops up until the sequence of skill premia has converged in each country.

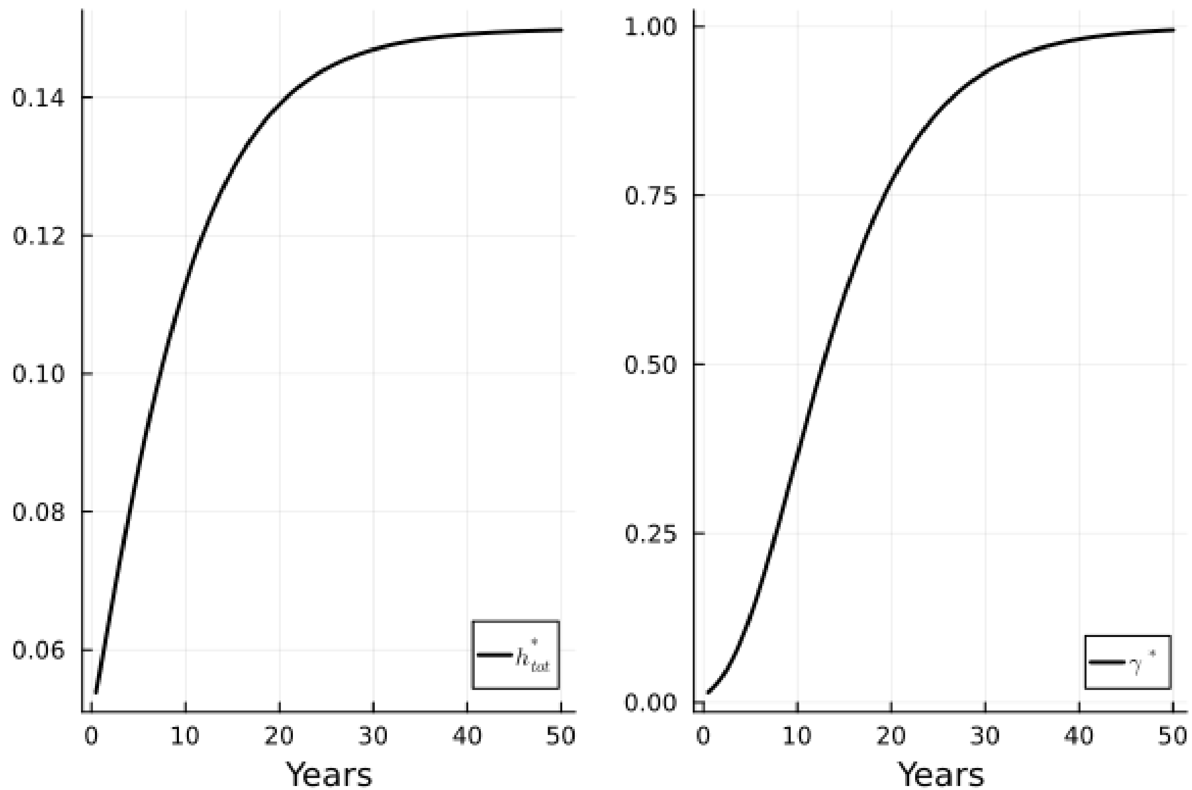
Full convergence. In the case of full convergence I let the research productivity and skill endowment of the poor country converge to the level of the rich country, i.e., ,$\gamma^* \rightarrow \gamma, h_{\text{tot}}^* \rightarrow h_{\text{tot}}$. I assume that the emerging market fully converges to the fundamentals of the advanced economy within 30 years. I impose a process that mimics convergence in the Solow model, and uses the formula $h_{\text{tot},t}^* = \left\{h_{\text{tot}}^{1-\alpha_{\text{conv}}}\left(1 - e^{-(1-\alpha_{\text{conv}})t}\right) + \left(h_{\text{tot},0}^*\right)^{1-\alpha_{\text{conv}}} e^{-(1-\alpha_{\text{conv}})t}\right\}^{\frac{1}{1-\alpha_{\text{conv}}}}$ where α_{conv} governs the speed of convergence, and an analogous expression applies to γ_t^*. I assume $\alpha_{\text{conv}} = .87$, which induces convergence patters as depicted in figure 16.

