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

Florian Trouvain
University of Michigan

Job Market Paper

November 6, 2022

[most recent version here]

Abstract

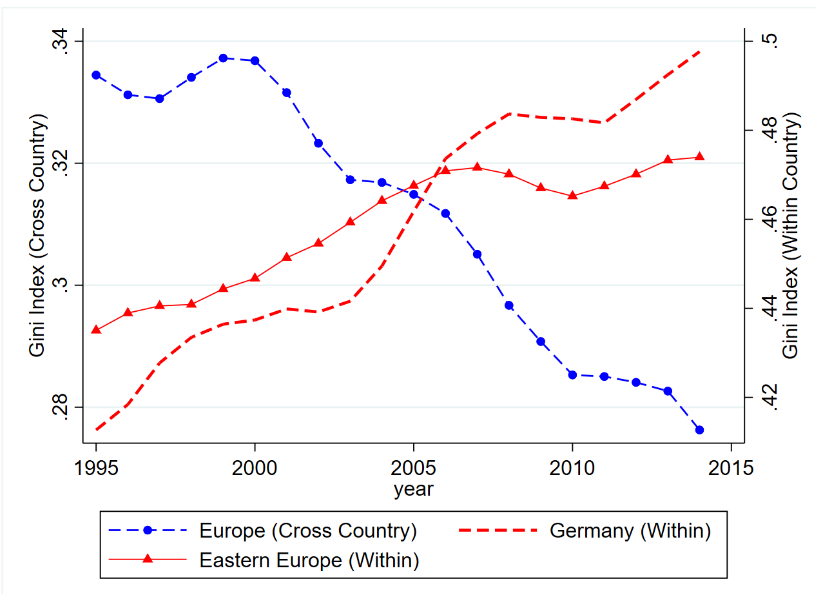
Economic Growth since the mid 1990s is characterized by i) declining cross-country inequality, ii) rising within-country inequality, and iii) overall weak growth in advanced economies. I provide a unifying explanation for these facts by developing a theory of long-run growth that focuses on the interaction of innovation and technology adoption in a globalized world. I model both activities as skill-intensive, and study how goods market integration with emerging markets shapes the returns to innovation vis-a-vis technology adoption. While the development of frontier technology in advanced economies is boosted by globalization, increasing innovation comes at the cost of rising inequality and reduced domestic technology adoption. When ideas are getting harder to find, the growth drag from reduced adoption dominates positive innovation effects, which explains slow TFP growth and stagnant wages for non-college workers in advanced economies. The mechanism is corroborated by cross-sectional evidence from German micro data, which leverages regional specialization in innovation vs. production together with the fall of the Iron Curtain.

*E-mail: ftrouv@umich.edu. 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. I am extremely grateful to my advisors for generous feedback, crucial advice, and patient guidance throughout the past 5 years: John Leahy, Dmitriy Stolyarov, Linda Tesar, and Brian Wu. I thank Dominick Bartelme and John Bound for pivotal encouragement and advice. For their insights I thank Andres Blanco, Mike Blank, Charlie Brown, Paco Buera, Max Dvorkin, Jonathan Eaton, Carlos Garriga, Josh Hausman, Elhanan Helpman, John Laitner, Emir Murathanoglu, Pablo Ottonello, B. Ravikumar, Paulina Restrepo-Echavarria, Hannah Rubinton, Juan Sanchez, Ana Maria Santacreu, Brit Sharoni, Yongs Shin, Sebastian Sotelo, Jagadeesh Sivadasan, and Mark Wright. I have greatly benefited from visiting the St. Louis Fed as Ph.D fellow, and I am grateful for detailed comments and generous hospitality. I thank the German Academic Scholarship Foundation for financial support.

1 Introduction

Three key features of economic growth from the mid 1990s up until the COVID-19 pandemic can be summarized as follows. First, cross-country income inequality has declined. Second, within-country income inequality has risen, both in advanced economies and emerging markets. And third, growth in advanced economies was slow, with real wages being stagnant for non-college workers, in contrast to fast per capita growth in emerging markets. Figure 1 illustrates global convergence and within-country divergence by plotting a cross-country and a within-country indexes Gini-indices over time. The plot focuses on Europe, where Eastern European economies represent emerging markets but similar plots could be produced for the world as a whole. While Eastern Europe experienced annual per capita growth of around 5%, Western Europe fared less well. For example, Germany grew at a rate below 1%, which I single out here as it will be the focus of my empirical application later on, but the growth and inequality patterns are similar across advanced economies.

Figure 1. 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 simple averages and cross country inequality is measured in terms of GDP per capita for each country using PWT V10.

In this paper I develop a model of long-run technological change that provides a unifying explanation for these *cross-country and within-country* patterns of growth. I introduce a technology adoption friction into an otherwise standard endogenous growth model with two types of labor, high skilled and production labor.

I use a two-sector structure where the research sector invents new technology that is embodied in capital goods. The production sector produces a final consumption good by combining capital and labor with technology from the research sector. I assume that incorporating novel technology in the production sector is a *costly and skill-intensive process*, which is a crucial assumption that shapes aggregate growth as I discuss.[1] The rate at which new technology is adopted is determined by firms in the production sector, which solve a dynamic problem. This gives rise to an equilibrium adoption gap, i.e. there is a lag between when a new technology is invented, and when it is used in the production sector. This gap is going to be a function of the endogenous allocation of skilled labor between the two sectors performing innovation and adoption related activities.

A central insight from the model is that the presence of an adoption friction leads to a novel complementarity between innovation and technology adoption. Innovators take into account that their ideas will become profitable only after they are adopted. The present discounted value of future profits thus directly depends on the time it takes for ideas to be adopted, which in turn is a function of the speed of adoption in the production sector. Ceteris paribus, higher adoption effort pushes up the net present value of innovation by reducing the waiting time of an innovator to sell their technology. In contrast, since both innovation and adoption are skill-intensive activities and draw on the same scarce resource, skilled labor, a factor market rivalry emerges. In the closed economy, the factor market rivalry is dominated by the complementarity between innovation and adoption so that the two activities move in lockstep. The intuition is that the innovation sector cannot "run away" from the production sector since the latter constitutes the innovators' client base.

This complementarity can break down in the open economy. Globalization can lead to uneven economic growth in advanced economies where the innovation sector and skilled labor gain, while adoption activity and production worker wages stagnate. This happens when advanced economies integrate with emerging markets and advanced economies have a comparative advantage in developing frontier technology. Market integration, by which I mean free trade in ideas and final goods, then changes the returns to innovation vis-a-vis technology adoption within advanced economies. This breaks the complementarity between innovation and adoption as I argue next.

First, goods market integration provides emerging markets with access to modern technology. This leads to fast technology adoption and strong catch-up growth, which reduces cross-country income inequality. Second, given that frontier technology is produced in advanced economies, fast technology adoption in emerging markets has a feedback effect on the returns to innovation in advanced economies: as more countries make use of modern technology, the profits that innovators reap from developing new ideas increase due to a simple market-size effect. High profits for innovators in advanced economies, and fast adoption in emerging markets, are thus two sides of the same coin. This leads to additional entry into innovation, and increases skilled labor demand in the research sector, which in turn pushes up the skill premium in advanced economies. However, due to a market clearing condition for skilled

[1] Adoption here implies the ability to use a capital good but the monopoly of the innovator is always protected.

labor, the expansion of the innovation sector must come at the cost of reducing technology adoption in the *domestic* production sector. This novel complementarity between innovation and adoption, and the extent to which it can be reversed in the open economy, is the major theoretical contribution of this paper. To be precise, innovation and adoption are still complementary but the complementarity is playing out on a *global scale* where fast adoption in the emerging markets raises the returns to innovation, while *locally,* factor market competition leads to brain drain in the production sector within rich countries.

A desirable feature of the theory is that divergence between innovation and adoption, and rising inequality unfolds only after integration between asymmetric countries where one party is the main supplier of innovation. This explains why globalization since the 1990s has different effects compared to the process of trade integration among rich countries since WW2. In a symmetric country model where both countries are equal in innovative capacity, the benefit of exporting ideas exactly cancels with competition from abroad, leaving the returns to innovation unchanged. In contrast, poor countries that only adopt and are unable to contribute to the technological frontier creates the bias that changes the returns to innovation in advanced economy. Comparative advantage in innovation is thus crucial for the argument to hold.

The theory leads to ex-ante ambiguous effects of globalization on aggregate growth in advanced economies. This ambiguity results from the fact that productivity depends on both innovation and adoption. Gains from temporarily faster growth of the technological frontier in an open economy can be fully undone by a lack of domestic technology adoption. This is in stark contrast to benchmark endogenous growth models, which tend to respond very positively to market size effects due to idea-based scale effects. While the model features scale effects in innovation as well, there are constant returns in technology adoption in the sense that doubling the production labor force requires twice the amount of skilled labor in adoption so as to keep the adoption gap unchanged. This is useful in two ways. First, it breaks the strong positive relationship between market integration and growth. Second, it avoids counterfactual scale effects *across countries at a point in time.*[2] That is, it avoids a counterfactual scenario where a large developing economy like India is richer than a small advanced economy like Belgium, because the total number of skilled workers is higher in India.

Another useful feature of the model is that it can explain rising inequality in emerging markets too, which is consistent with a rising within-country Gini index for all countries, rich and poor, as shown in figure 1.[3] Since technology adoption is a skill-intensive activity, rising returns to adoption in the integrated equilibrium from the point of view of the emerging market lead to higher demand for skilled labor and a rising skill premium.

The model finds a direct empirical counterpart in the growth experience of advanced economies

[2] Technology adoption is the key determinant of cross-country income differences, and adoption will be a function of skill intensity and not the total number of skilled workers.

[3] Goldberg and Pavcnik (2007) pointed out that inequality has been increasing in many developing economies after major trade liberalizations.

and emerging markets since the mid 1990s. I focus on the case of Germany, which produces frontier technology and experienced a large and sudden integration shock with Eastern Europe after the fall of the Iron Curtain. I document empirically the joint evolution of patenting activity and employment reallocation to "innovative" establishments, in combination with weak aggregate growth, stagnant wages, and rising inequality over the period from 1995 – 2015. I calibrate the model to assess whether it can generate both rising inequality and a productivity slowdown consistent with the data. I do so by focusing on goods market integration between Germany and Eastern Europe after the fall of the Iron Curtain in a simplified two-country setting. I compare the closed economy steady state with the open economy one. In the closed economy steady state I assume that Germany is on a balanced growth path which is characterized by a constant skill premium and constant allocation of skilled labor between innovation and adoption. At the time of integration, effective market size increases since Eastern Europe begins to adopt German technology, which induces additional entry into innovation and a reallocation of labor from adoption to innovation within Germany.

The model predicts a quantitatively large cumulative drop in TFP of 17%, relative to trend. Production worker wages and TFP are directly linked in the model, which allows the mechanism to explain stagnant real wages for the non-college workforce. At the same time, integration leads to cumulative wage gains for skilled labor of 11%, adding up to an increase in the skill premium of 33%. The model predicts an increase in the Gini index of 6pp, which accounts for 75% of the observed increase from 1995 to 2015 in Germany. Consistent with the data, employment in the innovation sector expands and boosts the development of frontier technology. The model predicts that this expansion comes at the cost of a rising domestic adoption gap. This leads to uneven effects of market integration with skilled labor and the emerging market as a whole benefiting, while production labor in rich countries loses in real terms. Aggregating up worker income within advanced economies predicts an aggregate growth slowdown of roughly .5% predicted over a twenty year period, i.e. 10 log points in total.

This growth slowdown is not *hard-wired* into the model. It depends crucially on the functional form of the adoption technology and on the strength of the dynamic knowledge spillover, a central parameter in any idea-based growth model. When introducing a stronger knowledge spillover, a limiting case being Romer (1990)'s initial formulation, market integration delivers gains for everyone. If, on the other hand, ideas "are getting harder to find" as in Jones (1995), a growth slowdown becomes possible. I use the recent estimate of Bloom et al. (2020) to pin down this parameter, which leads to strong diminishing returns in research activity, and means that my growth model is of the semi-endogenous kind. Given strong diminishing returns in research, reallocating skilled labor into innovation has only a modest positive effect on the technological frontier. Yet, the adverse effect on technology adoption can be large depending on how important the adoption of technology is for productivity growth. The model captures this relationship in a simple elasticity of TFP with respect to skilled labor devoted to technology adoption. I use a link implied by the model between cross-country inequality and technology adoption that allows me to pin down this parameter. Adoption is important, and the lack thereof explains the negative effects of market integration on advanced economies. Consistent with the adverse effect of globalization in my calibration is that this economy is inefficient. There are externalities both in innovation and adoption, and the de-central equilibrium features too little adoption in the closed economy. This inefficiency is amplified in the open economy.

In a final step I complement the analysis with additional evidence using administrative worker level data as well as patent data and population counts across German counties. I leverage regional specialization in innovation vs. production, together with the fall of the Iron Curtain, to provide regression-based evidence consistent with uneven and innovation-biased growth, and weak technology adoption. The idea is to take the two-sector structure and project it into space based on a region's initial specialization in innovation. In the period after market integration, growth is biased towards innovative, high-income regions, which experience relatively high wage and skilled labor growth. These patterns were completely reversed before 1995, with growth being biased in favor of laggard regions, a trend that is more broadly true across advanced economies in the early post war period. The empirical evidence thus corroborates the main point of the theory: globalization had a dramatic impact on the rate and distribution of economic growth across workers and regions, and a model with an endogenous adoption gap is well-suited to capture this heterogeneity in a tractable way.

The rest of the paper proceeds as follows: Section 1.1 situates the paper in the literature. Section 2 presents a model of innovation and adoption. Section 3 introduces the open economy version. Section 4 offers a quantitative exercise after calibrating and estimating key parameters of the model. Section 5 provides additional empirical evidence to support the central mechanism. Section 6 concludes.

1.1 Relationship to the literature

This paper relates to four different streams of the literature. First, the paper builds on and relates to the large literature on economic growth. I combine theories of innovation and growth, following Romer (1990) and Jones (1995), with Nelson and Phelps (1966)'s work on technology adoption. Recent work that models innovation and adoption jointly are Konig et al. (2021), building on König, Lorenz, and Zilibotti (2016), as well as Benhabib, Perla, and Tonetti (2021) and Sampson (2019). These papers have in common that they develop heterogeneous firm models where high productivity firms push out the technological frontier, while laggard firms learn from high productivity firms to improve their productivity. In contrast to their work, my model features a two-sector structure with innovation and production being distinct activities, as in Acemoglu et al. (2018). This gives rise to a novel complementarity on the market for ideas where fast adoption leads to more innovation. In contrast, I consider how innovation and adoption compete for skilled labor in general equilibrium, which allows to match patterns in the data that were out of reach for benchmark models, namely rising inequality and weak growth after market integration. The paper is also related to Sala-i-Martin and Barro (1997), Acemoglu, Aghion, and Zilibotti (2006), and Benhabib, Perla, and Tonetti (2014) which study models where laggard countries face a choice between adoption and innovation. I extend this line of work by considering how adoption impacts the return to innovation in advanced ones.

Second, a number of recent papers have studied the recent productivity slowdown. [4] One strand of this literature focuses on the effect of declining population growth on economic growth (Peters and Walsh, 2019; Jones, 2020; Hopenhayn, Neira, and Singhania, 2018; Engbom et al., 2019), which is negative due to scale effects in innovation. While I agree that this is a central force and build directly on the semi-endogenous growth model of Jones (1995), my theory highlights a new channel of weak technology adoption which I show is quantitatively powerful. This adoption margin also provides a micro-foundation for empirical work that finds weak technology adoption to be an important driver of slow productivity growth, see Andrews, Criscuolo, and Gal (2015). In addition, the model explains the rising share of innovative activity in the economy. Note that in the benchmark model of Jones (1995), falling population growth leads to a declining share of resources devoted to innovation.[5] In the data, however, patenting activity picked up, and regional economies specialized in innovation outperformed others, see Moretti (2012). The effect of globalization on innovation in my model resolves this tension. An alternative explanation for the productivity slowdown marries models of Schumpeterian growth (Aghion and Howitt, 1990; Grossman and Helpman, 1991b) with biased technology shocks that favor large incumbents and suppress competition, which in turn slows down productivity growth, see for instance De Ridder (2019), Akcigit and Ates (2019), and Aghion et al. (2019). The strong scale effects inherent in these theories mean that they have to abstract away from globalization or population growth.

Third, a vast literature in international trade models how openness and comparative advantage lead to sectoral specialization and interact with economic growth, see Feenstra (2015) for a textbook introduction. Acemoglu (2003) considers how openness interacts with the direction of technological change but abstracts away from technology adoption. Much of the literature on trade and growth suggests that market integration increases the long-run growth rate (Rivera-Batiz and Romer (1991), Sampson (2016), Grossman and Helpman (2018), Hsieh, Klenow, and Nath (2019) or Perla, Tonetti, and Waugh (2021). My theory is consistent with this work in that integration is pro-innovation, but it may not always be pro-growth from the point of view of an advanced economy. For emerging markets, integration is always growth-enhancing as access to technology improves, consistent with empirical evidence on cross-country differences in technology adoption (Comin and Hobijn, 2010a; Comin and Mestieri, 2014) and the large empirical literature on development and trade, see Irwin (2019) for a review. Recent work combining quantitative trade models with endogenous and semi-endogenous growth theory are Cai, Li, and Santacreu (2022), Somale (2021), and Lind and Ramondo (2022). This

[4]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.

[5]This is most easily seen in a version of the model of Jones (1995) without population growth. The assumption that ideas are getting harder to find leads to prohibitively high entry cost into research, and idea production disappears unless there is an offsetting positive labor supply shock.

strand of the literature builds on the influential work of Eaton and Kortum (1999) and tends to find pro-growth effects of market integration in multi-sector multi-country models.

Fourth, this paper relates to a large literature in labor economics that studies the rising skill premium, starting with the seminal work of Katz and Murphy (1992), Bound and Johnson (1992), and Krueger (1993).[6] In my model the skill premium not only matters as distributional accounting device but has a direct effect on productivity. A rising skill premium leads to less adoption effort in equilibrium, with adverse effects on low-skilled workers. This margin helps rationalize stagnant wage growth for non-college workers that is hard to obtain in the benchmark model of skill-biased technological change of Katz and Murphy (1992) due to the strong complementarity between low-skilled and high-skilled workers. [7] A related literature has focused on the task content of work and automation (Autor, Levy, and Murnane, 2003; Acemoglu and Restrepo, 2018a), and I introduce this as one of several extensions of the baseline model. In short, a more skill-intensive task content leads to less labor available for technology adoption so the two mechanisms can complement each other to generate wage stagnation and weak productivity growth.

2 A Tractable Theory of Innovation and Adoption

2.1 Environment

Household Problem: Time is continuous and there are three types of households in the economy, capitalists, high skilled workers, and production workers. Each group grows at a common exogenous rate g_L. Workers supply their labor inelastically which leads to an economy wide endowment of L efficiency units of production labor and H efficiency units of high skilled labor. Factors earn income at a wage rate w and w_H, respectively. I denote the relative price of skill, i.e. the skill premium, as $s = \frac{w_H}{w}$. Workers are hand-to-mouth agents that consume all their labor income instantly, while capitalists only earn returns from the assets they hold, following Angeletos (2007). This assumption leads to a constant aggregate saving rate in the economy in steady state and during transition periods.[8] Without loss of generality, I assume that the measure of capitalists is equal to L. Dynastic capitalists solve a forward-looking consumption-saving problem

[6] A related literature in international trade studies the impact of globalization on inequality. See Helpman, Itskhoki, and Redding (2010), Liu and Trefler (2008), Sampson (2014), or Burstein and Vogel (2017) for recent work with firm heterogeneity. See Wood (1994) and Leamer (1994) for Heckscher-Ohlin type models, and Helpman (2016) for an overview.

[7] Note that in the benchmark model of skill-biased technological change, biased productivity growth towards skilled labor raises wages *for all workers*, albeit for some more than others. This is inconsistent with observed wage stagnation for a large share of workers in advanced economies, see Acemoglu and Autor (2011) for a discussion of this point.

[8] In the steady state, however, there is no difference between this model and one with forward-looking workers. The structure here helps simplify the transition dynamics but could be given up at the cost of adding an additional state variable.

$$\begin{array}{c} \max_{\{c,B\}} \int_0^\infty e^{-(\rho-g_L)t} \log c_t \, dt \\ \text{s.t. } \dot{B}_t = r_t B_t - C_t. \end{array} \tag{1}$$

B denotes total assets in the economy, which will include both physical capital and shares in firms. Changes in total assets $\dot{B}_t$ denote net savings and r is the net return on all assets. Per capita consumption of capitalists is denoted by $c_t = \frac{C_t}{L_t}$ and the discount factor satisfies $\rho - g_L > 0$. Solving the consumption-saving problem leads to the standard Euler equation (2) where capitalists' per capita consumption grows at rate

$$\frac{\dot{c}}{c} \quad = \quad r_t - \rho. \tag{2}$$

Note that all variables that are not exogenous parameters should have a t subscript that I drop for readability.

Final Goods Production: A competitive final good sector combines differentiated intermediate goods $i \in \Omega_M$ to produce final output Yaccording to

$$Y \quad = \quad L^{-\delta_Y} \left(\int_{\Omega_M} (q_i)^{\frac{\sigma-1}{\sigma}} \, di \right)^{\frac{\sigma}{\sigma-1}} \quad , \tag{3}$$

where the elasticity of substitution between differentiated intermediate goods equals σ. $L_t^{-\delta_Y}$ is an additional productivity shifter. Note that the market structure in the production sector is one of monopolistic competition. Population growth thus leads to additional productivity growth due to increasing returns in the production sector.[9] I take this effect out by assuming $\delta_Y = \frac{1}{\sigma-1}$, which exactly offsets increasing returns in the production sector.[10]

The final good serves as the numeraire. It can be used for consumption or turned into physical capital one for one. Denoting aggregate consumption, i.e. the sum of capitalist and worker consumption, as $\tilde{C}$ and investment in physical capital as I, the usual resource constraint follows

$$\dot{K} \quad = \quad Y - \tilde{C} - \delta_k K, \tag{4}$$

where the physical capital stock K depreciates at rate δ_k.

Intermediate Goods Production: I often refer to the set of intermediate goods producers as firms in the production sector. In this production sector symmetric firms of infinitesimal size

[9] I cannot abstract away from population growth since it is needed to sustain long-run technological change in this semi-endogenous growth model,

[10] There are two reasons to do this. First, a strong variety growth effect in the production sector would imply that much of long run growth is driven by an increasing measure of firms in the production sector, and not by novel technology. Second, without this adjustment large countries would be systematically more productive than small ones, which is hardly the case in the data, see Klenow and Rodriguez-Clare (1997) and Caselli (2005) on cross-country income differences. This ad-hoc adjustment does not change any of the qualitative insights of the theory. A micro-foundation could be provided by adding a fixed factor, say land, into a constant-returns-to-scale aggregate production function.

compete monopolistically. The problem of an intermediate goods firm can be split into a static profit maximization problem and a dynamic adoption problem.

Static firm problem: Firm $i \in \Omega_M$ produces according to a Cobb-Douglas production function that combines differentiated capital goods $x_j \in \Omega_{A_i}$ with production labor l_i,

$$q_i = \left(\int_{j\in\Omega_{A_i}} \left(\frac{x_{ij}}{\alpha}\right)^{\alpha} dj\right)\left(\frac{l_i}{1-\alpha}\right)^{1-\alpha} . \tag{5}$$

The set Ω_{A_i} contains all capital goods that the firm i is able to use. Note that this is a subset of all capital goods that are in principal available where the total set is denoted by Ω_{A_F} and $\Omega_{A_i} \subseteq \Omega_{A_F}$ or $A_i \leq A_F$ which is the same inequality expressed in terms of the measure of each set.[11] The measure of capital goods that the firm has access to will be pinned down by the dynamic adoption choice but can be taken as given when solving the static problem. I assume that capital goods are symmetric so that $\int x_{ji}dj = A_i\overline{x}$ where $\overline{x} = x_j \,\forall j \in \Omega_{A_i}$. Equal spending across capital goods in an implication of profit maximization and capital good symmetry, i.e. there are no quality differences across capital goods. Note that this symmetry also implies $p_{xj} = p_x \,\forall j$. Production firms rent these capital goods anew each period.

The amount of ideas the production firm has access to depends on what I call "know-how". Define the variable A_{iK} as a measure of "know-how" (K for "know-how"). This is the set of capital goods the firm knows how to use, a key state variable in the dynamic adoption problem. While $A_{iK} = A_i$ are the same number in equilibrium because all capital goods that the firm knows how to use are going to be used, it is useful to distinguish them. Strictly speaking, A_{iK} represents organizational capital, while A_i is the equilibrium measure of capital goods in use.

The intermediate goods firm in the production sector thus solves

$$\begin{aligned}
&\max_{p_i,q_i,\{x_{ji}\},l_i} \pi_i = p_i q_i - c\left(q_i\right)\\
&\quad s.t.\\
&\qquad q_i = Y p_i^{-\sigma}\\
&\qquad q_i = \left(\int_{j\in\Omega_{A_{iK}}} \left(\frac{x_{ij}}{\alpha}\right)^{\alpha} dj\right)\left(\frac{l_i}{1-\alpha}\right)^{1-\alpha}
\end{aligned}$$

taking factor prices p_x and w and aggregate variables as given, i.e. the solution concept is one of monopolistic competition. This static problem is well-known, and leads to a constant markup over marginal cost, which in turn are a weighted geometric average where the weights are given by the

[11] I will establish a link between available capital goods and innovation following Romer (1990) later on, where each capital good embodies a unique idea. The sets Ω_M, Ω_{A_i}, and Ω_{A_F} will all be evolving endogenously over time.

Cobb Douglas production function,

$$mc \quad = \quad (p_x)^{\alpha} \left(\frac{w}{A_{iK}}\right)^{1-\alpha}. \tag{6}$$

Note the variety effect encoded in A_{ik} that is baked into the production function. Intuitively, given a fixed level of capital expenditure, a firm prefers to spend this money on many different capital varieties because there are diminishing returns in each individual capital good variety. An increase in A_{iK} thus makes the firm more productive and pushes down marginal cost.[12] The price of a differentiated intermediate good reads

$$p = \frac{\sigma}{\sigma - 1} mc, \tag{7}$$

Factor demand for production labor and capital goods are proportional to revenue $\tilde{r} = Yp^{1-\sigma}$

$$\begin{aligned} wl &= \tilde{r}\tfrac{\sigma-1}{\sigma}(1-\alpha) \\ p_x\overline{x} &= \tilde{r}\tfrac{\sigma-1}{\sigma}\alpha, \end{aligned} \tag{8}$$

and operating profits, defined as revenue minus variable cost, $\pi^o = \tilde{r} - wl - p_x\overline{x}$, are proportional to revenue as well

$$\pi^o \quad = \quad \tfrac{\tilde{r}}{\sigma}. \tag{9}$$

Dynamic adoption problem: The adoption of new capital goods is a costly process that is carried out by forward-looking firms. This part of the model is novel, and I discuss crucial assumption and implications below, while laying out the environment here. Adopting new capital varieties where new is to be understood from the point of view of the intermediate goods producer requires skilled labor. The process of technology adoption takes the simple form of increasing the size of the set $\Omega_{A_{iK}}$ by adding capital goods from the set $\{x_j : j \in \Omega_{A_F} \wedge j \notin \Omega_{A_{iK}}\}$. I assume that the adoption process takes the following functional form

$$\dot{A_{iK}} \quad = \quad \zeta A_F^{1-\theta} A_{iK}^{\theta} h_i^{\beta} - A_{iK}\delta_I, \tag{10}$$

where $\theta \in (0, 1)$, $\zeta > 0$, and $\beta \in (0, 1)$. The law of motion is similar to Lucas (2009a) or Sampson (2019) where the term $1 - \theta$ captures an "advantage of backwardness" (Gerschenkron, 1962). This allows for temporary growth spurts when the distance between current technology and frontier is large. Importantly, adoption-driven growth only happens as long as the firm hires skilled labor h_i. Lastly, capital goods disappear at the Poisson rate δ_I, which represents a random death shock to the idea that will be embodied in the capital good as I discuss below. Note that (10) implies that the

[12] This variety effect was originally introduced in Dixit and Stiglitz (1977) on the demand side. See Ethier (1982) for a supply side interpretation.

firm has control over its own knowledge stock A_{iK}, and takes the evolution of the frontier level of technology A_F and other firms' productivity $A_{j:j\neq i}$ as given. Moreover, the constant ζ needs to be sufficiently small to rule out a corner solution at $A_{iK} = A_F$.[13]

The dynamic problem of the firm can be stated using the HJB approach, where r denotes the interest rate, δ_{ex} a Poisson death shock to production firms, and V is the value function of the firm,

$$(r_t + \delta_{ex}) V(A_{iK}, t) \quad - \dot{V} = \quad \max_{h_i} \pi_t^o (A_{iK}) + \partial_{A_{iK}} V(A_{iK}, t) \left[\dot{A_{iK}} \right] \quad - w_H h_i. \tag{11}$$

The current level of know-how A_{iK} is the key state variable of the firm. It impacts its current profit flow but also affects the law of motion of adoption. Other aggregate state variables, such as total demand or the measure of firms, are captured in t. This model of technology adoption has a fixed cost flavor as the adoption choice does not interact with the static profit maximization decision which renders the model tractable. Given constant returns to scale on the firm level, adoption related overhead costs necessitate a model of imperfect competition in the production sector since a competitive production sector would not be able to generate the profits needed to sustain technology adoption.[14]

Free entry: I close the production sector by assuming free entry after paying a fixed entry cost in terms of production labor.[15] I assume that entrants reach the know-how of incumbents as they enter, which captures a strong knowledge spillover within each country in anticipation of the open economy setting later on. This implies that I can drop the i subscript since the spillover ensures all firms are identical and thus make identical choices. The free entry condition reads

$$f_e w \quad \geq \quad V(A_K, t). \tag{12}$$

The inequality is binding when there is positive entry, which gives rise to an endogenous measure of intermediate goods firms. This measure is denoted by M and changing over time according to

$$\dot{M} \quad = \quad \tfrac{L^E}{f_e} - M\delta_{ex} \tag{13}$$

where L^E and L^P are production labor devoted to entry or production.

The assumption of strong local knowledge spillover merits some discussion. is adoption setting merits additional discussion. The benefit of the symmetric firm model is that it substantially simplifies the innovator problem. Since innovator take into account how fast their ideas are adopted, a model of heterogenous firms means that the entire distribution is part of the state space. Abstracting away

[13]Another strategy is to use the original Nelson-Phelps specification $\dot{A_K} = (A_F - A_K)\,\psi(h)$, 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 the corner solution. The downside is that the speed of convergence to the steady state, conditional on β, is fixed. My specification has an additional degree of freedom in θ which allows me to match the speed of convergence across countries.

[14]This argument has been made in Schumpeter (1942) and Romer (1990) with regard to innovation. The argument also applies to adoption once it is modeled as a costly activity in a model with constant returns to scale.

[15]I discuss a version of the model where entry costs are paid in terms of a composite good that uses both production workers and skilled workers in section 3.1, together with several other extensions.

from this layer of heterogeneity allows me to focus on cross-country and within-country inequality, and how they relate to innovation and technology adoption in a global world.

An alternative is to allow firms to enter with an imperfect knowledge spillover.[16] In the appendix A.2 I consider such a setting where entrants enter with a below-average productivity but they make endogenous adoption decisions that allow them to converge to the state of the art technology in the long run. [17]

Innovation: Innovators expend skilled labor to add novel technology to the stock of ideas, following Romer (1990). Denote by A_F the technological frontier, which is simply the total number of ideas ever invented in this model of horizontal differentiation. I assume that innovators can produce a flow of $\frac{1}{f_R} A_F^{\phi}$ new ideas with one unit of skilled labor, where f_R represents a fixed entry cost. A knowledge spillover is captured in the parameter ϕ but I allow for this spillover to be weak, i.e. $\phi < 1$, following Jones (1995)'s semi-endogenous growth logic. The aggregate flow of ideas equals

$$\dot{A_F} = \frac{1}{f_R} A_F^{\phi} H_F - \delta_I A_F, \tag{14}$$

where H_F denotes the amount of skilled labor devoted to the development of new ideas. Moreover, the fixed cost includes a congestion force as in Jones (1995)

$$f_R = \frac{H_F^{1-\lambda}}{\gamma} \tag{15}$$

where γ represent an exogenous research productivity and $\lambda \in (0, 1]$ parameterizes the congestion force.[18] Innovators are infinitesimal, so they take aggregate variables and factor prices as given.

Free Entry: Entry occurs up until the net present value of an innovation equals the entry cost

$$V_I A_F^{\phi} \quad \leq \quad f_R w_H. \tag{16}$$

where (16) is binding whenever there is entry into innovation. This gives rise to an endogenous measure of ideas in equilibrium, and since $\phi < 1$, positive population growth is needed to sustain technological change.

Present Discounted Value of an Idea: In contrast to Romer (1990), where the adoption of new ideas is immediate, the benefit from innovation only comes with a delay. This delay is endogenous, and

[16] See Luttmer (2007), Lucas (2009b), Sampson (2016) and Buera and Oberfield (2020) for models with knowledge spillover.

[17] I show that this leads to a model with an endogenous firm size distribution. Importantly, in the steady state, after integrating out firm heterogeneity to compute aggregate outcomes, the qualitative predictions of the model remain unchanged. Once the equilibrium has reached a stationary steady state, a shock to the cost of technology adoption, say a rising skill premium, will shift the average of the stationary distribution to the same extent as firms in the homogeneous firm model. In follow up work I do focus on how rising skill prices and technological frontier growth interact with the firm size distribution, with a more flexible adoption technology and market structure on the firm side.

[18] A justification for this congestion force is that there may be useless duplication, i.e. two researchers coming up with the same idea.

depends on adoption in the production sector, which is the key new feature of the model. Note that the present discounted value of an innovation can be written as the usual discounted sum of future profits

$$V_I \quad = \quad \int_{t+\tau_t}^{\infty} \exp\left(-\int_t^u (r_v + \delta_I)\, dv\right) \pi_u^I du \tag{17}$$

where π^I represents the flow profits (royalty) and u and v are arguments of integration. Denote with $\tau \in \mathbb{R}^+$ the endogenous waiting time it takes for an idea to become profitable, i.e. τ is the timer interval between entry and first profit. Since the cost of innovation are incurred at time t, the discount factor runs from t onward.

I first turn to the flow profits. I first turn to the flow profits which lead to a static pricing problem for the owner of the patent-protected idea. I follow Romer (1990) and assume that idea-embodying capital goods are produced with physical capital alone according to a linear production function. For simplicity, I assume that capital can be turned into capital goods one for one. Note that demand for each capital good has the familiar CES structure which follows from the intermediate goods firm optimal demand for capital goods, i.e. the production sector represents the demand side here. In particular, the static problem at each instant reads

$$\begin{aligned}
\max_{p_{xj}, X_j} \qquad & \pi_j = p_{xj} X_j - c(X_j) \\
s.t. \qquad & \\
& X_j = P_x X \left(\frac{p_{xj}}{P_x}\right)^{-\frac{1}{1-\alpha}} \\
P_x \quad & = \left(\int_{j \in \Omega_{A_{iK}}} (p_{xj})^{-\frac{\alpha}{1-\alpha}}\right)^{-\frac{1-\alpha}{\alpha}} \\
& c(X_j) = X_j (r + \delta_k)
\end{aligned}$$

where $\int X_j dj = X$, and $X_j = x_j M$ is aggregate demand for capital goods, and aggregate demand for a particular type j which simply is firm demand times number of firms. The cost function $c(.)$ is linear, and P_x is an aggregate price index that will look very simple in equilibrium since there is no heterogeneity across capital goods. The last line then uses the fact that the rental rate of physical capital equals the interest rate plus depreciation. Again, the reader familiar with models of monopolistic competition will anticipate that the equilibrium price equals

$$p_{x_j} = \frac{1}{\alpha}(r + \delta_k) \quad \forall j. \tag{18}$$

which is a constant markup over marginal cost.[19] After solving out for the endogenous price index and aggregating over all intermediate goods firms, the flow profits are equal to a constant share of total

[19] In this model the capital share and the markup are tied together as in Romer (1990) or Jones (1995). One could easily change this by modeling the production function of intermediate goods firms using a double-nest with two different elasticities, i.e. $y = \left(\frac{(\int x^\rho dj)^{\frac{1}{\rho}}}{\alpha}\right)^{\alpha} \left(\frac{l}{1-\alpha}\right)^{1-\alpha}$ so that the markup is related to ρ while the capital share is still a function of α.

revenue divided by the total measure of active ideas, which in turn is proportional to the wage bill in the economy due to Cobb-Douglas production[20]

$$\begin{aligned} \pi_I &= \frac{P_x X(1-\alpha)}{A} \\ &= \frac{\alpha L^P w}{A} \end{aligned} \tag{19}$$

Now I turn to the endogenous waiting time. I partition the set of capital goods A_F into the set $\Omega_A \in [0, A]$ and $\Omega_F \in (A, A_F]$. A capital good in set Ω_A is in use, while a capital good in set Ω_F is waiting to be adopted. For simplicity, I assume that among all available but unused ideas, the idea that has been developed first is going to be adopted first. Moreover, all ideas, adopted and waiting to be adopted, are subject to the Poisson death shock at rate δ_I that already showed up in the law of motion of adoption.[21] Simply put, innovators wait in line till they are up. And they are up when all innovators, which invented before them, are adopted or disappeared due to the Poisson shock.[22] This means that the time it takes for an idea to be adopted is endogenous and in particular depends on adoption effort in the production sector. This waiting time can be defined mathematically as follows. First, define the measure of ideas that stand between the adoption of an idea invented at time t as $W(t) := A_F - A$. Define the time of adoption $t + \tau_t$ for inventor cohort t. While there are new ideas invented, they will only be adopted after cohort t and are thus irrelevant for cohort t's waiting time. Note that the measure W is shrinking over time for two reasons. Ideas die at rate δ_I, so a flow $W\delta_I dt$ is disappearing at every instant.[23] Second, a flow $A_t(\delta_I + g_A)\,dt$ is adopted every instant, which could be negative or positive.[24] To achieve net variety growth g_A the intermediate goods firm needs to adopt $A_t(\delta_I + g_A)\,dt$ varieties to make up for the loss of ideas due to the random death shock. This adoption leads to a reduction in W as well. In figure ??, I illustrate the flows. One can think of $A_t(\delta_I + g_A)\,dt$ as outflow from the left (in red), and $W\delta_I dt$ as outflow that hits the whole measure W evenly (blue). Based on this argument τ is implicitly defined by $W(t, t+\tau) = 0$, together with an initial condition $W(t,t) = A_F - A$, a trajectory of A_t that the innovators takes as given, and the differential equation

$$\dot{W} = -\delta_I W - A(\delta_I + g_A)\ . \tag{20}$$

[20] Formally, $\pi_I = \int_i r_i^x \underbrace{\left[\frac{p_x^{-\frac{1}{1-\alpha}}}{P_x^{1-\frac{1}{1-\alpha}}}\right]}_{x_{ij}} [(p_x) - (r + \delta_k)]\,di = P_x X \int_i \left[\frac{p_x^{-\frac{1}{1-\alpha}}}{P_x^{1-\frac{1}{1-\alpha}}}\right] \left[(p_x)\left(1 - \frac{\epsilon-1}{\epsilon}\right)\right] = \frac{R^X}{A}(1-\alpha) = \frac{\alpha}{1-\alpha}\frac{L^P w}{A}(1-\alpha) = \frac{\alpha L^P w}{A}$.