The previous computational algorithm applies almost unchanged. Changes in skilled labor endowments shows up in the market clearing condition. And the free entry condition into research needs to be updated slightly $\tilde{v}_{\mathrm{I},k,t} = \frac{1}{\gamma_{k,t}} h_{\mathrm{F},k,t}^{1-\lambda} s_{k,t}$ where the exogenous research productivity is now time varying for the emerging market. Laws of motion of idea creation also need to be updated to take account of the changing research productivity.

B.3 Initial Idea-Ownership in Integrated Equilibrium

The issue here is that in autarky countries may invent identical technology, which begs the question: who owns the technology in the initial integrated equilibrium. I say initial because in the long-run the

Figure 16. Full Convergence in East



initial allocation doesn't matter and the share of ideas is fully pinned down by differences in research specialization and country size. I next offer a micro-foundation that pins down this initial ownership allocation.

Suppose all ideas only invented by the West are held by the West. Ideas that overlap in the interval $[0, A^*_{\text{F}}]$ are split across East and West as follows. Suppose that the same idea across countries differs by some quality-metric $q > 0$ such that the effective quality is $q_c = 1 + \epsilon_c + u_c$. The parameter $\epsilon_c > 0$ is a country-specific quality shifter. Let the shock u be bounded between $[-1, 1]$, mean-zero, and with a vanishingly small variance. Further, suppose that an innovator needs to pay a fixed cost f_o for as long as the idea exists. As argued in Acemoglu et al. (2018), such a setup converges in the limit to the simple baseline monopoly pricing when $f_o \to 0, \mathbb{V}[u] \to 0$, and $\epsilon_c \to 0 \, \forall c$, but the ratioo $\frac{\epsilon_c}{\epsilon_k} \to b > 0$ converges to something bounded away from zero. This sustains an ownership structure at time zero where the advanced economy's share of global patents equals ζ, with ζ being a function of b and the initial gap $\frac{A_{\text{F}}}{A^*_{\text{F}}}$.

C Empirical Appendix

C.1 Calibrating λ

To calibrate the model I have to pin down λ,which does not have a direct antecedent in the literature. As explained in the main body of the paper the parameter calibrated in Jones (1995) is slightly different. In principal, cross-country research specialization among a set of countries with similar research productivities γ allows me to identify λ since the theory implies that the steady state share of ideas is a function of specialization in research, and total country size measured in terms of unskilled labor

$$\log \chi_c = \alpha_0 + \lambda \log H_{\text{F}} + (1 - \lambda) \log L.$$

I use data from the OECD, and combine it with Barro and Lee's skill measure. I proxy the steady state patent share (a stock) with average patents over the period from 2011-2019 (a flow). Note that patents are an imperfect proxy that likely understate innovative activity and technological change. As long as the paten share is a linear function of the larger share of actual ideas develop in some country, the approach works nonetheless.

The measure of patents is based on patents captured by the Patent Cooperation Treaty (PIC), which are global patents that simultaneously protect intellectual property among several countries. This selection helps make patents comparable across countries, and is closest to the notion of "global ideas" in the paper. Clearly, patents are a crude measure of innovative activity, and there is no reason to believe the economy is exactly in steady state. Nor should we think that there is no heterogeneity in innovative productivity among rich countries in Europe. Overall, the approach is imperfect.

Perhaps surprisingly, the empirical results are extremely sensible without any massaging by using just a simple simple linear regression. After making an ad-hoc adjustment for the bias induced by unobserved heterogeneity in patenting productivity,[67] I arrive at an estimate $\lambda = .9$ which is a reasonable value for a first pass.

I provide robustness results next where I use total employment instead of unskilled labor but the estimation is otherwise unchanged.

[67] Sampson (2023) argues heterogeneity in innovative productivity is a lot smaller among advanced economies relative to the gap between advanced economies and emerging markets.

Figure 17. Regression for λ

	log patents	log patents	log patents	log patents	log patents
log researchers	1.093***		0.980**		0.878**
	(0.0644)		(0.335)		(0.390)
log employment		1.070***	0.113	1.093***	0.122
		(0.0745)	(0.353)	(0.0657)	(0.390)
log (research/employment)				0.980**	
				(0.335)	
N	13	13	13	13	13
R^2	0.961	0.944	0.961	0.961	

The outcome variable is log patents. Employment is defined as total unskilled labor L consistent with the theory. Robust standard error are computed, * p<0.10, ** p<0.05, *** p<0.01. The last column runs a constrained regression enforcing that the slope coefficients have to add up to one.