[21] This assumption is useful to generate churn among innovators even when there is no population growth and TFP is constant but is not consequential for any qualitative insight.

[22] Whether the adoption is deterministic or stochastic is not central for any of the results that follow and I sketch out a stochastic version in the appendix. Markets are complete in the model so the stochasticity of adoption does not matter and washes out in the aggregate.

[23] This death shock can also hit cohort t and is taken into account when computing the net present value of an invention.

[24] A production firm will never drop ideas on purpose so a negative growth rate is bounded by $-\delta_I$ which is the case when no adoption effort is exerted.

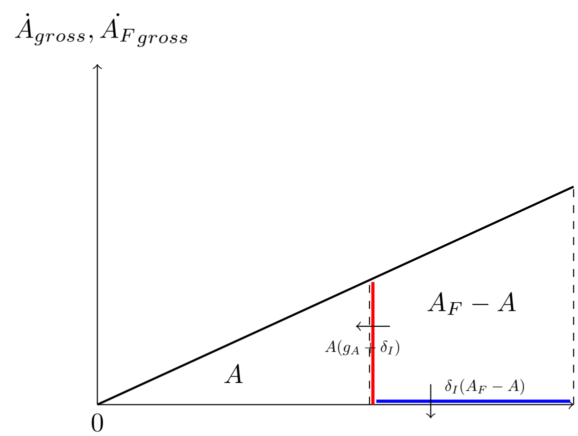


Figure 2. Outflow of X

This formulation leads to a closed form solution for τ as I show below.

2.2 Equilibrium Concept

I define an equilibrium on the balanced growth path of this semi-endogenous growth model as follows.

Definition 1. A balanced growth path equilibrium with constant population growth $g_L = g_H$ and $\phi < 1$ consist of a sequence of prices $\{w_t, w_{Ht}, r_t, p_{xt}, p_{it}, V_t, V_{It}\}$ and allocations $\{L_t^P, L_t^E, H_{Dt}, H_{Ft}, X_t, K_t, M_t, A_t, A_{Ft}, C_t\}$ for $t \in \mathbb{R}$ that grow at a constant rate over time (possibly zero), and a constant adoption gap $\Gamma = \log A_F - \log A$, where

- Final goods producer maximizes profit
- Intermediate goods firms maximize the net present value of their operation subject to (5) and (11) where they take factor prices and aggregate variables as given and free entry holds
- Innovators maximize the net present value of their operation, and free entry holds
- Dynastic capitalists solve the consumption-saving problem given budget constraint and transversality condition
- All factor, goods, and asset markets clear and resource constraints are respected
- together with a set of initial conditions $\{M_0, A_0, A_{F0}, K_0\}$ that are strictly positive.

This completes the equilibrium description. To solve for transition dynamics later on, I define normalized variables by the total amount of production labor to obtain a stationary system of equations. The normalizations reflect that per capita growth in this semi-endogenous growth model is sustained by population growth, see Jones (1995). Let $m := \frac{M}{L}$, $l^P := \frac{L^P}{L}$,$l^E = \frac{L^E}{L}$, $a_F = \frac{A_F^{1-\phi}}{L^\lambda}$, $a = \frac{A^{1-\phi}}{L^\lambda}$, $h_D = \frac{Mh}{L}$, and $h_F = \frac{H_F}{L}$. Moreover, define the normalized technology level $z := \frac{A}{A_F}$ which will be constant on the balanced growth path with a constant adoption gap $\Gamma = -\log z$.

2.3 Solving the Model

Dynamic adoption problem: The intermediate goods firm hires skilled labor in order to adopt new varieties of capital. To solve this firms problem (11), I first need to normalize the HJB equation to render it stationary. Since entry cost grow with the wage rate, the appropriate normalization is w. Moreover, I rewrite the law of motion of adoption in terms of z, the relative technology level. The normalized problem reads

$$v\left(r+\delta_{ex}-g_w\right) = \max_h \frac{\pi_t\left(z\right)}{w} - sh + \left(\partial_z v\right)\dot{z} + \dot{v} \tag{21}$$

$$\text{s.t. } \dot{z} = \zeta z^\theta h^\beta - \left(g_F + \delta_I\right) z, \tag{22}$$

where $\frac{\dot{A_F}}{A_F} = g_F$, and $\dot{z} = \dot{v} = 0$ in the steady state. A solution to (21) needs to satisfy the first order condition

$$\left(\partial_z v\right)\beta\zeta z^\theta h^{\beta-1} \quad = \quad s. \tag{23}$$

Equation (23) captures the trade-off between the cost of adoption and the benefit of a higher productivity level. Perhaps surprisingly, the key price that shows up in this first order condition is the relative price of skill s. Intuitively, profits are proportional to w while adoption cost depend on w_H. The skill premium is thus the relevant relative price that determines the firm's adoption choice. The higher the skill premium, the more costly technology adoption is.

In the appendix I derive the differential equation that characterizes optimal adoption[25], leading to the following law of motion for skilled labor growth on the firm level

$$\frac{\dot{h}}{h} \quad = \quad \frac{1}{1-\beta}\left\{\underbrace{\rho+\delta_{ex}}_{\text{effective discounting}} + \underbrace{\left(1-\theta\right)\left(g_F+\delta_I\right)}_{\text{effective depreciation}} - \underbrace{\left\{\frac{\beta z^\theta \zeta h^{\beta-1}}{s}\left[\frac{\pi_t}{w}\frac{\left(1-\alpha\right)\left(\sigma-1\right)}{z}\right] + \frac{\dot{s}}{s}\right\}}_{\text{marginal benefit of extra unit of skilled labor}}\right\}. \tag{24}$$

Equation (24) is similar in spirit to the well-known q-theory of investment, and I show the mathematical equivalence in the appendix A. Just like in the firm investment literature, firms make a forward looking decision that depend on the current stock z (capital in the investment literature) and the level of investment h that runs into diminishing returns since $\beta < 1$. The speed of adjustment is

[25] For simplicity and expositional purposes I used the steady state interest rate $r = g_w + \rho$. The reader can substitute out ρ if preferred.

going to depend crucially on β and θ, a point I will return to when calibrating the model. Imposing $\dot{h} = \dot{z} = \dot{s} = 0$ in the steady state leads to a simple solution for the demand for skilled labor of intermediate goods firms.

Proposition 1. *Suppose* $\frac{\rho+\delta_{ex}}{\delta_I+g_F} + (1-\theta) > \beta(\sigma-1)(1-\alpha)$ *holds, then a unique saddle-path stable steady state equilibrium obtains for a fixed relative price of skill* s *and a fixed frontier growth rate* g_F*.*

The inequality in proposition 1 guarantees existence and uniqueness of the solution. It ensures that the future benefit of improving ones productivity are sufficiently small relative to effective discounting. If this is the case, the firm's demand for skilled labor for adoption purposes equals

$$h = \frac{1}{s}\frac{\beta(1-\alpha)(\sigma-1)(g_F+\delta_I)}{\rho+\delta_{ex}+(1-\theta)(g_F+\delta_I)}\left[\frac{\pi}{w}\right]. \tag{25}$$

The demand for skilled labor is proportional to normalized profits, and falling in the skill premium. Moreover, it is positively related to the sensitivity of profits with respect to productivity, $(\sigma-1)(1-\alpha)$, as a large demand elasticity will make firms benefit more from a technological improvements, ceteris paribus.[26]

The qualitative transition dynamics in partial equilibrium (fix r and s) can be studied using a phase diagram. The law of motion of z implies a positive link between skilled labor and relative technology level z. After inspecting equation (24) one can see that the marginal product of an additional unit of skilled labor falls as z increases as long as $\theta < 1$, a mechanism similar to a diminishing marginal product of capital in the neoclassical model. This implies a negative relationship between h and z in the steady state, leading to a unique pair $\{z, h\}$. I plot the qualitative dynamics after a 10% increase in the relative price of skill in figure 3. The dashed blue line shows the new locus in the steady state, and the arrows indicate the transition path. There is an strong initial jump down to a lower level of skilled labor, which is a direct response to the increase in the skill premium. The equilibrium converges to a new steady state by raising skilled labor investment slightly. These qualitative dynamics are identical to q-investment dynamics after an increase in the interest rate.

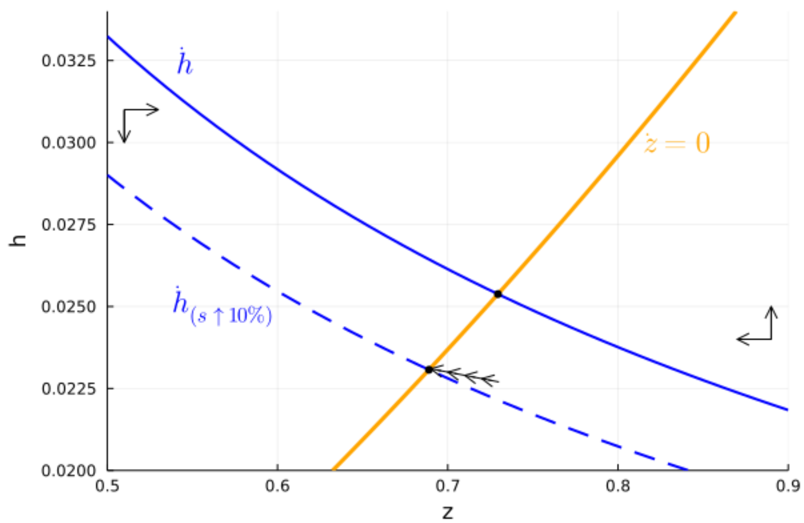
The value of a firm in the production sector equals the sum of its discounted profits

$$V_t = \int_t^\infty \exp\left(-\int_t^u (r_v + \delta_{ex})\, dv\right) w_u \left[\frac{\pi_u}{w_u} - s_u h_u\right] du.$$

Following the steps in Peters and Walsh (2019) and using the discipline imposed by the free entry

[26] The ceteris paribus assumption is crucial here, since under monopolistic competition among homogeneous firms, all firms make the same investment choice and so their individual improvements are undone by a reduction in the aggregate price index. Since the aggregate price index is normalized to unity, this adjustment occurs through an increase in the real wage.

Figure 3. Phase Diagram



condition, one can show that the normalized value function, $v = \frac{V_t}{w_t}$, is equal to

$$v \quad = \quad \frac{\frac{\pi}{w} - sh}{r_t + \delta_{ex} - g_w} \tag{26}$$

as long as the free entry condition binds. The value of the firm is thus directly tied to net profits $\frac{\pi}{w} - sh$ and appropriately discounted by taking into account the cost of capital, the death probability, and wage growth.[27] Define $\kappa_1 := \frac{\beta(1-\alpha)(\sigma-1)(g_F+\delta_I)}{\rho+\delta_{ex}+(1-\theta)(g_F+\delta_I)}$, and $\kappa_2 := \frac{1}{1-\kappa_1}$ to simplify notation and impose $v = f_e$ in the steady state. Together with (26) and (25) the normalized operating profits on the balanced growth path are pinned down,

$$f_e\left(\rho + \delta_{ex}\right)\kappa_2 = \frac{\pi}{w}, \tag{27}$$

where I used $r = \rho + g_w$. Equation (27) is directly related to the flow cost of entry, $f_e(\rho + \delta_{ex})$, but it features an additional term $\kappa_2 > 1$,[28] a consequence of the additional overhead costs due to technology adoption.

Using the fact that in a homogenous firm model operating profits are equal to $\pi = \frac{Y}{M}\frac{1}{\sigma}$, together with $Y\frac{\sigma-1}{\sigma}(1-\alpha) = L^P w$ from Cobb-Douglas production, I can pin down the ratio of normalized production labor and equilibrium measure of intermediate goods firms m

[27] The free entry condition ties the value of entry to the wage rate, and hence higher future wages must mean higher future firm values as long as the free entry condition is binding.

[28] As long as proposition 1 holds, $\kappa_2 > 1$ will hold as well.

$$f_e\left(\rho+\delta_{ex}\right)\kappa_2 \quad = \quad \frac{l^P}{m}\frac{1}{(1-\alpha)(\sigma-1)}. \tag{28}$$

Together with the normalized firm entry resource constraint

$$\dot{m} \quad = \quad \frac{l^E}{f_e}-\left(\delta_{ex}+g_L\right)m, \tag{29}$$

the steady state normalized measure of firms reads

$$m \quad = \quad \frac{1}{f_e\left[(\rho+\delta_{ex})(1-\alpha)(\sigma-1)\kappa_2+g_L+\delta_{ex}\right]}. \tag{30}$$

Note that out of steady state, the endogenous firm measure is not constant.[29]

Steady State Adoption Gap: This model features a constant adoption gap in the steady state. It is easy to see how the adoption gap is increasing in the skill premium by combining the adoption technology (21) with the firm's demand for skill (25). Taking logs leads to

$$\log z \quad = \quad -\frac{\beta}{1-\theta}\log s+\frac{1}{1-\theta}\log\left(\frac{\zeta}{(g_F+\delta_I)}\left(\frac{\pi}{w}\kappa_1\right)^{\beta}\right). \tag{31}$$

Expression (31) highlights the response of the relative technology level z to an increase in the skill premium. A 1% increase in the skill premium reduces the relative technology level z by $\frac{\beta}{1-\theta}$%. Intuitively, both diminishing returns in adoption (β) as well as advantage of backwardness ($1-\theta$) jointly determine the strength of this response. Skilled labor in adoption is important when β is large so that adoption effort does not run into diminishing returns. Similarly, the effect is strong when θ is large, which is the extent to which current knowledge helps to adopt new knowledge. And when skilled labor is important in technology adoption, the effects of a rising skill premium are severe. The skill premium is not only an accounting device to keep track of inequality, but takes on an additional role whereby it directly impacts productivity. This point is related to models of directed technological change as in Acemoglu (2002), but the mechanism is quite different. A rising skill premium simply makes adoption more expansive, and thus hurts technology adoption.

Innovation: Innovators need to take into account that their idea is adopted with a lag τ and only then becomes profitable. The present discounted value of an idea reads

$$V_{It} \quad = \quad \int_{t+\tau}^{\infty}\exp\left(-\int_t^u\left(r_x+\delta_I\right)dx\right)\pi_u^I du \tag{32}$$

where the flow profits equal $\pi^I=\frac{\alpha L^P w}{A}$. Define $\tau' := \frac{\partial \tau_t}{\partial t}$ as the instantaneous change in the waiting

[29] This is an implication of (26) which states that firms need to earn sufficiently high operating profits to make up for technology adoption cost. When adoption is high, entry needs to stall so that incumbents can still break even despite large adoption costs.

time. When the free entry condition is binding, the value function can be written in simplified form

$$V_I = \exp\left(-\int_t^{t+\tau}[r_x+\delta_I]\,dx\right)(1+\tau')\,\frac{\pi^I_{t+\tau}}{r+\delta_I-g_{w_H}-(1-\lambda)\,g_{H_F}+\phi g_F} \tag{33}$$

which holds on and off the balanced growth path.[30] The expression combines the flow profits in period $t+\tau$ with an appropriate discount factor that takes into account a standard term $\frac{1}{r+\delta_I-g_{w_H}-(1-\lambda)g_{H_F}+\phi g_F}$, an extra discount factor $\exp\left(-\int_t^{t+\tau}[r_x+\delta_I]\,dx\right)$ that runs from t to $t+\tau$ since ideas become profitable only at $t+\tau$ while costs are incurred at t, and an additional term $1+\tau' = \frac{\partial[t+\tau]}{\partial t}$. This term incorporates changes in the waiting time off the balanced growth path. Details are in the appendix A.3.1.

The waiting time τ, which is essential to compute the value of an innovation, turns out to be a simple expression in the steady state that is proportional to the adoption gap,

Proposition 2. *In a steady state the waiting time depends on the ratio of the adoption gap $-\log z$ and the gross adoption rate $(g_A+\delta_I)$*

$$\tau \quad = \quad -\frac{\log z}{g_A+\delta_I}. \tag{34}$$

Proof in A.3.4.

Intuitively, equation (34) takes physical units of productivity ($\log A_F - \log A = -\log z$) and projects them into time units τ by dividing through the gross adoption rate $g_A+\delta_I$ measured at a point in time. As the adoption gap disappears ($z\to 1$), the waiting time shrinks to zero.

In the steady state, the present value of an innovation thus simplifies to[31]

$$V_I \quad = \quad \frac{1}{\tilde\rho+g_F+\delta_I}\left(\frac{\alpha L^P w}{A_F}\right) z^{\frac{\tilde\rho}{g_A+\delta_I}} \tag{35}$$

where $\widetilde{\rho} := \rho - g_L > 0$ is the effective discount factor of the dynastic household and I substituted out τ using (34). Clearly the net present discounted value of innovation depends on adoption effort in the production sector through its effect on z. If there was no adoption, z would be zero and there would be no innovation either.

Using the normalized notation, and combining (35) with the free entry condition, $f_R w_H A_F^{-\phi} = V_I$, leads to the research arbitrage condition that binds whenever there is positive entry,

$$\frac{1}{\gamma} \quad = \quad \frac{1}{s}\frac{\alpha l^P_t}{\tilde\rho+g_F+\delta_I}\left(\frac{h_F^{\lambda-1}}{a_{Ft}}\right) z^{\frac{\tilde\rho}{g_A+\delta_I}}. \tag{36}$$

[30] Unless otherwise indicated growth rates are in time t, i.e. the denominator I have $\phi g_F = \phi g_{Ft}$.

[31] Profits accrue only from date $t+\tau$ on but on the balanced growth path I can write the expression in terms of date t variables since $\exp\left(-\int_t^{t+\tau}[r_x+\delta_I]\,dx\right)\frac{L^P_{t+\tau}}{L^P_t}L^P_t = \left(-\int_t^{t+\tau}[r_x+\delta_I-g_{L^P}]\,dx\right)L^P_t$. Moreover, I use the fact that $A = A_F z$ to substitute out A.

Two important economic mechanisms are captured in (36). First, the skill premium is again the central relative price to determine entry into innovation. While the innovator pays a fixed cost in high skilled wages w_H, their profits later on are proportional to the wage in the production sector w so that the crucial price signal is the ratio of high and production labor wages. Note that a_F, the relative measure of ideas, needs to decline as the skill premium increase. As entry gets more expensive, a downward adjustment in the number of ideas ensure that innovators still break even.

Market Clearing: The final step in solving the models requires finding the relative price of skill that clears the market for skilled labor. Normalized skilled labor demand $h_D = \frac{hM}{L}$ in the production sector is readily derived

$$\begin{aligned} h_D &= m h_i \\ &= \tfrac{1}{s}\Lambda^D \end{aligned} \tag{37}$$

where Λ^D collects elements that are constants in the steady state.[32] Using a normalized version of the law of motion of ideas (14) where $h_F := \frac{H_F}{L}$, I get

$$\frac{\gamma_R h_F^{\lambda}}{(g_F + \delta_I)} = a_F. \tag{38}$$

Combining (36) with (38) leads to the research sector's normalized demand for skilled labor

$$\begin{aligned} h_F &= \tfrac{1}{s}\left(\frac{g_F + \delta_I}{\tilde{\rho} + g_{A_F} + \delta_I}\right) \alpha l^P (z)^{\frac{\tilde{\rho}}{\delta_I + g_F}} \\ &= \tfrac{1}{s}(z)^{\frac{\tilde{\rho}}{\delta_I + g_F}} \Lambda^F. \end{aligned} \tag{39}$$

Adding up (37) and (39) and imposing market clearing, I obtain the following equation that implicitly defines the relative price of skill

$$\left\{ \tfrac{1}{s}(z)^{\frac{\tilde{\rho}}{\delta_I + g_F}} \Lambda^F + \tfrac{1}{s}\Lambda^D \right\} = h^{tot} \tag{40}$$

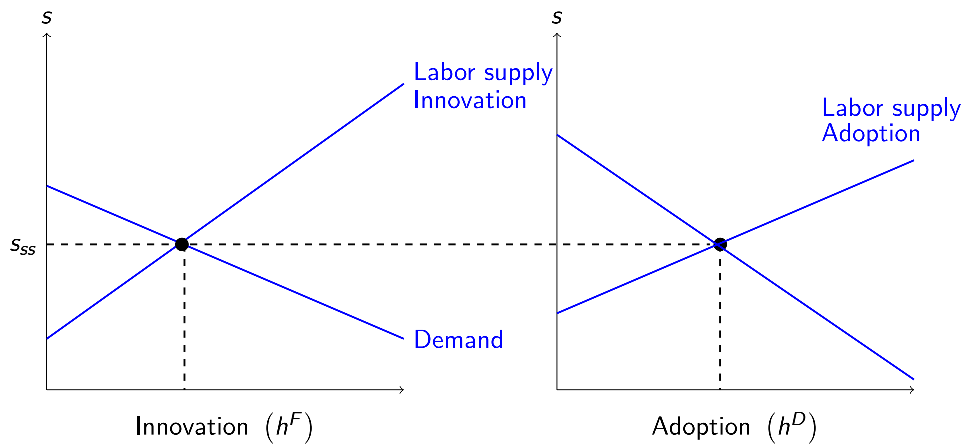
where $\frac{H}{L} = h^{tot}$. Note that z itself is a function of the price of skill so this equation needs to be solved numerically. Throughout the paper I focus on equilibria where h^{tot} is sufficiently scarce so that $s > 1$.

This market clearing condition connects adoption activity and innovation activity as they compete for the same scarce resource, skilled labor. A simple diagram 4 helps to illustrate their interactions. Both adoption activity and innovation activity are downward sloping in the skill premium. While aggregate labor supply is fixed, it is upward sloping for each sector individually and equilibrium is reached when the relative price of skill clears both markets.

Aggregation: This economy behaves similar to a neoclassical economy where long run growth

[32] That is, $\Lambda^D = \kappa_1 \frac{\pi}{w} m$ whose values I have derived in the previous section. Importantly, these are constant in the steady states.

Figure 4. Market Clearing for Skilled Labor



can be summarized as follows.

Proposition 3. *The balanced growth path is characterized by the following long run growth rates: firm growth in the production sector is equal to population growth,* $g_M = g_{L,}$, *technology frontier growth is equal to* $g_F = \frac{\lambda}{1-\phi} g_L$, *the adoption gap is constant so* $g_A = g_F$, *wage growth equals* $g_w = g_A$, *and capital accumulates at a growth rate* $g_K = g_L + g_A$.

Note that both l_i and h_i are constant in the steady state but aggregate demand for low and high skilled labor rises in line with overall population growth through the extensive margin. This means that long-run per capita growth is characterized by a constant z together with an ever-expanding stock of ideas A_F. The production sector aggregates up to a neoclassical production function where the term $A_F z$ represents labor productivity,

$$Y \quad = \quad \left(\frac{K}{\alpha}\right)^{\alpha} \left(\frac{z A_F L^P}{1-\alpha}\right)^{1-\alpha} \tag{41}$$

and total demand for capital goods matches physical capital $MA\overline{x} = K$. A standard link between the rental rate of capital and the capital labor ratio emerges, but markups must be applied twice, due to imperfect competition in both the innovation and production sector

$$RK = \alpha Y \underbrace{\frac{\sigma-1}{\sigma}\alpha}_{\text{markup}},$$

which leads to a constant capital-effective labor ratio on the balanced growth path

$$k_{ss} = \left\{\left(\frac{\sigma-1}{\sigma}\alpha\right)\frac{\alpha}{[(\rho+g_w+\delta_K)]}\right\}^{\frac{1}{1-\alpha}}$$

where $k := \frac{K}{zA_F L^P}$. The real wage for production and high skilled workers reads

$$\begin{aligned} w &= (1-\alpha)\frac{\sigma-1}{\sigma} z A_F k_{ss}^{\alpha} \\ w_H &= sw \end{aligned}. \tag{42}$$

The model nests the model of Jones (1995) as for the right sequence of parameters, the productions sector becomes perfectly competitive ($\frac{\sigma}{\sigma-1} \to 1$) while the adoption frictions vanishes ($z \to 1$), see appendix A.4. Just as in Jones (1995), or any other growth model, real income is low when the level of technology A_F is low. I allow for an additional mechanism that generates low real per capita income: a lack of technology adoption reflected in a low z. This feature allows the model to match cross-country inequality which mostly depends on the distribution of country-specific z-levels as I show in the next section. It also allows for the possibility of a growth slowdown in the face of rising innovative effort and frontier technology growth, a case when z and A_F move in opposite directions.

2.4 Complementarity between Innovation and Adoption

Endogenizing both innovation and adoption leads to novel interactions between the two. First, based on the innovator problem and in particular equation (36), the partial equilibrium elasticity of the total measure of ideas A_F with respect to the adoption gap in the steady state equals[33]

$$\frac{\partial \log A_F}{\partial \log z} = \frac{\lambda}{1-\phi}\frac{\tilde{\rho}}{g_A+\delta_I}. \tag{43}$$

As production firms raise their adoption effort and push up z, the present discounted value of an innovation increases due to a falling waiting time τ. This leads to additional entry into innovation and pushes up the total stock of ideas A_F. The strength of this complementarity depends on the ratio of effective discounting and the gross adoption rate ($\frac{\tilde{\rho}}{g_A+\delta_I}$), interacted with the overall sensitivity of idea output to skilled labor input ($\frac{\lambda}{1-\phi}$).

This complementarity remains important in general equilibrium. While innovation and adoption

[33] The elasticity here is to be understood relative to some alternative balanced growth trend since A_F is growing over time.

are rivalrous due to competition for skilled labor on factor markets, they are complementary in the sense that positive productivity shocks to one activity lead to an expansion in the net output of the other activity as well. To see this, I consider how innovation and adoption respond to three different fundamental shocks in the model, showing that the two activities tend to move in lock-step even in general equilibrium.

First, consider an increase in γ, the research productivity. From equation (38), one can immediately infer that the steady state demand for skilled labor in research is independent of the fixed research cost. Market clearing remains unchanged, nor does the skill premium move. The measure of ideas grows at an elevated rate for some time and since z remains constant, technology adoption must occur at an elevated rate as well. The takeaway is that biased exogenous productivity growth favoring the research sector does not lead to divergence between innovation and adoption. A result that will change in the open economy as I show later.

Next suppose that ζ increases which effectively makes adoption easier. This leads to a larger z, which in turn leads to a reallocation of labor from adoption to innovation and a higher skill premium. It can be shown that both z and A_F increase, highlighting the complementarity of the two activities. The intuition is that a declining adoption friction leads to higher innovator profits, which leads to a reallocation of labor into innovative activity. One way to see this is to compute the ratio of skilled labor devoted to innovation relative to adoption by combining (37) and (39)

$$\frac{H^F}{H^D} \quad = \quad \left(\frac{g_F+\delta_I}{\tilde{\rho}+g_F+\delta_I}\right)\kappa_1\alpha\left(\sigma-1\right)\left(1-\alpha\right)\left(z\right)^{\frac{\tilde{\rho}}{\delta_I+g_A}}. \tag{44}$$

Given that z increases, this implies a reallocation of labor from adoption to innovation. Yet, both adoption and innovation expand in the sense that the stock of ideas increases while the adoption gap declines.

Finally, Suppose that the relative supply of skilled labor shrinks. For example, a changing task content of work could lead to such a scenario, which I explore more carefully in subsection 3.1. A negative shock to the relative supply of skilled labor leads to a rising skill premium, which hurts both innovation and adoption. Note that the effect on innovation is stronger, as seen in equation (44). Since z is falling due to a rising skill premium, innovation is hurt twice, first directly since its main input, skilled labor, has gotten more expensive, and second due to the effect of a rising adoption gap on the net present value of an innovation. Again, innovation and adoption move in the same direction, and innovation responds even stronger than adoption.

All of this points to a general complementarity between the two activities. These different scenarios highlight that it is difficult for innovation activity to run away from the rest of the economy, precisely because the rest of the economy represents the client base for innovators. The next section shows how this complementarity breaks down in the open economy.

3 Open Economy

In this section I study the implications of my model in a simple two-country open economy setting. Countries have different fundamental research productivity γ and different relative skill endowments h^{tot} but are otherwise identical. In particular, preferences and non-research technology are the same. Countries produce the same final goods, and I focus on an integrated equilibrium with frictionless trade in final goods and ideas. There is no migration, and I abstract away from intermediate goods trade in the production sector, but this latter assumption is not relevant for the innovation-adoption tradeoff or inequality.[34] Lastly, I assume that capital goods are produced locally using capital accumulated by the domestic economy, and I impose that trade is balanced at all times. By assuming that capital goods are produced locally, I abstract away from offshoring. Importantly, even if a domestic capital good is produced abroad for foreign use, the domestic inventor still receives a royalty.[35] The model of trade will thus be one where poor countries trade final goods in order to use ideas produced in advanced economies.

I focus on steady state results in this theoretical section. In what follows, the asterisk $*$ denotes foreign variables where the domestic economy is assumed to be the advanced economy, and W denotes world aggregates.

Cross Country Income Differences: Before I solve for an equilibrium allocation it is useful to understand how this growth model with adoption margin leads to an endogenous cross-country income distribution where $c \in C$ is a country-index. Since all countries adopt technology from the same global frontier, which is the sum of ideas in each country, $A_F^W = \sum_c A_{F_c}$, productivity differences in A arise solely due to differences in technology adoption alone. Consider the productivity ratio of the home and foreign economy, $\frac{A^*}{A} = \frac{z^* A_F^W}{z A_F^W} = \frac{z^*}{z}$, which directly pins down the relative wages of production workers

$$\frac{w^*}{w} = \frac{z^*}{z}. \tag{45}$$

Since the adoption gap directly leads to a TFP gap, the model is consistent with the large literature on development accounting, which finds that differences in living standard are driven by productivity differences (Caselli (2005), Klenow and Rodriguez-Clare (1997)). Differences in specialization in innovation complicate the mapping from adoption to GDP slightly, and lead to country-specific skill

[34] Intermediate goods trade a la Krugman can be added without any complication. Moreover, I shut down the usual final goods differentiation assumption a la Armington or similar-looking models from Eaton and Kortum (2002) or Melitz (2003). A vast literature has studied the gains from trade in these models, see Costinot and Rodríguez-Clare (2014) for an overview. The focus of my analysis, however, rests on understanding the productivity slowdown, so gains from trade are not helpful in that endeavor.

[35] The production location of capital goods is related to a recent literature on multinational production and offshoring, see for instance Antras, Fort, and Tintelnot (2017) or Arkolakis et al. (2018). Since capital goods are assembled using capital, which in turn is produced using labor, the location of production for capital goods matters for wages and welfare. I avoid this complexity by assuming capital goods are produced locally. Moreover, while offshoring can potentially help understand negative wage effects, Arkolakis et al. (2018) show that quantitatively offshoring raises production worker wages. More importantly, in quantitative trade models offshoring delivers gains from trade which makes the productivity slowdown in Germany after the mid 1990s even more puzzling.

premia that are positively related to research activity, as I show below. However, since most labor is unskilled, and their relative wage is fully pinned down by z, differences in adoption are the primary driver of global inequality. The model ties a country's position on the global productivity distribution directly to how skilled labor on the firm-level is devoted to technology adoption

$$z_c \propto h_c^{\frac{\beta}{1-\theta}}. \tag{46}$$

The two country restriction is only important to solve for transition dynamics and the steady state results generalize to a setting with $|C| > 2$. The model is consistent with the view that human capital accumulation is central to the process of economic development (Lucas (1988), Lucas (2009b)) but it maintains the position that long-run growth requires idea-based technological change (Jones, 2005).[36]

Note how the adoption margin solves the problem of cross country "scale effects", i.e. the counterfactual implication of most growth (and international trade) models that given identical relative endowments and technology, the larger economy is more productive (see Ramondo, Rodríguez-Clare, and Saborío-Rodríguez (2016)). Note that adoption effort is unrelated to the total size of the labor force and only depends on the *share* of skilled labor devoted to adoption relative to production labor. There is thus no reason why a larger country should be more productive than a small one. This holds true since labor force growth leads to additional firm creation but leaves the ratio of skilled labor to production labor in the production sector unchanged. This extensive margin effect is reminiscent of Young (1998)'s work on growth without scale effects. If the measure of firms were fixed, a larger country would have a relatively higher skill share per firm which again would lead to troubling scale effects. Endogenizing the measure of firms in the production sector is thus essential for this model to deliver a sensible global income distribution. Scale effects do matter in innovation, so population growth and size show up there but since the technological frontier is global this effect cannot be identified in the cross-section.

Equilibrium in the Open Economy 2 Country setting: I assume for simplicity that the knowledge spillover A_F^{ϕ} is global,[37] which leads to the following law of motion of ideas in the home

[36] I abstract away from endogenous technological change that leads to different skill-requirements in production, a point made in Caselli and Coleman (2006) and Acemoglu and Zilibotti (2001). While I abstract away from differences in the aggregate production function, specialization into innovative activity leads to a similar pattern whereby skill-intensive innovation soaks up the relatively larger amount of skilled labor compared to an emerging market. See Malmberg (2017), Rossi (2022), Schoellman (2012), as well as Hendricks and Schoellman (2018) for empirical work on cross-country skill premia and development accounting. In my quantification I will pick parameter values that will imply that while a high skill premium lowers productivity and production worker wages, it will raises the real income of skilled labor. That is to say, scarcity dominates the negative effect of weak adoption within each country, and a poor country thus has very high income skilled workers. This implication can be avoided by introducing an additional layer of country heterogeneity, for instance one could let the adoption parameters be country specific ζ_c. Among rich countries, the implication seems more appropriate where skilled labor flocks to the US even though real income of low income households is below many other advanced economies.

[37] See Grossman and Helpman (1991a) for an in-depth discussion of this issue. Global knowledge spillover seem a natural assumption in a model of long-run growth so that the spillover is a function of the world stock of ideas.

economy.

$$\dot{A_F} = \frac{\left(A_F^W\right)^{\phi} H^F}{f_R} - \delta_I A_F. \tag{47}$$

The absence of trade cost ensures that $V_I = V_I^* = V^W$, and in combination with the free entry condition $f_R w_H \left(A_F^W\right)^{-\phi} = V$, it follows that the ratio of skilled labor devoted to innovation equals

$$\frac{h_F}{h_F^*} = \left(\frac{\frac{\gamma}{w_H}}{\frac{\gamma^*}{w_H^*}}\right)^{\frac{1}{1-\lambda}} \tag{48}$$

where I used that both countries have the same amount of production labor and $f_R = \frac{H_F^{1-\lambda}}{\gamma}$. The share of ideas produced in each country is denoted by χ, so that $\chi + \chi^* = 1$. Using the resource constraint in idea production (47), it follows that[38]

$$\left(\frac{\chi}{\chi^*}\right) = \frac{\gamma}{\gamma^*}\left(\frac{h_F}{h_F^*}\right)^{\lambda}. \tag{49}$$

Combining this expression with (48) and noting that $\frac{w_H}{w_H^*} = \frac{s}{s^*}\frac{z}{z^*}$ leads to

$$\left(\frac{\chi}{1-\chi}\right) = \left(\frac{\gamma}{\gamma^*}\right)^{\frac{1}{1-\lambda}} \left(\frac{s}{s^*}\frac{z}{z^*}\right)^{-\frac{\lambda}{1-\lambda}}. \tag{50}$$

Equation (50) highlights how the global share of ideas produced in the home economy is positively related to comparative advantage in research, and negatively related to the *cross-country* relative price of skill (not to be confused with the within-country skill premium). The negative link arises as innovation is less attractive when skilled wages are relative high, all else equal.[39]

To compute the price of skill in each country, one needs to solve for a set of global skilled labor market clearing conditions jointly. The steps are the same as in the closed economy except that an innovation earns profits in both countries now. In particular, I need to find the skill premium in each country, s and s^*, which pins down z and z^* and thus also delivers $\frac{w_H^*}{w_H} = \frac{s^*}{s}\frac{z^*}{z}$. Market clearing in the steady state in the rich and poor country reads

$$\begin{aligned} \left\{\frac{\chi}{z}\Lambda^{FO}\left((z)^{\frac{\tilde{\rho}}{g_A+\delta_I}+1} + (z^*)^{\frac{\tilde{\rho}}{g_A+\delta_I}+1}\right) + \Lambda^D\right\} &= s h^{tot} \\ \left\{\frac{\chi^*}{z^*}\Lambda^{FO}\left((z)^{\frac{\tilde{\rho}}{g_A+\delta_I}+1} + (z^*)^{\frac{\tilde{\rho}}{g_A+\delta_I}+1}\right) + \Lambda^D\right\} &= s^* h^{tot,*} \end{aligned} \tag{51}$$

where $\Lambda^{FO} = l^P \alpha \frac{g_F+\delta_I}{\tilde{\rho}+\delta_I+g_A}$. Note that both χ and z are functions of s and s^* so an equilibrium involves

[38]Proof: $\dot{A_F} = \gamma\left(A_F^W\right)^{\phi} H_F^{\lambda} - \delta_I A_F \Leftrightarrow \frac{g_F+\delta_I}{\gamma} = \frac{H_F^{\lambda}}{A_F} A_F^W \left(A_F^W\right)^{\phi-1} \Leftrightarrow \frac{g_F+\delta_I}{\gamma}\chi = \frac{H_F^{\lambda}}{\left(A_F^W\right)^{1-\phi}}$. Now you can do the same for the foreign economy and compute the ratio.

[39]Skilled wages are an equilibrium outcome. They might be high in a very innovative country but the country is not very innovative because skilled wages are high.

a set of skill premia that solve (51).

Balanced trade implies

$$\underbrace{\{Y+Y^*\}\frac{\sigma-1}{\sigma}\alpha\left(1-\alpha\right)\left(\chi-\chi^*\right)}_{\text{Innovator Profits/Royalty}}=\underbrace{Y^*-C^*-I^*}_{\text{Final good exports}} \tag{52}$$

where the emerging market trades final goods to make up for its net-import of ideas. Since capital goods are produced locally, only the royalty needs to be matched with exports, leading to (52).[40] In this open economy equilibrium comparative advantage allows for specialization in research activity, with distributional consequence and feedback effects on the level of domestic technology adoption.

A central results is that integration between symmetric countries, however, has no distributional effects as it leaves the skill premium unchanged, and delivers the standard variety gains from trade as in Krugman (1980).

Proposition 4. *Symmetric integration with $\gamma=\gamma^*$ & $h^{tot}=h^{tot,*}$ does not change the skill premium s nor the adoption gap z, but leads to welfare gains from trade.*

To understand this result, note that the market clearing condition is effectively unchanged by halving the share of research performed in the economy $\chi=\frac{1}{2}$ but simultaneously doubling the market size term $1+\left(\frac{z^*}{z}\right)^{\frac{\tilde{\rho}}{g_A+\delta_I}+1}=2$, which exactly cancels and leads to unchanged skill premia, and thus unchanged adoption gaps. Of course, there are twice as many capital goods available which raises productivity. This result is the same as in Krugman (1980) where trade integration leads to variety gains but leaves the measure of firms in each economy unchanged because foreign market access cancels exactly with foreign competition. This result can be generalized to many countries of different size as long as each country has the same research productivity and skill ratio.[41] It also highlights how globalization since the 1990s is fundamentally different from the early post-war integration efforts among the US and advanced European economies. Trade integration among similar countries induces no bias.[42] Heterogeneity across countries in terms of their fundamental research productivity changes this result.

To see this, in proposition 5 I consider what happens if the research productivity γ of the home economy increases.