Table 2. Regression for λ robustness

	log patents	log patents	log patents	log patents
log researchers	1.093***		0.942**	
	(0.0644)		(0.360)	
log employment		1.074***	0.152	1.093***
		(0.0746)	(0.383)	(0.0654)
log (research/employment)				0.942**
				(0.360)
N	13	13	13	13
R^2	0.961	0.945	0.961	0.961

The outcome variable is log patents. The main difference to the previous table is that employment is now total employment. Robust standard error are computed, * p<0.10, ** p<0.05, *** p<0.01. The last column runs a constrained regression enforcing that the slope coefficients have to add up to one.

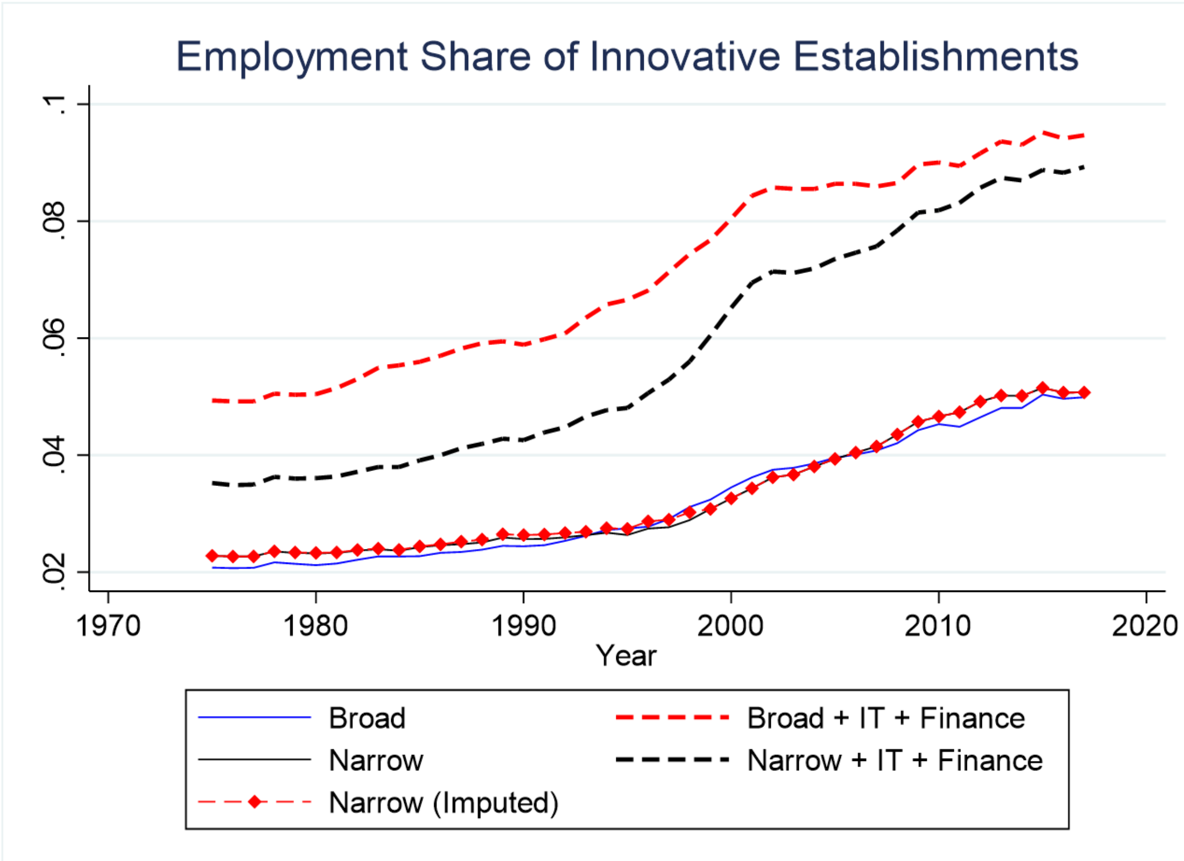
C.2 IAB Data

The IAB data is provided by the Institute of Employment Research in Nuremberg. While data access is restricted, the application process is fairly straightforward and open to researchers. The dataset combines worker level data including wage, skill, and demographics with establishment level data including total employment and sectoral classification.

Sectoral Classification. The model features a two-sector structure so I classify establishments into innovative and production establishments. I use different definitions for robustness. Figure 18 plots the evolution of the aggregate employment share of innovative establishments.

The sectoral classifications are defined in the do-file changing_innovation_activity.do. Below is a

Figure 18. Employment Share in West-Germany



Employment shares in production and research sector are based on IAB data for West Germany with additional details in the appendix. This classification can be carried out using different sectoral codes (broad vs. narrow where the narrower codes are unavailable for earlier periods) and a question arises whether IT and finance belong into the research or production sector.

short summary of the sectors that appear in the definition.

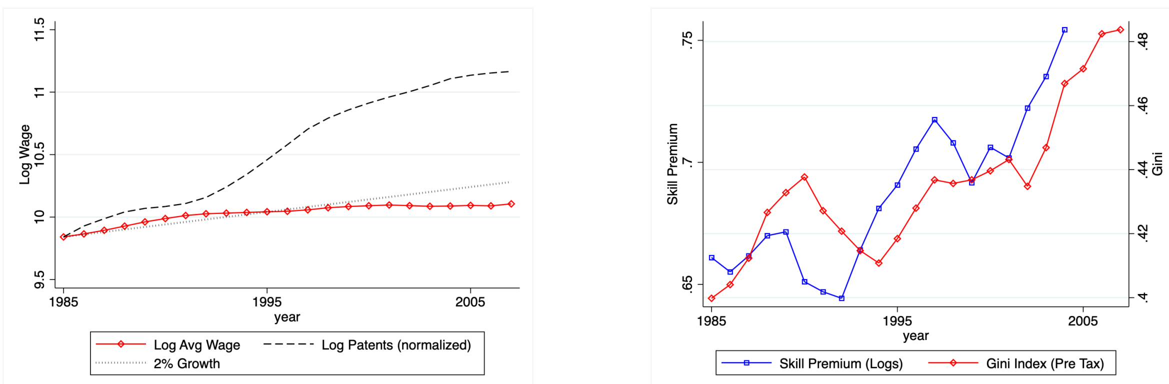
An important statistic omitted in the main part of the paper is the increase in the skill premium plotted in figure 20. Note that the IAB data is not suitable to document this increase due to heavy top coding. The skill premium appears stable in the beginning of the sample period and then widens substantially since the mid 90s.