[40] Note that total aggregate profits that accrue to innovators are proportional to total spending on capital goods, i.e. $\sum_c\int\pi_{j,c}dj=\frac{P_xX}{(1-\alpha)^{-1}}+\frac{P_xX^*}{(1-\alpha)^{-1}}$. These are the royalties that are paid each instant, and using the Cobb-Douglas assumption the following is true $P_xX=\frac{\sigma-1}{\sigma}\alpha Y$. Lastly, the term $(\chi-\chi^*)$ represents the gap between royalties received versus royalties paid, a difference that needs to be matched by final good exports.

[41] A crucial assumption for this result to be true is the CES technology that ensures that markups don't respond to market size. See Krugman (1979) and Melitz and Ottaviano (2008) for models of international trade with variable markups.

[42] This result is almost identical to neoclassical trade theory, which requires factor differences for trade to have any effect. Note that increasing returns in technology leads to additional returns as in Krugman (1980), and the presence of an adoption gap matters when considering the effects of asymmetric integration.

Proposition 5. *An increase in the home economy's absolute advantage in research γ, given that $\beta + \theta < 1$, leads to an increase in the skill premium in the home economy while the skill premium in the foreign economy falls. Proof see appendix A.6.*

Proposition 5 is important when contrasting against the closed economy result from the previous section where I argued that improvements in the research technology have no effect on the allocation of skilled labor across sectors. This is no longer true in the open economy where an improvement in the research technology means that a larger share of world research will be performed in the home economy. This raises the demand for skilled labor, and in turn pushes up the skill premium. Comparative advantage and openness shape the interaction between innovation and adoption in ways that are absent in the closed economy. The inequality $\beta + \theta < 1$ bounds the negative effect on an increase in the skill premium on productivity, which matters for the theoretical result here and the quantitative application below. This inequality is also respected when matching data moments in the quantitative application.

Special Case: *Suppose that $\lambda = 1$ & $\gamma^* = 0$*

To obtain sharp implications, and to simplify the quantitative application, I focus on a particularly tractable scenario where all research is performed in the advanced economy. This represents an extreme case of *asymmetric integration*, by which I mean that research is only produced in the advanced economy. While the foreign economy uses technology, and its skilled labor is fully devoted to adoption of technology, it does not contribute any ideas to the global technological frontier. 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, but with almost no patenting activity in emerging markets and Eastern Europe.[43] For more recent years, this assumption may be less appropriate as China is starting to contribute to the global technological frontier.[44]

One issue that arises is how to model the emerging market in the closed economy. If there is no access to technology, income would be zero, which is counterfactual. For simplicity, I assume that some innovation trickles through, perhaps due to government funded imitation effort that is left unmodeled. This allows the economy to produce capital goods on the range $[0, z_0^*)$ with a large but constant initial adoption gap z_0^*.[45]

After goods market integration, the skill premium in the advanced economy increases. This increase

[43] The contribution of Eastern Europe at the time is so small that it ends up in a residual category. Germany on the other hand is the country with most patents in Europe.

[44] Bergeaud and Verluise (2022) provide evidence from patent data suggesting that China is contributing as much as the USA to the technological frontier in recent years. Studwell (2013) offers a different perspective, based on a case study of the High Speed Rail Technology in China, where superficial improvements and a relaxation of safety standard were hiding a fundamental lack of innovation.

[45] Moreover, for simplicity suppose that there is no markup on these capital goods but efficiency is lower in the emerging market so that the price per effective unit of capital is the same. That is, to produce a unity of capital good, $\frac{1}{\alpha}$ units of physical capital are needed. This avoids dealing with price heterogeneity across capital goods in the integrated equilibrium but is otherwise inconsequential.

is driven by demand for frontier technology in the emerging market, and directly related to the fact that foreigners adopt technology. A potential innovators takes into account that profits accrue both at home and abroad, and the free entry condition into innovation now includes foreign profits as well

$$V_I = \underbrace{\left(\frac{\alpha}{\tilde{\rho}+g_A+\delta_I}\right)\frac{L^P w}{A^F} z^{\frac{\tilde{\rho}}{g_A+\delta_I}}}_{\text{same as closed economy}} \left\{ 1 + \underbrace{\frac{L^{P,*}w^*}{L^P w}\left(\frac{z^*}{z}\right)^{\frac{\tilde{\rho}}{g_A+\delta_I}}}_{\text{additional market size effect}} \right\}. \tag{53}$$

Equation (53) reveals that the strength of the idea demand shock depends on i) the foreign adoption gap $(z^*)^{\frac{\tilde{\rho}}{g_A+\delta_I}}$, and ii) foreign GDP summarized in $L^{P*}w^*$ *relative* to domestic variables. I assumed equal sized countries so L^P cancels and the reader can confirm that this expression is consistent with the more general result in (51) when using $\frac{w^*}{w} = \frac{z^*}{z}$. The model can easily accommodate countries of different size as equation (53) shows, and a larger foreign labor force exerts more pull on domestic innovation.

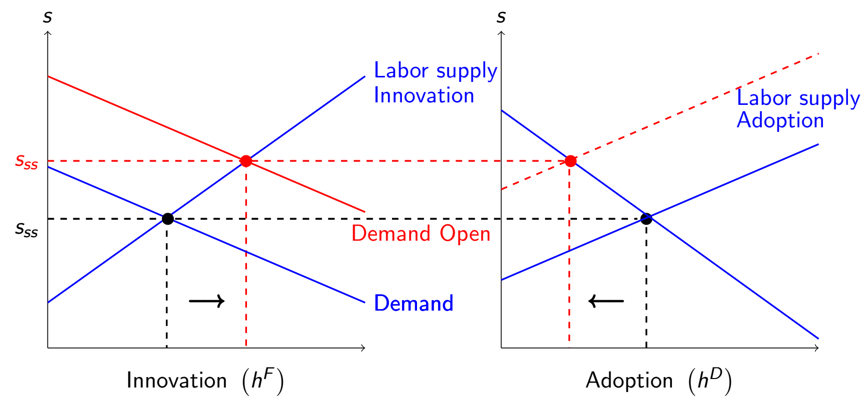
Note that market integration directly increases the market size of innovators, which raises profits that are arbitraged away by increasing entry into innovation. Importantly, convergence in the emerging market further raises the returns to innovation. Note that a rising wage rate ($w^* \uparrow$) and a declining adoption gap ($z^* \uparrow$) both push up the value of an idea. In a model where technology is endogenous, fast adoption in emerging markets and rising returns to innovation in advanced economies are two sides of the same coin.

Adoption-driven growth in emerging markets thus leads rising demand for skilled labor in advanced economies driven by an expansion of the research sector. In general equilibrium this brings about an increase in the relative price of skill in the advanced economy and a reallocation of skilled labor from adoption to innovation. Since capital supply and firm entry is perfectly elastic, the factor that is capturing the benefits from market integration in advanced economies is skilled labor. Figure 5 summarizes the main argument of this paper in a simple supply-demand plot.

While innovation and adoption were characterized by a strong complementarity in the closed economy, factor market rivalry and competition for skilled labor characterizes the relationship between innovation and domestic adoption in the open economy. While innovation is still responding to adoption, it is responding to *foreign adoption* which drives a wedge between innovation and domestic adoption that did not emerge in the closed economy.

To summarize, in the advanced economy innovation takes off, adoption recedes, and inequality increase after market integration. These are qualitative insights that hold in general in this type of model given asymmetric integration. The emerging market catches up with the advanced economy, the extent to which depends on how much skill they have available to adopt technology. The faster they adopt, the stronger is the pull on innovation in the advanced economy. In order to compute

Figure 5. Market Clearing for Skilled Labor in Open Economy



the aggregate rate of growth and the exact increase in the skill premium, I simulate the model after pinning down the relevant parameters. Before I turn to this quantitative application, I conclude the theoretical section by considering how the theory relates to recent work on skill-biased technological change and the effect of declining population growth that have served as main explanations for rising inequality and weak productivity growth, as well as several extensions of the baseline model.

3.1 Discussion and Extensions

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 see how my model relates to this large literature.

First, a more realistic model would include skill-biased technological change as virtually all sectors in Germany (and other countries) become more skill-intensive over time.[46] I abstract away from this secular trend to show what my approach can add to this well-known literature. A useful feature of the model is that it breaks the positive link between inequality and growth that is inherent to most theories of skill-biased technological change. As pointed out in Acemoglu and Autor (2011), skill-biased technological change generates wage growth *for all workers.* The reason is the strong complementarity

[46] I highlight in the data section how empirically skill-growth was faster in one sector than the other. Yet, it is the case that the share of skilled labor is increasing in all sectors consistent with secular skill-biased technological change, see figure ?? in the appendix.

between high and low skilled workers which ensures that biased technological change benefits everyone. 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 a rise in the skill premium can hamper economic growth by reducing equilibrium adoption effort.

Second, a related literature has focused on the task content of work (Autor, Levy, and Murnane, 2003) and automation (Acemoglu and Restrepo, 2018b) which is able to generate more inequality with less overall aggregate growth.[47] It is still true, however, that technological change pushes out the production possibility frontier so the growth slowdown remains puzzling. Combining task-based models with the endogenous technology adoption margin is promising to generate negative aggregate effects as I show next.

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 production and skilled labor enters the production function ($\tilde{\beta} = 0$ is the baseline case in the paper).[48] The model remains mostly unchanged except now there is an additional term $\tilde{\Lambda}_{\tilde{\beta}}$ when solving for skilled labor market clearing,

$$\frac{1}{s}\left(\tilde{\Lambda}_F z^{\frac{\tilde{\rho}}{g_F+\delta_I}} + \tilde{\Lambda}_D + \tilde{\Lambda}_{\tilde{\beta}}\right) = h^{tot}. \tag{54}$$

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 the innovation and adoption.

Note, however, that an increase in the relative price of skill driven by a changing task content of 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. This is at odds with the data as I show below, and the effect of globalization on the returns to innovation will resolve this tension.

Population Growth Slowdown and Business Dynamics: A compelling explanation for sluggish productivity growth is based on the effect of declining population growth on TFP in (semi)endogenous growth models (Jones, 2020; Peters and Walsh, 2019). Note that a population growth slowdown in the benchmark model of Jones (1995) would not be able to generate increasing levels of innovative effort, nor would it lead to rising inequality (even if there were two types of labor as in Romer (1990)). Slower population growth induces slower productivity growth which requires a smaller share of labor

[47] Another seminal paper on real wage losses of low skilled workers is Caselli (1999) which focuses on learning barriers and capital reallocation.

[48] Acemoglu and Restrepo (2020) show how to micro-found this Cobb-Douglas production function in a model of automation.

devoted to the production of new ideas.[49] Yet, an increasing share of employment is devoted to research activity, see Bloom et al. (2020) for the US, and evidence that I compile for Germany in section 4. The open economy model, where a push for innovative effort is driven by a rising global demand for ideas, rationalizes rising research activity in advanced economies. Moreover, the skill premium plays an important role in my theory by impacting equilibrium adoption effort, a margin that is abstracted away from in most of the literature on endogenous growth. This margin allows me to directly addresses recent empirical findings of Andrews, Criscuolo, and Gal (2016) highlight stalling adoption as an important factor for the growth slowdown, which seems unrelated to the decline in population growth.

An important assumption in the baseline model is that the entry cost into the intermediate goods sector are paid in production labor. The downward sloping relationship between the skill premium and the demand for skilled labor for adoption purposes on the firm level is directly related to the fact that long-run firm profits are proportional to the cost of entry, which in turn is proportional to production worker wages.

If one were to generalize the entry cost to be a Cobb-Douglas aggregator, i.e. $f_e w^{\mu} w_H^{1-\mu}$, the elasticity of a rising skill price on adoption would become

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

which creates a weaker response of the skill premium on technology adoption since $\mu < 1$. Note, however, that there would be an additional effect on firm entry since rising entry costs require a smaller number of firms in equilibrium to ensure profits are sufficiently high. I abstract away from this margin for simplicity. However, missing firm entry and slowing firm dynamics have been documented and found to be important for weak economic growth, see for instance Decker et al. (2017), Decker et al. (2020), or Karahan, Pugsley, and Şahin (2019). A rising skill premium will negatively affect firm entry whenever firm-entry is a relatively skill-intensive activity so the framework might be useful to understand this pattern as well.[50]

In appendix B, I also consider the effects of immigration, and I argue how more general factor intensity differences across sectors changes the results, i.e. innovation may also require some production labor. I also discuss how endogenizing the high-skilled labor supply, i.e. h^{tot} becomes an upward sloping function in s, changes the results.

[49] See footnote 17 in Jones (1995).

[50] This point is related to Salgado (2020) where skill-biased technological change leads to less entry into entrepreneurship.

4 Quantitative Results

4.1 Calibration of the Model

To calibrate the model I need to pin down a number of parameters.

Growth $\{g_L, \phi, \delta_I\}$: In this semi-endogenous growth model long-run growth is fully driven by the interaction of population growth with the knowledge spillover embedded in the idea production function $g_F = \frac{g_L}{1-\phi}$. I follow Jones (1995) and more recently Bloom et al. (2020) and set $\phi = -1$. Population growth in Germany has been low at a rate below 0.2% from 1980 − 2015, based on data from the PWT. On the other hand, growth in skilled labor, which is the crucial input in idea creation and adoption, has been growing at a rate of 3.1% over the same time period, using the PWT in combination with the Barro-Lee data set. Presumably, not all skilled labor is "skilled enough" to play a role in the idea-generating process so picking a population growth rate of 3% seems likely to high. On the other hand, improved educational attainment might reasonably have an impact on production labor where workers are supplying more effective units. Weighing these considerations against each other, with the goal in mind to settle for a reasonable medium-run growth rate in "effective" population, I decided to assume a long-run population growth rate of 2% with fixed high-skill-to-production labor share. This implies a long-run per capita growth rate of 1%. I also need to pick the rate at which capital goods disappear. I assume $\delta_I = .04$, which implies that a capital good is in use for 25 years on average.

Convergence $\{\theta, \beta, \alpha\}$: 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 a number of covariates and in particular human capital, is close to $-.02$. I linearize the law of motion of z around its steady state to pin down θ to match these cross country convergence patterns. The linearization leads to

$$\frac{\dot{z}}{z} \approx \underbrace{(1-\theta)(\delta_I + g_F)}_{=\hat{\beta}_B}(\log z_{ss} - \log z_t) + \beta(\delta_I + g_F)(\log h_t - \log h_{ss})$$

so that given $g_F = 1\%$ and $\delta_I = 4\%$, a reasonable estimate for θ is thus 0.6 which ensures that $\hat{\beta}_B = -.02$. This leads to slow convergence dynamics relative to a neoclassical model.[51] While θ plays a similar role to the capital share in the neoclassical model by shaping the speed of convergence, the interpretation is different and relates to the advantage of backwardness that generates fast productivity growth in emerging markets. Moreover, I set α to be equal to .5. Once one takes into account that there are overhead labor costs both in terms of production labor for firm entry one arrives at the usual share of capital in total income of 33%.[52]

[51] Mankiw, Romer, and Weil (1992) extend the Solow model to include human capital to increase the share of reproducible factors which allows them to slow the convergence dynamics.

[52] If the capital share is measured as firms' spending on capital goods, then $\frac{p_x X}{Y} = \alpha * \frac{\sigma-1}{\sigma} = .5 * 2/3 = 1/3$.

To pin down β I rely on cross-country income differences. Real wage differences for production workers across countries are fully captured by z_c

$$z_c = \left(\frac{\zeta h_c^{\beta}}{g_F + \delta_I}\right)^{\frac{1}{1-\theta}} \tag{55}$$

so the real wage in any country is proportional to $h^{\frac{\beta}{1-\theta}}$. Conditional on a distribution of the relative amount of skilled labor devoted to adoption across countries $\{h_c\}$, the parameters $\{\theta, \beta\}$ jointly translate this initial distribution into observed cross country wage gaps. A small β leads to small cross country income differences.[53] Taking logs of (55) and adding a measurement error u allows me to back out β by running the following regression

$$\log z_{ct} = \alpha + \delta_t + \frac{\beta}{1-\theta} \log h_{ct} + u_{ct}. \tag{56}$$

The slope coefficient through the lens of the model equals $\frac{\beta}{1-\theta}$ where I proxy for production worker wages using GDP per capita and I proxy for h using the share of college-educated workers in each country, i.e. h^{tot}. Since most countries don't perform frontier innovation this simplification should not bias the results dramatically in a large cross section of countries.[54]

I combine data from Barro and Lee (2013) with the PWT and run the regression for the year 2015 to capture the post-integration steady state where more countries have moved toward a market-based open economy. I obtain a coefficient (robust standard error) of .9 (.06) with an R-squared of 65%, as can be seen in figure 6. Given that θ is .6, β has to be around .35. I am able to explain much of the variation in cross-country income differences even though I assume that all countries have access to exactly the same adoption technology and preferences, which I view as desirable from a theoretical point of view. Clearly, this exercise is not a causal one and merely serves as a first step to transparently obtain an estimate for β through the lens of the model. I will assess the quality of this initial cross-sectional based estimate when computing transition dynamics in the simulated model for Germany and compare them to growth dynamics observed in the data, which will lend additional credibility to my calibration.

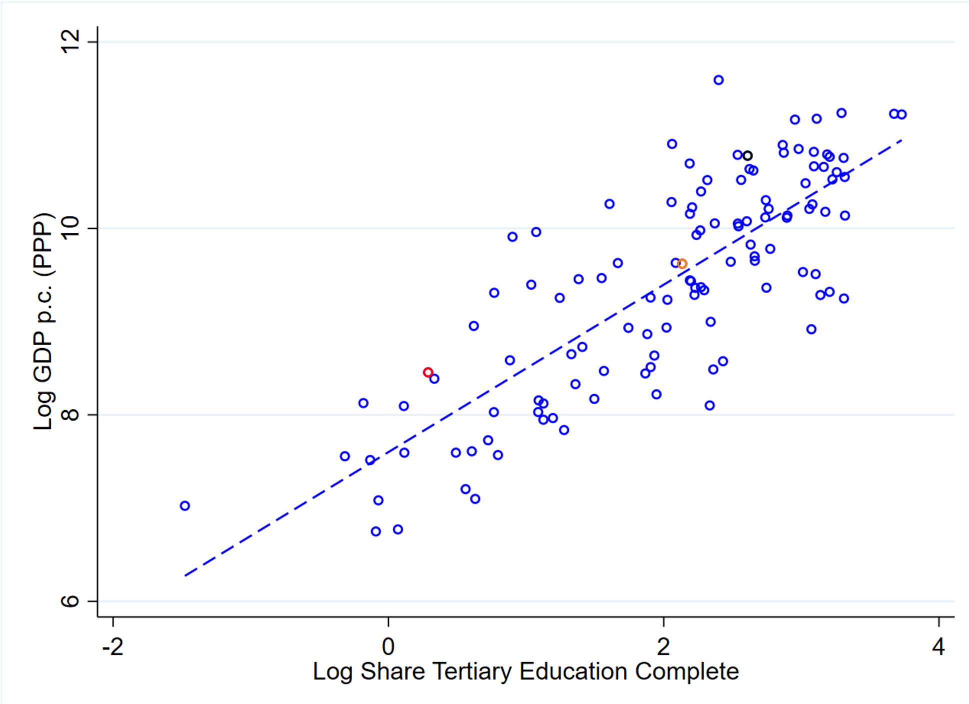
Elasticity of Substitution $\{\sigma\}$: I take the elasticity of substitution from Broda and Weinstein (2006) and pick a value of 3 which is close to the median estimate in their study.[55]

[53] In a closed economy, the logic of the model does not work since the economy would not be able to adopt frontier technology and its human capital would allow no inference on its level of technological sophistication. Soviet Russia – with strong scientists yet weak technological capabilities – is a case in point.

[54] A more sophisticated measure could try to incorporate country differences in innovation which would for instance help position of the US on top of the world income distribution for instance.