Figure 19. Sectoral classification

Industry Digit	variable_na	sector code	label	baseline	baseline plus Finance and IT
5-Steller	(w93_5)	65110	Zentralbanken	0	1
5-Steller	(w93_5)	65121	Kreditbanken einschliesslich Zw	0	1
5-Steller	(w93_5)	65124	Genossenschaftliche Zentralban	0	1
5-Steller	(w93_5)	65126	Realkreditinstitute	0	1
5-Steller	(w93_5)	65127	Kreditinstitute mit Sonderaufg	0	1
5-Steller	(w93_5)	65210	Institutionen fuer Finanzierung	0	1
5-Steller	(w93_5)	65220	Spezialkreditinstitute	0	1
5-Steller	(w93_5)	65231	Kapitalanlagegesellschaften	0	1
5-Steller	(w93_5)	65233	Sonstige Finanzierungsinstitut	0	1
5-Steller	(w93_5)	67110	Effekten- und Warenterminboerse	0	1
5-Steller	(w93_5)	67120	Effektenverm und	0	1
5-Steller	(w93_5)	67130	Sonstige mit dem Kreditgewerbe	0	0
5-Steller	(w93_5)	72201	Softwareberatung	0	1
5-Steller	(w93_5)	72202	Softwareentwicklung	0	1
5-Steller	(w93_5)	72301	Datenerfassungsdienste	0	1
5-Steller	(w93_5)	72302	Datenverarbeitungs- und Tabell	0	1
5-Steller	(w93_5)	72303	Bereitstellungsdienste fuer Tei	0	1
5-Steller	(w93_5)	72304	Sonstige Datenverarbeitungsdie	0	1
5-Steller	(w93_5)	72400	Datenbanken	0	1
5-Steller	(w93_5)	72500	Instandhaltung und Reparatur v	0	0
5-Steller	(w93_5)	72601	Informationsvermittlung	0	1
5-Steller	(w93_5)	72602	Mit der Datenverarbeitung verb	0	1
5-Steller	(w93_5)	73101	Forschung und Entwicklung im	1	1
5-Steller	(w93_5)	73102	Forschung und Entwicklung im	1	1
5-Steller	(w93_5)	73103	Forschung und Entwicklung im	1	1
5-Steller	(w93_5)	73104	Forschung und Entwicklung im	1	1
5-Steller	(w93_5)	73105	Forschung und Entwicklung im	1	1
5-Steller	(w93_5)	73201	Forschung und Entwicklung im	1	1
5-Steller	(w93_5)	73202	Forschung und Entwicklung im	1	1
5-Steller	(w93_5)	74111	Rechtsanwaltskanzleien mit Not	0	0
5-Steller	(w93_5)	74112	Rechtsanwaltskanzleien ohne No	0	0
5-Steller	(w93_5)	74113	Notariat	0	0
5-Steller	(w93_5)	74114	Patentanwaltskanzleien	1	1
5-Steller	(w93_5)	74115	Sonstige Rechtsberatung	0	0
5-Steller	(w93_5)	74121	Praxen von Wirtschaftspruefern,	0	0
5-Steller	(w93_5)	74122	Praxen von vereidigten Buchprue	0	0
5-Steller	(w93_5)	74131	Marktforschung	1	1
5-Steller	(w93_5)	74132	Meinungsforschung	0	0
5-Steller	(w93_5)	74141	Unternehmensberatung	1	1
5-Steller	(w93_5)	74142	Public-Relations-Beratung	1	1
5-Steller	(w93_5)	74151	Beteiligungsgesellschaften mit	0	1
5-Steller	(w93_5)	74152	Sonstige Beteiligungsgesellsch	0	1
5-Steller	(w93_5)	74153	Geschlossene Immobilienfonds m	0	0
5-Steller	(w93_5)	74154	Geschlossene Immobilienfonds m	0	0
5-Steller	(w93_5)	74155	Komplementaergesellschaften	0	1
5-Steller	(w93_5)	74156	Verwaltung und Fuehrung von	1	1
5-Steller	(w03_5)	65110	Zentralbanken	0	1
5-Steller	(w03_5)	65121	Kreditbanken einschliesslich Zwei	0	1
5-Steller	(w03_5)			0	
5-Steller	(w03_5)	65124	Genossenschaftliche Zentralbanken	0	1
5-Steller	(w03_5)	65126	Realkreditinstitute	0	1
5-Steller	(w03_5)	65127	Kreditinstitute mit Sonderaufgabe	0	1
5-Steller	(w03_5)	65210	Institutionen fuer Finanzierungsl	0	1
5-Steller	(w03_5)	65220	Spezialkreditinstitute	0	1
5-Steller	(w03_5)	65231	Kapitalanlagegesellschaften	0	1
5-Steller	(w03_5)	65233	Sonstige Finanzierungsinstitution	0	1
5-Steller	(w03_5)	67110	Effekten- und Warenboersen	0	1
5-Steller	(w03_5)	67120	Effektenvermittlung und	0	1
5-Steller	(w03_5)			0	
5-Steller	(w03_5)	72210	Verlegen von Software	0	1
5-Steller	(w03_5)	72221	Softwareberatung	0	1
5-Steller	(w03_5)	72222	Entwicklung und Programmierung vo	0	1
5-Steller	(w03_5)	72223	Sonstige Softwareentwicklung	0	1
5-Steller	(w03_5)	72301	Datenerfassungsdienste	0	1
5-Steller	(w03_5)	72303	Bereitstellungsdienste fuer Teiln	0	1
5-Steller	(w03_5)	72305	Sonstige Datenverarbeitungsdienst	0	1
5-Steller	(w03_5)	72400	Datenbanken	0	1
5-Steller	(w03_5)	72601	Informationsvermittlung	0	1
5-Steller	(w03_5)	72602	Mit der Datenverarbeitung verbund	0	1
5-Steller	(w03_5)	73101	Forschung und Entwicklung im	1	1
5-Steller	(w03_5)	73102	Forschung und Entwicklung im	1	1
5-Steller	(w03_5)	73103	Forschung und Entwicklung im	1	1
5-Steller	(w03_5)	73104	Forschung und Entwicklung im	1	1
5-Steller	(w03_5)	73105	Forschung und Entwicklung im	1	1
5-Steller	(w03_5)	73201	Forschung und Entwicklung im	1	1
5-Steller	(w03_5)	73202	Forschung und Entwicklung im	1	1
5-Steller	(w03_5)	74111	Rechtsanwaltskanzleien mit Notari	0	0
5-Steller	(w03_5)	74112	Rechtsanwaltskanzleien ohne Notar	0	0
5-Steller	(w03_5)	74113	Notariate	0	0
5-Steller	(w03_5)	74114	Patentanwaltskanzleien	1	1
5-Steller	(w03_5)	74115	Sonstige Rechtsberatung	0	0
5-Steller	(w03_5)	74121	Praxen von Wirtschaftsprueferinne	0	0
5-Steller	(w03_5)	74122	Praxen von vereidigten Buchpruefe	0	0
5-Steller	(w03_5)	74123	Praxen von Steuerberaterinnen und	0	0
5-Steller	(w03_5)	74124	Praxen von Steuerbevollmaechtigte	0	0
5-Steller	(w03_5)	74125	Buchfuehrung (ohne Datenverarbeit	0	0