[55] This may suggest a profit share that is very large, compared for instance to the estimate of Basu and Fernald (1997). This is not the case, however, as the net profits of the firm are *not* given by revenue over demand elasticity, $\frac{r}{\sigma}$, because there is an additional overhead adoption cost that needs to be subtracted. In fact, depending on the size parameters of the model the profits of the firm expressed as a percent of revenue can become vanishingly small when close to violating the inequality stated in proposition 1. Precisely this inequality suggests a smaller σ is appropriate so at to avoid "too

Figure 6. Cross Country Inequality & Skilled Labor Ratios



Data from PWT 10.0 and Barro and Lee (2013). I drop countries with less than 1 mio people, and focus on the log share of completed tertiary education. I plot the link between log real per capita GDP (PPP) and the log the share of completed tertiary education for 2015. The red dot represents Congo, orange is Brazil, and black is Germany.

Skill-Share and adoption efficiency $\{\frac{H}{L}, \zeta\}$: I pick the skill-intensity to be equal to .15 which slightly above the average share of the population over the period $1980 - 2015$ that has a college education, according to Barro and Lee (2013).[56] Aiming for a skill premium of 2 in 1994, I pick the parameter ζ to be equal to .23. Intuitively, a large ζ pushes up the skill premium. As adoption gets easier, the waiting time declines and innovation becomes more profitable, raising the overall demand for skilled labor. This can be seen from equation (44).

Foreign Economy $\{z^*, L^*\}$: The strength of the market size shock depends on the size of the foreign market. As I have shown in the theory section, what matters is foreign GDP or $w^* L^{P*}$, and the foreign wage rate is a function of z^*. Moreover, z^* also matters as it shows up in the adoption friction which pins down how long it takes for a domestic innovation to become profitable on the foreign market. On the one hand, the rise of the East and Far East, to borrow a term from Dauth, Findeisen, and Suedekum (2014), involves literally billions of people so the market size shock should be massive. On the other hand, Germany is not the only producer of frontier technology and comes in third after

much" adoption.

[56]When only requiring some tertiary education, the ratio goes up to 20%, which is still low compared to other advanced economies.

the US and Japan. I proceed by assuming that market integration happens between two equal sized countries where I assume that the foreign relative technology level shifts from $z^*_{1995} = .2$ to $z^*_{2015} = .4$, and $L^P = L^{P*}$. This development story is consistent with a relative skill share of .04, where the initial equilibrium z^*_{1995} was produced by some force outside of the model, for instance state-guided imitation in the Soviet Union. I assume that from the point of view of the domestic economy, the initial demand from abroad is zero (due to the Iron curtain) but some technology trickled through nonetheless so that I don't overstate the degree of poverty in Eastern Europe during the 1990s.

4.2 Quantitative Results and Aggregate Evidence

For now I compare two steady states, where the initial equilibrium is in autarky while the new steady state is characterized by an integrated goods market between advanced and emerging economies. I then compute the log difference of wages in the two equilibria and divide through by 20 to obtain the average growth rate over a 20-year period. I am in the process of computing the full transition dynamics and will add them soon. I focus on German and Poland when comparing theory to data. Germany is a useful benchmark because i) it is a large open economy with ii) an active innovation sector that iii) experienced a major integration shock due to the fall of the Iron Curtain. I pick as base of comparison Poland, which is in the middle of the pack among Eastern European countries, poorer than the Czech Republic but richer than Romania.

Table 1 and table 2 summarize wage growth for production and skilled labor, both in the model and in the data.[57] The most remarkable result is that in the new steady state real wages for production workers in Germany are permanently lower compared to the balanced growth path under autarky. While real wages stagnate, wages for skilled labor grow at a rate of 1.5%, i.e. 50% above the long run growth rate of 1%.

Table 1. Wage Growth Production Labor (1995 – 2015)

	Model	Data
Germany	0.19%	0.13%
Poland	4.5%	3.8%

While the model does a good job at matching stagnant wage growth for production workers, the wage growth for employment in innovation (skilled labor) measured in the IAB data is weak compared to the model. This relates to an ongoing debate on the rise of the skill premium in Germany and top

[57] I keep this exercise as simple as possible and focus on raw wage data, without controlling for observables such as age and sex.

Table 2. Wage Growth Skilled Labor (1995 – 2015)

	Model	Data
Germany	1.54%	0.4%
Poland	NA%	5.1%

The wage data for Poland comes from the LIS income databased LIS (2022), and I select production workers as employed workers with low and medium levels of education (educ == 1, educ == 2) while skilled workers are defined as the one with a level of education of 3. I compute a simple weighted average with population weights for 1995 and 2015 to capture a broad trend in the economy. To see changes in inequality, I run a regression of log real income on a standard controls (age fixed effects, sex, marriage) on year fixed effects and year-specific education dummies. I know allow for education dummies to take on all three categories, and the both the premium for middle and high levels of education increases by 7% and 12%, respectively. Regression results are reported in appendix XX. I apply a CPI to compute real wage growth. I use the series POLCPIALLMINMEI (annual average cpi all goods Poland) from Fred, downloaded on August 8 2022. The data for Germany comes from the IAB BHP establishment panel, additional information on this datasource is contained in the appendix. Figure 11 summarizes divergent trends in daily wages across workers in sectors in innovation and production where additional details on this sectoral classification is in the appendix.

coding issues in the Germany data.[58] Using the Klems data in figure 7 shows a sizable increase in the skill premium of more than 10% that is consistent with a rising Gini index. In any case, inequality has gone up in Germany just like in almost any other advanced economy, and the 1.54% growth for skilled labor seem reasonable considered against the broad trend in advanced economies and a more general interpretation of what the skill premium stands for in my model: skilled labor that is able to develop new ideas (innovation), or implement new ideas in a new context (adoption). Not everyone who has obtained a college degree will fall into this category. The growth rates can directly be translated into a rise of the skill premium of 28 log points in Germany, i.e. 33%, and the Gini index would increase by 6pp which accounts for 75% of the observed increase in the data (compare to 8pp in the data).

The steady state skill premium in Poland that is implied by the growth spurt from $z_0 = .2$ to $z_T = .4$ is $s^{POL} = 3.7$. I don't compute growth in the skill premium since it is unclear what the right baseline is. If one assumes that there is no skill premium in communist Poland, the skill premium would increase by 270%. This is definitely an overstatement but it seems highly likely that in a planned economy the skill premium was depressed since adoption activity was restricted. The implied skill-to-labor ratio that is needed through the lens of the model to generate the growth spurt is much lower than the actual share in Poland, which is highly educated like many other Eastern European economies. As mentioned in the open economy theory section, the model needs a poorer economy to have a larger skill premium so as to deliver a wide adoption gap.[59] One key feature of recent economic growth is

[58] See Dustmann, Ludsteck, and Schönberg (2009) who find a rise in the skill premium, while Doepke and Gaetani (2020) find none. A major problem is that the administrative wage data is top coded, and even if a right tail is imputed, likely to be understating the true extent of wage growth in the right tail of the distribution.

[59] If one used that value, wages for production workers in Poland would be *higher* than in Germany. In the current calibration I maintain that German production workers earn 30% higher wages in the new steady state. Given that all innovation occurs in Germany, and maintaining a 30% wage gap, the model wants Germany to have a much larger college

the rise of inequality in developing economies, see Goldberg and Pavcnik (2007). The model matches this fact since technology adoption is an inherently skill intensive activity. After market integration, there is much technology to adopt, so the demand for skill goes up, which ultimately drives up the skill premium in the emerging market.

The way I have defined TFP($= A_F * z$) means that production worker wage growth is identical to TFP growth. Note that GDP growth is higher than the measure of physical productivity TFP($= A_F * z$) due to gains from trade that show up in the wage rate of skilled workers.[60] Simply adding up model-based wage growth across workers implies that overall the economy experiences a growth slowdown that I project onto this 20year horizon. In that period GDP growth would only be .4% relative to the long-run trend of 1%.

The key insight to understanding the very negative results on production worker wages in Germany is to consider the trade-off between innovation and adoption. The aggregate effect of reallocating labor from one activity to the other depends on the strength of the dynamic knowledge spillover parameterized in ϕ. As skilled labor moves from adoption into innovation, the adoption gap has to increase, which has a first-order negative effect on real wages. In general, more innovation pushes up the real wage due to growth of the technological frontier A_F. If, however, ideas are difficult to find with a knowledge spillover of $\phi = -1$, this effect is relatively weak and dominated by the lack of adoption from the point of view of production workers. In a sense, the economy sacrifices high-return adoption effort for innovative activity that runs into diminishing returns quickly.

Benchmarking against Romer and Jones: The discrepancy between the disappointing growth effects in advanced economies in the model, and the strong pro-growth effects found in for instance Rivera-Batiz and Romer (1991), and more recently in the quantitative trade literature,[61] is explained by the weak knowledge spillover, and how it interacts with technology adoption. One can resurrect the strong pro-growth effects of integration by dropping the domestic adoption margin *and/or* raising the the knowledge spillover. When keeping the adoption margin, production workers in rich countries become indifferent between autarky and integration when $\phi = .21$. Adoption still takes a hit but growing innovation, relative to the long run trend, exactly offsets declining technology adoption so that the real wage is unchanged relative to trend. GDP would be higher due to rising skilled worker wages.[62]

share than Poland. The picture looks less bleak when taking into account more backward regions in Eastern Europe such as Albania. Another possibility is that the same level of schooling leads to different effective units of skilled labor, effectively rendering skilled labor more scarce in Poland, see Schoellman (2012) and Hendricks and Schoellman (2018).

[60] A recent literature studies theoretical measurement questions of incorporating the gains from trade properly into domestic productivity statistics, see for instance Baqaee and Farhi (2019), which I don't pursue here.

[61] See for instance Alvarez, Buera, and Lucas Jr (2013), Sampson (2016), Lind and Ramondo (2022), Perla, Tonetti, and Waugh (2021), and Buera and Oberfield (2020) for models building on ideas flows and knowledge spillovers, especially Lucas (2009b) and Kortum (1997). See Hsieh, Klenow, and Nath (2019) for a Schumpeterian growth model with strong scale effects, and Cai, Li, and Santacreu (2022) and Somale (2021) building on the quantitative global growth model of Eaton and Kortum (2001).

[62] I maintain a long-run growth rate of 1%, so I have to adjust the population growth as follows $g_L = (1 - \phi) * 0.01$, since it is not sensible to compare economies with different autarky long run growth rates.

Shutting down the adoption friction and assuming that there is only one type of labor L with a competitive production sector as in Jones (1995) leads to the following results in the open economy. Market integration, even while maintaining $\gamma^* = 0$, effectively looks like a labor supply shock that leads to increasing productivity and real wages at the rate $d\log w = \frac{1}{1-\phi} d\log L$. Both economies experience a productivity boom due to increasing specialization in research in the advanced economy. Frictionless trade replicates the full-integration equilibrium now with factor price equalization, which is identical to an autarky equilibrium when doubling the labor force. This leads to a transitory growth spurt that adds up to an average increased rate of per capita growth computed over a 20 year horizon of 2.75%, almost 2% above trend. Since there is only one type of labor in Jones (1995) there is no inequality.

The model of Romer (1990) does feature two types of labor, and skilled labor is used both in innovation and production. There is no adoption gap, however, and skilled labor enters production in a Cobb-Douglas fashion, something that I considered in the extension on skill-biased technological change 3.1. Note that the original model of Romer features strong scale effects, like the Schumpeterian counterparts in the early literature(Aghion and Howitt (1990), Grossman and Helpman (1991b)). This leads to a higher long-run rate of growth after market integration. This may seem like an odd prediction nowadays but was indeed the hope in the early 1990s.[63] Combining the models of Romer and Jones may thus be the best comparison. Even in that scenario with diminishing returns in research productivity, the fact that i) the advanced economy is specializing more in innovation, and ii) there is no adoption friction, implies that measured TFP defined as A_F must increase due to market integration. That is, TFP in the production sector increases after controlling for the changing composition of labor, an adjustment that is indeed made in growth accounting decompositions.

Efficiency: The adverse effects on growth suggests that the economy is inefficient. First, there are markups in the innovation sector and the production sector, which leads to insufficient size of the sector whichever has the higher markup, see Baqaee and Farhi (2020) for an in-depth analysis of this issue. Moreover, there is a knowledge spillover embodied in ϕ that renders research inefficient. While in Romer (1990) the knowledge spillover is the key source of underinvestment in research, it works the other way around here because $\phi < 0$. This means that a so-called "fishing out of ideas" occurs where each innovator does not take into account that they make future research harder. Lastly, there is a local knowledge spillover in production whereby incumbents do not internalize that their adoption efforts raise the productivity of future entrants. The fact that the economy responds so poorly to the market size shock suggests that over-investment in research *relative to adoption* is a problem. The paper does not imply that too much research is performed relative to producing final output, the tradeoff studied in Romer (1990) or Jones (1995). What the paper suggests is that given a fixed amount of resources devoted to innovation and adoption, over-investment in research *relative to adoption* becomes a distinct possibility when ideas are becoming harder to find.

[63] See Baldwin (1989)'s take on the effects of European integration on growth

To see this, derive the planner solution in the autarky steady state.[64]The solution to (57) is

$$\frac{\beta}{1-\theta}\left(1-\phi\right) \quad = \quad \frac{H^{D}}{H^{F}} \tag{58}$$

and formalizes the previous discussion. When the knowledge spillover is weak, and ϕ is negative, the planner wants to allocate more skilled labor to adoption. In contrast, when adoption has no effect ($\beta \approx 0$) or the advantage of backwardness is strong ($\theta \approx 0$), the planner shifts more labor into frontier research. Given my calibration, the market equilibrium delivers a ratio of adoption labor to innovation labor of .65. while the planner solution is 1.75 in autarky, i.e. the planner allocates more labor to adoption purposes. Note that a planner that cares about world output faces a different choice. Suppose the planner puts equal weight on domestic and foreign income, and again only decides on the allocation of skilled labor within the advanced economy. Then the maximization problem becomes $\max 2 * \log\left(A_F\right) + \log z + \log z^*$ and the optimal solution would be $\frac{\beta}{2}\left(\frac{1-\phi}{1-\theta}\right) = \frac{H^D}{H^F}$. Intuitively, pushing out the frontier helps both the domestic and foreign economy so more labor is devoted to producing frontier technology. The planner solution now requires a ratio H^D/H^F of 0.875, which is closer to the decentralized equilibrium. Inefficiency thus depends on whether the planner has a domestic or global perspective.

The result of over-investment in research is in contrast to Jones (1995) which finds that underinvestment in research is the more likely outcome, even if $\phi = -1$. First, the two results are not directly comparable as I don't feature a dynamic tradeoff between consumption today vs. tomorrow. Instead, finding the optimal allocation between innovation and adoption that maximizes output today also maximizes output tomorrow. Second, even if I had stayed closer in modeling choice to Jones (1995), the results would be different because I have a markup in the production sector too. Jones's result of under-investment in research is driven by the large markup in the research sector relative to no markup in the production sector. Since in my model markups in production and innovation (3 vs. 2) aren't that different, this channel is less relevant and the inefficiency due to $\phi < 0$ becomes more important.

It is clear that subsidizing innovation in this model is counterproductive as it further amplifies the initial inefficiency. The skill premium widens even more, and growth slows as more labor is reallocated away form domestic technology adoption, which was under-supplied to begin with.

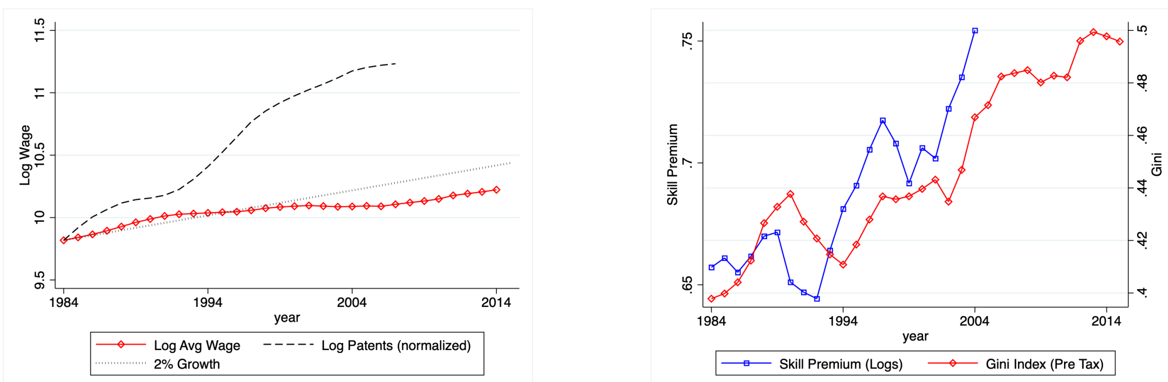
Aggregate Evidence: A unique feature of the theory is to reconcile rising innovative effort

[64]Given log utility this is found by maximizing $\log\left(A_F z\right)$. Next, suppose a planner allocates skilled labor between adoption and innovation but leaves the rest of the equilibrium unchanged, and in particular takes the measure of production firms as given. Then, the relative technology level z is proportional to $\left(H^D\right)^{\frac{\beta}{1-\theta}}$ while the total number of ideas is proportional to $\left(H^F\right)^{\frac{1}{1-\phi}}$. Picking up some constant parameters in Λ I obtain the following system

$$\begin{aligned} & \max \log\left(A_F z\right) \\ s.t. & \\ z \quad & = \Lambda_z \left(H - H^F\right)^{\frac{\beta}{1-\theta}} \\ A_F \quad & = \Lambda_F \left(H^F\right)^{\frac{1}{1-\phi}} \end{aligned} \tag{57}$$

measured in the patenting rate occurs against the backdrop of stagnant real wages and weak TFP growth as can be seen in figure 7. Wages grew at a rate above 2% up until 1995, from which point onward Germany experiences its worst two decades of economic growth since WW2 where per capita income growth fell to a meager 0.55% annually, even though patent activity, a proxy for innovation, grew exponentially and at an elevated rate as shown in figure 7. Careful evidence documenting a productivity slowdown is provided in Van Ark, O'Mahoney, and Timmer (2008) who use a growth accounting framework to obtain a residual measure of TFP. German TFP growth from 1995-2004 is estimated to be .3%, an all time low in post war history.[65] This pattern of robust innovative activity, weak productivity growth, and stagnant real wages is not unique to the German economy but seems to hold across a wide set of advanced economies, also the US.[66] This is a puzzle for benchmark models of endogenous growth, but the quantitative exercise has shown that the model proposed in this paper precisely explains these puzzling patterns. [67]

Figure 7. 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 production 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).

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 7 shows the evolution of the skill premium, and the Gini Index, both of which shoot up in the mid 1990s.[68]

[65] See table 4 in Van Ark, O'Mahoney, and Timmer (2008).

[66] Results for the US available upon request.

[67] Note that a declining labor share as argued in Karabarbounis and Neiman (2014) is not able to quantitatively account for weak wage growth. Using the KLEMS data, the labor share from 1995 to 2004 fell only from 67.7% to 65.6%. Assuming constant GDP per capita growth of 2%, this would have led to average wage growth in that period of 1.65%. Moreover, note that automation or investment specific technological change, the most popular explanation for a declining labor share, should lift GDP growth up. Recent work has cast doubt on the global decline of the labor share, see Gutiérrez and Piton (2020) and Koh, Santaeulàlia-Llopis, and Zheng (2020).

[68] See also the work of Card, Heining, and Kline (2013) on rising Germany inequality in the 90s and 2000s who find

Consistent with rising patenting activity, the baseline calibration predicts on average .73% of extra skilled employment growth in the innovation sector, while the production sector looses 1.40% per year on average over the 20 year period. On a balanced growth path ideas grow at the same rate as wages in the long run, and the decoupling of the two visible in 7 is explained by weak technology adoption during an episode of globalization that drives apart the returns of *local* adoption vs *global* innovation.