Industry Di	variable_na	sector code	label	baseline	baseline plus Finance and IT
3-Steller	(w93_3)	651	Zentralbanken u. Kreditinst.	0	1
3-Steller	(w93_3)	652	So. Finanzierungsinstitute	0	1
3-Steller	(w93_3)	722	Softwarehaeuser	0	1
3-Steller	(w93_3)	723	Datenverarbeitungsdienste	0	0
3-Steller	(w93_3)	724	Datenbanken	0	0
3-Steller	(w93_3)	726	Verb Ttg. der Datenverarb.	0	0
3-Steller	(w93_3)	731	F&E Naturwissenschaft	1	1
3-Steller	(w93_3)	732	F&E Recht, Wirtschaft usw.	1	1
3-Steller	(w93_3)	741	Beratungsunternehmen	1	1
3-Steller	(w93_3)	742	Architektur- u. Ingenieurbuero	0	0
3-Steller	(w93_3)	743	Tech., physik. u. chem.	0	0
3-Steller	(w93_3)	744	Werbung	0	0
5-Steller	(w03_5)	74131	Marktforschung	1	1
5-Steller	(w03_5)	74132	Meinungsforschung	0	0
5-Steller	(w03_5)	74141	Unternehmensberatung	1	1
5-Steller	(w03_5)	74142	Public-Relations-Beratung	1	1
5-Steller	(w03_5)	74151	Managementtaetigkeiten von Holdi	1	1
5-Steller	(w03_5)	74152	Managementtaetigkeiten von sonst	1	1
5-Steller	(w03_5)	74153	Geschlossene Immobilienfonds mit	0	0
5-Steller	(w03_5)	74154	Geschlossene Immobilienfonds mit	0	0
5-Steller	(w03_5)	74155	Komplementaergesellschaften	0	1
5-Steller	(w03_5)	74156	Verwaltung und Fuehrung von	1	1
5-Steller	(w08_5)	62011	Entwicklung und Programmierung	0	1
5-Steller	(w08_5)	62019	Sonstige Softwareentwicklung	0	1
5-Steller	(w08_5)	62020	Erbringung von Beratungsleistunge	0	1
5-Steller	(w08_5)	62030	Betrieb von Datenverarbeitungsein	0	1
5-Steller	(w08_5)	62090	Erbringung von sonstigen Dienstle	0	1
5-Steller	(w08_5)	63110	Datenverarbeitung, Hosting und da	0	1
5-Steller	(w08_5)	63120	Webportale	0	0
5-Steller	(w08_5)	63910	Korrespondenz- und Nachrichtenbı	0	0
5-Steller	(w08_5)	63990	Erbringung von sonstigen Informat	0	0
5-Steller	(w08_5)	64110	Zentralbanken	0	1
5-Steller	(w08_5)	64191	Kreditbanken einschliesslich Zwei	0	0
5-Steller	(w08_5)	64192	Kreditinstitute des Sparkassensek	0	0
5-Steller	(w08_5)	64193	Kreditinstitute des Genossenschaf	0	0
5-Steller	(w08_5)	64194	Realkreditinstitute	0	1
5-Steller	(w08_5)	64195	Kreditinstitute mit Sonderaufgabe	0	1
5-Steller	(w08_5)	64196	Bausparkassen	0	0
5-Steller	(w08_5)	64200	Beteiligungsgesellschaften	0	1
5-Steller	(w08_5)	64300	Treuhand- und sonstige Fonds	0	1
5-Steller	(w08_5)	64910	Institutionen fuer Finanzierungsl	0	1
5-Steller	(w08_5)	64921	Spezialkreditinstitute (ohne Pfan	0	1
5-Steller	(w08_5)	64922	Leihhaeuser	0	0
5-Steller	(w08_5)	64991	Investmentaktiengesellschaften un	0	1
5-Steller	(w08_5)	64999	Sonstige Finanzierungsinstitution	0	1
5-Steller	(w08_5)	66110	Effekten- und Warenboersen	0	1
5-Steller	(w08_5)	66120	Effekten- und Warenhandel	0	1
5-Steller	(w08_5)	66190	Sonstige mit Finanzdienstleistung	0	0
5-Steller	(w08_5)	66210	Risiko- und Schadensbewertung	0	0
5-Steller	(w08_5)	66220	Taetigkeit von Versicherungsmakle	0	0
5-Steller	(w08_5)	66290	Sonstige mit Versicherungsdienstl	0	0
5-Steller	(w08_5)	66300	Fondsmanagement	0	1
5-Steller	(w08_5)	69101	Rechtsanwaltskanzleien mit Notari	0	0
5-Steller	(w08_5)	69102	Rechtsanwaltskanzleien ohne Nota	0	0
5-Steller	(w08_5)	69103	Notariate		0
5-Steller	(w08_5)	69104	Patentanwaltskanzleien	1	1
5-Steller	(w08_5)	69109	Erbringung sonstiger juristischer	0	0
5-Steller	(w08_5)	69201	Praxen von Wirtschaftsprueferinne	0	0
5-Steller	(w08_5)	69202	Praxen von vereidigten Buchpruefe	0	0
5-Steller	(w08_5)	69203	Praxen von Steuerbevollmaechtigte	0	0
5-Steller	(w08_5)	69204	Buchfuehrung (ohne Datenverarbei	0	0
5-Steller	(w08_5)	70101	Managementtaetigkeiten von Holdi	1	1
5-Steller	(w08_5)	70109	Sonstige Verwaltung und Fuehrung	1	1
5-Steller	(w08_5)	70210	Public-Relations-Beratung	1	1
5-Steller	(w08_5)	70220	Unternehmensberatung	1	1
5-Steller	(w08_5)	72110	Forschung und Entwicklung im	1	1
5-Steller	(w08_5)	72190	Sonstige Forschung und Entwicklur	1	1
5-Steller	(w08_5)	72200	Forschung und Entwicklung im	1	1
5-Steller	(w08_5)	73110	Werbeagenturen	0	0
5-Steller	(w08_5)	73120	Vermarktung und Vermittlung von	0	0
5-Steller	(w08_5)	73200	Markt- und Meinungsforschung	1	1

Figure 20. Growth, Patents, and Inequality in Germany



Data for patents comes from the Crios Patstata database, see Coffano and Tarasconi (2014). Wage data is computed based on the PWT version 09, combining real national gdp (not PPP) with their measure of the labor share and dividing thorough by the total population. Patents are normalized so that the wage level and patent level coincide in 1984. GDP per capita growth does better than wages, but still grows substantially below trend, leading to an overall growth slowdown. Data for the skill premium, denoted as $\log\left(\frac{w_H}{w}\right)$ where the wage rates are the price of one hour of skilled or unskilled labor, comes from the KLEMS data version 07. Skill here refers to college-educated workers, group 3 in the Klems data. I do not make additional adjustments for efficiency units within skill group, which does not change the broad pattern. See the discussion and adjustments made in Buera et al. (2022) who also use the Klems data.The Gini index is pre tax and taken from the World Inequality Database of Alvaredo et al. (2020).