There are a number of concerns using patent data to proxy for innovation, not least that patents only reflect a very small share of innovation and productivity growth.[69] I alternative approach is to look at employment growth patterns across sectors where I assign establishments into an innovative and a production sector. I use the IAB BHP establishment sample that comprises a 50% random sample of German establishments with detailed sectoral classification. I define the innovation sector to be comprised of establishments that are active in sectors such as consulting, research, patent law, headquarter services, etc. The production sector is the rest of the economy. I thus follow a broad notion of "innovative" employment, and I am missing out on research activity in production firms.[70] A detailed discussion can be found in the appendix, but the idea is to map the simple two-sector structure in the theory into the sectoral classification in the data, so that differential sectoral employment growth rates can be interpreted as rising returns to innovation in an open economy. This approach is related to the firm-growth accounting framework in Garcia-Macia, Hsieh, and Klenow (2019) that highlights how to use employment growth patterns to tell apart different sources of productivity growth. In equilibrium, rising returns to innovation show up as elevated firm entry rates and rising total employment in the innovation sector.

The left panel in figure 8 shows a massive increase in the relative employment share in innovation. The model falls short of replicating a tripling of the relative share of research employment, which seems hard to get in any standard endogenous growth model.[71] Note that I abstract away from overall structural change in the economy toward services and away from agricultural and manufacturing production, which must explain some of the shift in employment. On the other hand, that abrupt acceleration suggests that rising employment is not exclusively driven by this secular trend.[72] The

establishment-specific wage premiums to be an key driver of inequality. Song et al. (2019) find similar results in the US, while Haltiwanger, Hyatt, and Spletzer (2022) argue that the industry plays the dominant role in the rise in inequality.

[69] Recent work highlights how patents are used to "defensively" to shut down competitors without producing novel content, see Argente et al. (2020). Note some of this literature is motivated by the fact that patenting activity has not translated into productivity growth. An obvious explanation for this weak transmission is provided in the model at hand because patents raise productivity only when the technology is widely adopted.

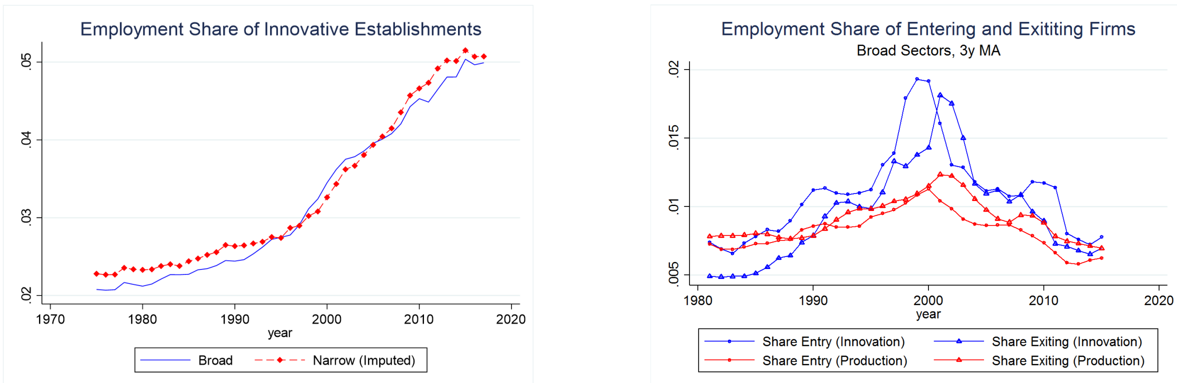
[70] An alternative strategy is to use occupational codes, but they are available for a relatively small sample of workers, and it is unclear, for example, whether an engineer is involved in research or production.

[71] Another test involves comparing the skill share across both sectors, which is indeed diverging since the 1990s and consistent with the theoretical prediction, see figure ?? in the appendix. Note, however, that even though the skill share in the production sector is falling relative to the skill in the research sector, there are increases in the total share of skilled labor in both sectors. A more realistic model would include an additional source of skill-biased task-changing technological change as discussed in Acemoglu and Autor (2011). This force would match the increase in both sectors, while globalization explains the divergence across sectors, which is my focus in this paper.

[72] See Buera et al. (2022) for related work on structural change and the skill premium. Another reason why the model might be off is that much innovation and research is carried out in production establishments in 1990, while stronger sorting (Card, Heining, and Kline, 2013) or outsourcing (Goldschmidt and Schmieder, 2017; Fort et al., 2020) leads to

right panel of figure 8 reports 3year moving average establishment entry rates across both sectors, showing that net entry and overall business dynamics take off in the innovation sector in the mid 1990s, consistent with rising returns to innovation that are arbitraged away in equilibrium by additional entry.

Figure 8. Employment in Innovation



The data is from the IAB BHP establishment panel. I discuss this dataset in the next section. I use sectoral classifications to assign establishments into innovative or productive establishment. Details on the classification are contained in the appendix. And I use information on entry and exit to compute the employment share of entrants, smoothed out using a 3year moving average. An entrant is a firm that did not exist in the previous year. An exiting firm is one that does not exist in the next year. The time series shows that entry and exit dynamics are high during the 90s and 2000s, with net entry into innovation.

Note that these reallocation patterns are fully consistent with a model of neoclassical trade whereby an advanced economy specializes in high-skill service industries and the emerging market specializes in labor intensive sectors after market integration, except for one critical difference: Technology is fixed in models of international trade, which leads to the usual gains from trade. In contrast, the exodus of skilled labor from the production sector has adverse effects on the *level of technology that is used in the production sector* in this model. A process of sectoral "brain drain" sets in that allows for more nuanced effects of openness on growth and inequality.

So far I have shown that a model with an endogenous adoption gap is able to account for the uneven and sluggish growth experience of advanced economies after market integration with emerging markets. Aggregate patterns are consistent with the theory, and hard to reconcile with benchmark growth models that do no feature an adoption margin. In the German context, these patterns are particularly stark as Germany produces frontier technology and integrates quickly with Eastern Europe after the fall of the Iron Curtain. A dramatic rise in exports from Germany from around 20% to 45% from 1995 to 2005, most of which is accounted for by integration between Germany and the "East"[73] as shown in Dauth, Findeisen, and Suedekum (2014), is consistent with and probably understates the extent of market integration that occurred in this short period as German multinationals are heavily invested

more fragmentation between innovation and production. My establishment measure of innovative employment would in that case understate the amount of research done in the early 1990s and thus overstate the growth rate.

[73]This includes former Soviet Satellite States such as Albania, Bulgaria, Croatia, the Czech Republic, Hungary, Poland, Romania, Serbia, Slovakia, Latvia, and Lithuania.

in Eastern Europe and idea flows through this channel do not show up as German exports. [74] I next use this sudden market integration shock in combination with German micro data to offer additional cross-sectional evidence on the rising returns to innovation and weak domestic technology adoption.

5 Cross Sectional Evidence

To obtain cross-sectional predictions I project the two sector structure of the theory into space, and leverage county or local labor markets (3 counties on average) variation in specialization in innovation relative to production. The theory predicts that after market integration and Eastern Europe's growth take off in 1995, regions specialized in innovation experience a positive shock due to the rising global demand for ideas. This should lead to elevated skilled employment growth and GDP growth in these regions. If comparative advantage in innovation is relatively fixed over time, and unevenly distributed across space, this provides an angle to compare how different regions respond to this market integration shock. Both assumptions are consistent with a large literature on the persistent clustering of innovation across space, see Feldman (1994). Moreover, in order to identify any effects labor must be mobile across space. Weak labor mobility in German regions suggest that the exercise is biased against finding any effects. To summarize, regions with greater initial specialization in innovation pre 1995 experience

1. higher (skilled) employment growth and
2. faster GDP growth.

This cross-sectional approach relates to a recent literature in empirical macroeconomics (Nakamura and Steinsson, 2014; Mian and Sufi, 2014; Autor, Dorn, and Hanson, 2013; Chodorow-Reich, 2019). I focus on West Germany to avoid dealing with the massive institutional change in East Germany after German unification around 1990.[75] The timing between inner-German integration in the early 1990s and goods market integration with Eastern Europe after 1994 diverges because the collapse of the Soviet Union had negative effects on many Eastern European economies at first. Most countries were able to recover at around 1994 at which point their growth spurt started. The case of Poland, which joined the WTO in 1995 and the European Union in 2004, summarizes well the overall trend toward integration with the West in Eastern Europe.

[74] In fact, in the model the crucial trade happens between royalties for idea usage against final goods. This assumption serves of course only as a crude stand-in for the many intricate ways in which German entrepreneurial activity creates value in Eastern Europe for their shareholders.

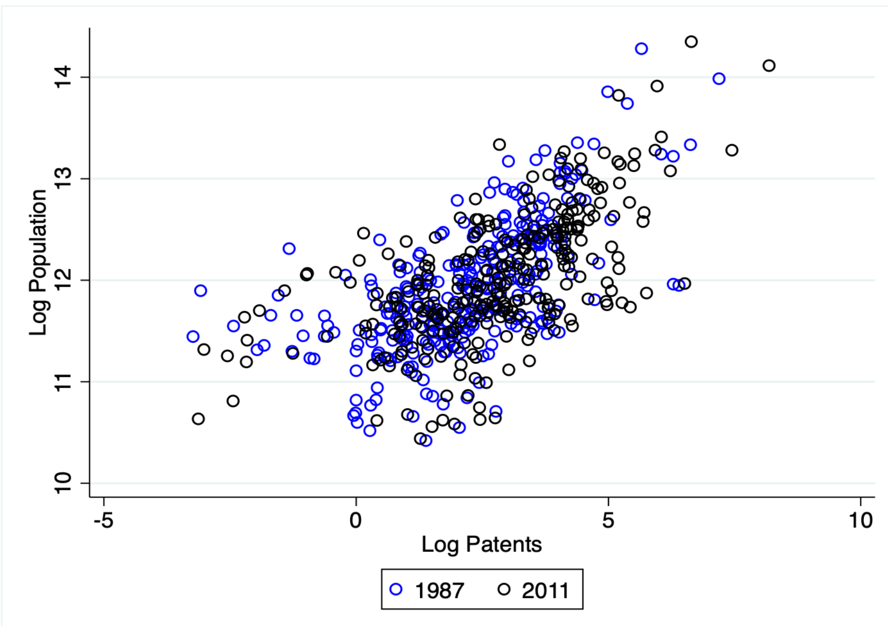
[75] See Findeisen et al. (2021) for work on employment reallocation in the East Germany. Note that most of the convergence within German between East Germany and West Germany occurs up until 1995, see Bachmann et al. (2022). Importantly, while some East German workers did migrate to the West, Findeisen et al. (2021) provide evidence that migration was not a central force after German unification. The fact that goods market integration with Eastern Europe unfolds after 1995, while German integration occurs in the late 1980s and early 1990s, is useful for my identification strategy since changes after 1995 are less likely to be driven by German unification.

5.1 Rising Returns to Innovation

Ideally, I would collect panel data across counties on GDP and skilled employment, neither of which is available consistently over the time horizon in question.[76] Instead, I focus on population growth as a proxy for GDP growth and skilled employment growth for which data is available over the relevant period for the years 1987, 1996, and 2011, assembled by Roesel (2022). I combine this data with patent data from the PATSTAT database (Coffano and Tarasconi (2014)). I measure specialization in innovation by patenting activity, focusing on a 3year moving average of total patents in each county.[77]

Figure 9 plots the positive correlation between log patents and log population across counties (Kreis-level).

Figure 9. Population & Patents Across Counties in West Germany



The figure plots the cross-country correlation between the log of patents and the log of population.

A transparent and simple test is to regress population growth on initial patents in a county, controlling for initial population over the area of a county (density) so as to compare two regions that have the same population-to-space ratio but differ in terms of their specialization in innovation measured by a different number of patents in the base period,

[76]Data from the IAB is in principal available on the county level but the sampling variation is too large to allow for a meaningful regression analysis on that level of granularity. I show results from the IAB sample below that are consistent with the predictions of the model, but are measured on a more aggregate level and in a more descriptive fashion.

[77]Using administrative data from the IAB I confirm below that indeed skilled labor growth and wage growth is biased in favor of high-income innovative regions but the data is not granular enough to be useful in the regression setting here.

$$\Delta_k \log pop_{rt} = \alpha + \gamma_t + \left(\beta + \underbrace{\delta_{t>1995}}_{>0}\right) Patents_{rt} + \log\left(pop_{rt}/area_{rt}\right) + u_{rt}. \tag{59}$$

Controlling for density is essential since there is mean reversion in population growth, and density tends to be tightly correlated with GDP per capita. I report the results in table ??, which confirms that initial patent specialization is a strong predictor of population growth from 1996 onward but not in the first period. Additional information and robustness is contained in the appendix. While using total patens in levels and controlling for log density provides the best fit, I run a version of (59) using log patents, which leads to a semi-elasticity that is easier to interpret: a 14% increase in patents approximately leads to an increase of population growth by 0.1 percentage points.[78]

Table 3. Innovation and Population Growth

	Population Growth
patents (β)	-0.000151
	(-1.56)
(1996-2011) $\times$ patents (δ)	0.000745***
	(5.99)
Time FE	Yes
Pop per Sq KM	Yes
Observations	613
R^2	0.676

Clustered standard errors at county level. T stats in parantheses.
* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

A potential confounder is skill-biased technological change, see the recent wok on urban biased growth and technological change (Giannone, 2017; Eckert, Ganapati, and Walsh, 2020). While I cannot rule out this possibility completely, controlling for density absorbs some of the variation that enters through this mechanism. Moreover, to have a shot at explaining the overall weak growth performance in the 90s and 2000s, one needs an additional channel since skill-biased technological change tends to raise aggregate productivity. The adoption margin is key to resolving this puzzle, and I offer some

[78] An important aspect to the argument is that innovative activity responds to market size. A number of papers has shown this to be the case, see for instance Acemoglu and Linn (2004) or Costinot et al. (2019) or Aghion et al. (2018). In the appendix, see table ??, I regress changes in patents in a 3 digit sector to changes in export flows from 1996 to 2007 using total flows and flows to the East ((CZE, EST, HUN, LTU, LVA, POL, SVK)). The correlation is around .74. To compute this correlation, I use the concordance provided by Lybbert and Zolas (2014) to map each patent's technology class to a sector (Nace 1 Rev & HS2) to match it with trade flows from the BACI database.

evidence on the adoption channel next.

5.2 Missing Technology Adoption

While measurement of technology adoption is challenging,[79] one can look for tell-tale patterns in the data. First, I document changing convergence patterns across high and low income local labor markets in Germany using data from the IAB BHP establishment panel that I can map into local labor markets, which is comparable to a commuting zone in the US.[80] While the sampling variation is too large to to tease out regression-based effects as reported in the previous section, the data is very useful to document broad trend breaks in wage growth and employment growth across local labor markets.

I focus on the period 1985 – 2006 which allows me to consider two separate regimes, one pre and post market integration with Eastern Europe with 1995 marking the dividing line. [81]

Figure 10 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, following Baumol (1986) and Martin and Barro (1997). While wage growth in the early period from 1986 – 1994 was on average higher for laggard regions. These growth patterns reverse completely in the 2000s, leading to relatively faster wage growth for high-income places while laggard regions stagnate.[82] To the extent that laggard regions are more focused on production, and frontier regions host most of the innovation, the changing growth patterns are consistent with rising returns to innovation in the aftermath of market integration. Importantly, faster growth of the frontier could not compensate for weak growth in the hinterlands.

Through the lens of the model, stagnation in laggard regions represents"falling back" due to a rising

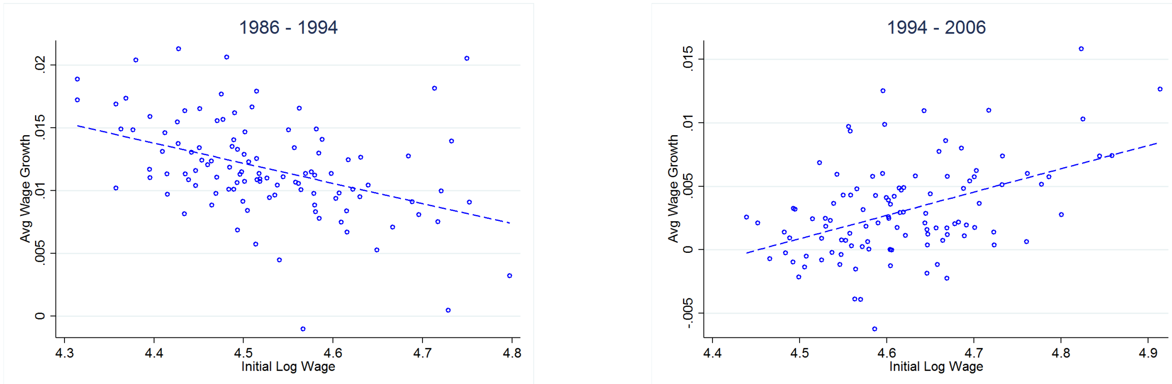
[79]See the approach of Comin and Hobijn (2010b) to measure the use of technologies on the country level, or recent work by Bloom et al. (2021) that use data from company earnings call in combination with machine learning and text analysis tools. Both papers suggest that the degree of technology adoption differs substantially across countries and regions.

[80]The data contains the county in which the establishment is located, as well as sectoral information, and the number and composition of workers, including detailed information on educational attainment. I use Kosfeld and Werner (2012)'s definition of local labor markets which leaves me with 109 regions. A local labor market contains roughly 3 counties on average.

[81]While the data starts in 1975, starting at the beginning is problematic for two reasons. First, the large Oil crisis in the early 1980s constitutes the kind of business cycle variation that I abstract away from in this project. Second, there are structural breaks in the compensation of skilled labor form 1983-1984 that are mostly attributable to measurement issues and not so much to actual wage growth. See for instance Fitzenberger and Kohn (2006) who use a methodology for this structural break from 1983 to 1984. My sample cut avoids this issue altogether. Note that the period from 1975-1985 does not feature strong convergence dynamics in wages, but it does feature strong convergence dynamics in the skill ratio of each region. Overall, there is a clear trend of within-country regional convergence in Europe and the US from 1950 - 1990 as I show in the appendix.

[82]It is likely that this fast growth in high income places is still an understatement due to top-coding issues in the German data. The IAB provides average wages on the establishment level that use the imputation procedure in Card, Heining, and Kline (2013) to deal with the fact that as much as 10% of wage observations are top coded. This procedure relies on a log normal model of the wage distribution which is conservative when taking into account the the income distribution is characterized by a thick right tail. Moreover, a substantial share of income nowadays comes in the form of bonuses, especially for skilled labor, see Eisfeldt, Falato, and Xiaolan (2021). This is entirely missing in the data.

Figure 10. Regional Convergence



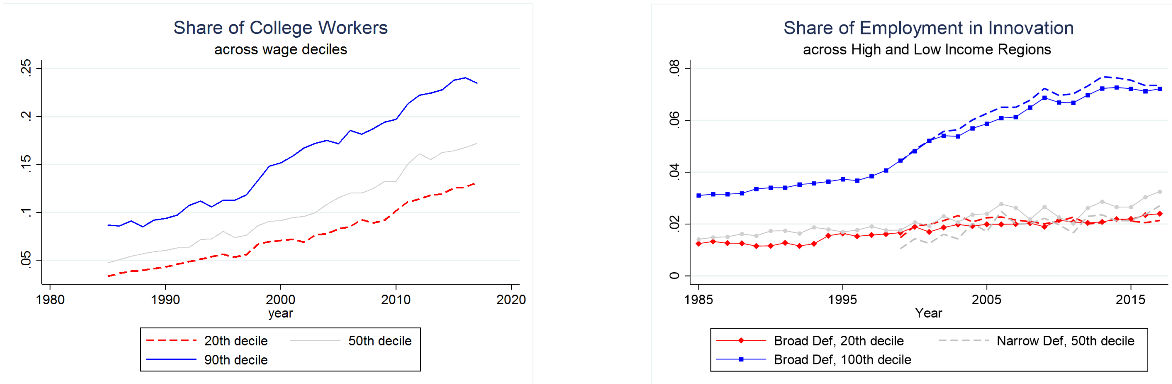
Using data from the BHP establishment sample, the figure plots average wage growth against initial the initial average wage in real terms. The plot shows how growth pre 1994 was biased towards lagging regions, while from 1994 onwards growth was biased towards high income regions. I stop short of the financial crisis, but have looked at convergence patterns from 206 - 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.

price of skill. The adoption gap widens up until the advantage of backwardness is sufficiently strong to make up for less adoption effort. A common concern is that international trade, and in particular import exposure following Autor, Dorn, and Hanson (2013), can lead to 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 in the appendix, which accounts for virtually none of the stagnation in laggard regions, see the appendix H.3 for additional results.

To corroborate the interpretation that a (relative) loss of skill hurts laggard regions, the left panel in figure 11 shows how the share of college workers in high income regions has been diverging in the decade starting in 1995. An acceleration in the college share in the high income regions gave way to stagnation in college share in low income regions, consistent with findings for the US economy (Berry and Glaeser, 2005). The general equilibrium structure of the model makes clear that the acceleration in one place comes at the cost of low-innovation regions that lose skill in relative terms.

Another way to get at the same fact is to correlate wage growth with total skilled employment growth. In the early period, skilled employment growth was fastest in laggard regions. In the later period, the pattern reversed and skilled labor was growing fastest in high income areas, consistent with findings for the US economy, see Berry and Glaeser (2005). Table 4 reports that skilled labor growth is robustly correlated with wage growth in both periods, while the direction of it has reversed completely. This is important because it highlights that explaining regional divergence based on models of skill-biased technological change faces the challenge that i) technological change was also skill-biased in the early post war period, but ii) that bias lead to relatively faster accumulation in laggard regions, not frontier ones. While technological change seems to be generally skill-biased over the 20th century (Goldin and Katz, 2010), adoption-driven growth of the hinterlands gave way to innovation-centric

Figure 11. Share of College Workers and Wage Growth of College Workers



These plots compute skill share and employment in innovation across high and low income regions 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.

urban growth since the 1990s (see Moretti (2019)). Consistent with adoption giving way to frontier innovation is that skilled employment growth used to be correlated with low skilled wage growth. This association disappeared in the more recent period. A model where adoption and innovation compete for skilled labor in a globalized world explains these changing growth patterns across space and workers all at once.

Table 4. Wage Growth & Total High Skill Employment Growth

		$g_H^{1986-1994}$		$g_H^{1994-2005}$		obs
		Coeff.	R^2	Coeff.	R^2	
1.	regional average wage growth	0.1326	**0.3177**	0.1665	**0.3733**	109
2.	regional average wage growth (low skill)	0.1043	**0.1644**	0.0621	**0.0312**	109

The table reports the results from bivariate regressions where wage growth is regressed on skilled employment growth for each period separately across local labor markets in West Germany, using the BHP establishment sample.

6 Conclusion

In this paper, I have developed an endogenous growth model that introduces a novel technology adoption margin. The key assumption in the model is that both innovation and adoption are essential for long-run productivity growth, and both require skilled labor as input. The equilibrium allocation of skilled labor depends on the returns of innovation vis-a-vis technology adoption, with the skill premium being the central price signal in the model.

When advanced economies have a strong comparative advantage in the development of frontier technology, global market integration changes the returns to innovation relative to adoption within rich countries. The innovation sector expands, and the skill premium rises while domestic technology adoption stalls. I calibrate a version of the growth model, which is similar to Jones (1995) but with two types of labor and endogenous technology adoption. In that scenario, weak domestic technology adoption entirely erases gains from additional innovation for the economy as a whole. The model is able to generate sizable real wage losses for production workers in rich countries, which are hurt by weak technology adoption. The theory matches weak aggregate growth in advanced economies while innovation picks up, which eludes benchmark endogenous growth models.

Empirical evidence from Germany where I leverage regional specialization in innovation vs. production together with the fall of the Iron Curtain is consistent with the key mechanism. Importantly, the broad patterns in the data – uneven growth across space and workers where the innovative sector runs away from the rest of the economy – generalize to other economies as well and have been documented in countries like the UK, France, or the USA.

I also highlight that integration will be beneficial to all groups if emerging markets transition fully and begin to contribute to the technological frontier. Like so often in models of endogenous technological change, openness and globalization can play a powerful role in sustaining long-run technological change due to the inherent non-rivalry of knowledge. Recent concerns about the adverse effects of the ability of emerging markets to compete with advanced economies in high-tech are misplaced through the lens of this model. Production workers are the group that would benefit the most from additional foreign innovation. Innovation from the East would push down the domestic skill premium, lead to a reallocation of skilled labor to less innovative regions within advanced economies, and overall faster real wage growth.

This paper takes a first step to analyze the interaction of innovation and adoption and its implication for inequality and productivity growth in a global world. Cross-country and cross-region growth patterns unfold in a highly integrated world economy. Much empirical and quantitative work remains to understand the two-way relationship between local and global growth patterns.

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 and Joshua Linn (2004). “Market size in innovation: theory and evidence from the pharmaceutical industry”. In: *The Quarterly journal of economics* 119.3, pp. 1049–1090.

Acemoglu, Daron and Pascual Restrepo (2018a). “The race between man and machine: Implications of technology for growth, factor shares, and employment”. In: *American Economic Review* 108.6, pp. 1488–1542.

— (2018b). “The race between man and machine: Implications of technology for growth, factor shares, and employment”. In: *American Economic Review* 108.6, pp. 1488–1542.

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

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.

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

Aghion, Philippe et al. (2018). *The impact of exports on innovation: Theory and evidence*. National Bureau of Economic Research New York, NY, USA.

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.

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 E, Francisco J Buera, and Robert E Lucas Jr (2013). *Idea flows, economic growth, and trade*. Tech. rep. National Bureau of Economic Research.

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

Andrews, Dan, Chiara Criscuolo, and Peter N Gal (2016). "The best versus the rest: the global productivity slowdown, divergence across firms and the role of public policy". In.

Angeletos, George-Marios (2007). "Uninsured idiosyncratic investment risk and aggregate saving". In: *Review of Economic dynamics* 10.1, pp. 1–30.

Antras, Pol, Teresa C Fort, and Felix Tintelnot (2017). "The margins of global sourcing: Theory and evidence from us firms". In: *American Economic Review* 107.9, pp. 2514–64.

Argente, David et al. (2020). "Patents to products: Product innovation and firm dynamics". In.

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

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 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.

Bachmann, Rüdiger et al. (2022). "Monopsony Makes Firms not only Small but also Unproductive: Why East Germany has not Converged". In.

Baldwin, Richard (1989). "The growth effects of 1992". In: *Economic policy* 4.9, pp. 247–281.

Baqaee, David and Emmanuel Farhi (2019). *Networks, barriers, and trade.* Tech. rep. National Bureau of Economic Research.

Baqaee, David Rezza and Emmanuel Farhi (2020). "Productivity and misallocation in general equilibrium". In: *The Quarterly Journal of Economics* 135.1, pp. 105–163.

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.

Basu, Susanto and John G Fernald (1997). "Returns to scale in US production: Estimates and implications". In: *Journal of political economy* 105.2, pp. 249–283.

Baumol, William J (1986). "Productivity growth, convergence, and welfare: what the long-run data show". In: *The american economic review*, pp. 1072–1085.

Becker, Sascha O et al. (2019). "Tasks, occupations, and wage inequality in an open economy". In.

Benhabib, Jess, Jesse Perla, and Christopher Tonetti (2014). "Catch-up and fall-back through innovation and imitation". In: *Journal of Economic Growth* 19.1, pp. 1–35.

— (2021). "Reconciling models of diffusion and innovation: A theory of the productivity distribution and technology frontier". In: *Econometrica* 89.5, pp. 2261–2301.

Bergeaud, Antonin and Cyril Verluise (2022). "The Rise of Chinaâs Technological Power: the Perspective from Frontier Technologies". In: Unpublished Manuscript.

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.

Bloom, Nicholas et al. (2021). *The diffusion of disruptive technologies*. Tech. rep. National Bureau of Economic Research.

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.

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 Jonathan Vogel (2017). "International trade, technology, and the skill premium". In: *Journal of Political Economy* 125.5, pp. 1356–1412.

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.

Caliendo, Lorenzo and Fernando Parro (2019). *The quantitative effects of trade policy on industrial and labor location*. Tech. rep. mimeo.

Card, David, Jörg Heining, and Patrick Kline (2013). "Workplace heterogeneity and the rise of West German wage inequality". In: *The Quarterly journal of economics* 128.3, pp. 967–1015.

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

— (2005). "Accounting for cross-country income differences". In: *Handbook of economic growth* 1, pp. 679–741.

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.

Chodorow-Reich, Gabriel (2019). "Geographic cross-sectional fiscal spending multipliers: What have we learned?" In: *American Economic Journal: Economic Policy* 11.2, pp. 1–34.

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 Bart Hobijn (2010a). "An exploration of technology diffusion". In: *American Economic Review* 100.5, pp. 2031–59.

Comin, Diego and Bart Hobijn (2010b). "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.

Costinot, Arnaud and Andrés Rodríguez-Clare (2014). "Trade theory with numbers: Quantifying the consequences of globalization". In: *Handbook of international economics*. Vol. 4. Elsevier, pp. 197–261.

Costinot, Arnaud et al. (2019). "The more we die, the more we sell? a simple test of the home-market effect". In: *The quarterly journal of economics* 134.2, pp. 843–894.

Dauth, Wolfgang, Sebastian Findeisen, and Jens Suedekum (2014). "The rise of the East and the Far East: German labor markets and trade integration". In: *Journal of the European Economic Association* 12.6, pp. 1643–1675.

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.

Decker, Ryan A. et al. (2020). "Changing Business Dynamism and Productivity: Shocks versus Responsiveness". In: *American Economic Review* 110.12, pp. 3952–90.

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

Doepke, Matthias and Ruben Gaetani (2020). *Why didn't the college premium rise everywhere? Employment protection and on-the-job investment in skills*. Tech. rep. National Bureau of Economic Research.

Dustmann, Christian, Johannes Ludsteck, and Uta Schönberg (2009). "Revisiting the German wage structure". In: *The Quarterly journal of economics* 124.2, pp. 843–881.

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

— (2001). "Technology, trade, and growth: A unified framework". In: *European economic review* 45.4-6, pp. 742–755.

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

Eckert, Fabian, Sharat Ganapati, and Conor Walsh (2020). "Skilled Scalable Services: The New Urban Bias in Economic Growth". In.

Eisfeldt, Andrea L, Antonio Falato, and Mindy Z Xiaolan (2021). *Human capitalists*. Tech. rep. National Bureau of Economic Research.

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

Ethier, Wilfred J (1982). "National and international returns to scale in the modern theory of international trade". In: *The American Economic Review* 72.3, pp. 389–405.

Federico, Giovanni and Antonio Tena Junguito (2016). "World trade, 1800-1938: a new data-set". In.

Feenstra, Robert C (2015). *Advanced international trade: theory and evidence*. Princeton university press.

Feenstra, Robert C, Robert Inklaar, and Marcel P Timmer (2016). "What is New in PWT 9.0". In: *Groningen Growth and Development Centre, University of Groningen*.

Feldman, Maryann P (1994). *The geography of innovation*. Vol. 2. Springer Science & Business Media.

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

Findeisen, Sebastian et al. (2021). *Transforming Institutions: Labor Reallocation and Wage Growth in a Reunified Germany*. Tech. rep. Unpublished Manuscript.

Fitzenberger, Bernd and Karsten Kohn (2006). "Skill wage premia, employment, and cohort effects: Are workers in Germany all of the same type?" In.

Fort, Teresa C et al. (2020). "Colocation of Production and Innovation: Evidence from the United States". In.

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

Gerschenkron, Alexander (1962). "Economic backwardness in historical perspective (1962)". In: *Cambridge MA*.

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

Goldberg, Pinelopi Koujianou and Nina Pavcnik (2007). "Distributional effects of globalization in developing countries". In: *Journal of economic Literature* 45.1, pp. 39–82.

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

Goldschmidt, Deborah and Johannes F Schmieder (2017). "The rise of domestic outsourcing and the evolution of the German wage structure". In: *The Quarterly Journal of Economics* 132.3, pp. 1165–1217.

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.

— (2018). "Growth, trade, and inequality". In: *Econometrica* 86.1, pp. 37–83.

Gutiérrez, Germán and Sophie Piton (2020). "Revisiting the global decline of the (non-housing) labor share". In: *American Economic Review: Insights* 2.3, pp. 321–38.

Haltiwanger, John C, Henry R Hyatt, and James Spletzer (2022). *Industries, mega firms, and increasing inequality*. Tech. rep. National Bureau of Economic Research.

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

Helpman, Elhanan, Oleg Itskhoki, and Stephen Redding (2010). "Inequality and unemployment in a global economy". In: *Econometrica* 78.4, pp. 1239–1283.

Hendricks, Lutz and Todd Schoellman (2018). "Human capital and development accounting: New evidence from wage gains at migration". In: *The Quarterly Journal of Economics* 133.2, pp. 665–700.

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.

Hsieh, Chang-Tai, Peter J Klenow, and Ishan B Nath (2019). *A Global View of Creative Destruction.* Tech. rep. National Bureau of Economic Research.

Irwin, Douglas A (2019). "Does trade reform promote economic growth? A review of recent evidence". In.

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

— (2005). "Growth and ideas". In: *Handbook of economic growth.* Vol. 1. Elsevier, pp. 1063–1111.

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

Karabarbounis, Loukas and Brent Neiman (2014). "The global decline of the labor share". In: *The Quarterly journal of economics* 129.1, pp. 61–103.

Karahan, Fatih, Benjamin Pugsley, and Ayşegül Şahin (2019). *Demographic origins of the startup deficit.* Tech. rep. National Bureau of Economic Research.

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.

Kleinman, Benny, Ernest Liu, and Stephen J Redding (2021). *Dynamic spatial general equilibrium.* Tech. rep. National Bureau of Economic Research.

Klenow, Peter J and Andres Rodriguez-Clare (1997). "The neoclassical revival in growth economics: Has it gone too far?" In: *NBER macroeconomics annual* 12, pp. 73–103.

Koh, Dongya, Raül Santaeulàlia-Llopis, and Yu Zheng (2020). "Labor share decline and intellectual property products capital". In: *Econometrica* 88.6, pp. 2609–2628.

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.

Kortum, Samuel S (1997). "Research, patenting, and technological change". In: *Econometrica: Journal of the Econometric Society*, pp. 1389–1419.

Kosfeld, Reinhold and Alexander Werner (2012). "Deutsche arbeitsmarktregionen–neuabgrenzung nach den kreisgebietsreformen 2007–2011". In: *Raumforschung und Raumordnung* 70.1, pp. 49–64.

Krueger, Alan B (1993). "How computers have changed the wage structure: evidence from microdata, 1984–1989". In: *The Quarterly Journal of Economics* 108.1, pp. 33–60.

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

Krugman, Paul R (1979). "Increasing returns, monopolistic competition, and international trade". In: *Journal of international Economics* 9.4, pp. 469–479.

Leamer, Edward E (1994). *Trade, wages and revolving door ideas.*

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

LIS (2022). "Luxembourg Income Study (LIS) Database, http://www.lisdatacenter.org". In.

Liu, Runjuan and Daniel Trefler (2008). *Much ado about nothing: American jobs and the rise of service outsourcing to China and India.* 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 (1988). "On the mechanics of economic development". In: *Journal of monetary economics* 22.1, pp. 3–42.

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

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

Lybbert, Travis J and Nikolas J Zolas (2014). "Getting patents and economic data to speak to each other: An âalgorithmic links with probabilitiesâ approach for joint analyses of patenting and economic activity". In: *Research Policy* 43.3, pp. 530–542.

Malmberg, Hannes (2017). "Human capital and development accounting revisited". In: *mimeo Institute for International Economic Studies.*

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.

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

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

Melitz, Marc J and Gianmarco IP Ottaviano (2008). "Market size, trade, and productivity". In: *The review of economic studies* 75.1, pp. 295–316.

Mian, Atif and Amir Sufi (2014). "What explains the 2007–2009 drop in employment?" In: *Econometrica* 82.6, pp. 2197–2223.

Moll, Benjamin (2014). "Productivity Losses from Financial Frictions: Can Self-Financing Undo Capital Misallocation?" In: *American Economic Review* 104.10, pp. 3186–3221. DOI: 10.1257/aer.104.10.3186. URL: https://www.aeaweb.org/articles?id=10.1257/aer.104.10.3186.

Moretti, Enrico (2012). *The new geography of jobs*. Houghton Mifflin Harcourt.

— (2019). *The effect of high-tech clusters on the productivity of top inventors*. Tech. rep. National Bureau of Economic Research.

Nakamura, Emi and Jon Steinsson (2014). "Fiscal stimulus in a monetary union: Evidence from US regions". In: *American Economic Review* 104.3, pp. 753–92.

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.

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)*.

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.

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.

Roesel, Felix (2022). "German Local Population Database (GPOP)(Version 1.0)". In.

Romer, David (2012). *Advanced macroeconomics, 4e*. New York: McGraw-Hill.

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

Rosés, Joan Ramón and Nikolaus Wolf (2018). *The economic development of Europe's regions: A quantitative history since 1900*. Routledge.

Rossi, Federico (2022). "The relative efficiency of skilled labor across countries: Measurement and interpretation". In: *American Economic Review* 112.1, pp. 235–66.

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

Sampson, Thomas (2014). "Selection into trade and wage inequality". In: *American Economic Journal: Microeconomics* 6.3, pp. 157–202.

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

— (2019). "Technology gaps, trade and income". In.

Schoellman, Todd (2012). "Education quality and development accounting". In: *The Review of Economic Studies* 79.1, pp. 388–417.

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.

Song, Jae et al. (2019). "Firming up inequality". In: *The Quarterly journal of economics* 134.1, pp. 1–50.

Studwell, Joe (2013). *How Asia works: Success and failure in the world's most dynamic region*. Open Road+ Grove/Atlantic.

Van Ark, Bart, Mary O'Mahoney, and Marcel P Timmer (2008). "The productivity gap between Europe and the United States: trends and causes". In: *Journal of economic perspectives* 22.1, pp. 25–44.

Wood, Adrian (1994). *North-South trade, employment and inequality: changing fortunes in a skill-driven world*. Clarendon Press.

Young, Alwyn (1998). "Growth without scale effects". In: *Journal of political economy* 106.1, pp. 41–63.

Appendix is preliminary and incomplete

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

$$\begin{aligned} \max \quad & \int_0^A \left(\frac{x(z)}{\alpha}\right)^{\alpha} dz \\ s.t. \quad & \int p_j x_j dj \leq I. \end{aligned} \tag{60}$$

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 $x_j = x_z \, \forall j, z$. As a consequence, the total quantity of each individual capital good variety must read $p_j x_j = \frac{p_x x}{A}$ 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

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

The problem further simplifies to

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

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

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

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

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

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

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

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

To solve the production firm's adoption problem, it is useful to rewrite the problem using a normalized value function $v=\frac{V}{w_t}$, as well as normalizing the state variable A^K 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 $r=\rho+g_F$, this leads to the following recursive formulation of the firm adoption problem,

$$\begin{aligned} v\left(\rho+\delta_{ex}\right) &= \max \dot{v}+\frac{\pi_t(z)}{w}-s_t h+v_z\dot{z} \\ s.t. & \\ \dot{z} &= z^{\theta}\zeta h^{\beta}-\left(g_F+\delta_I\right)z. \end{aligned} \tag{61}$$

A solution to the program (61) needs to satisfy the following first order condition

$$v_z\beta z^{\theta}\zeta h^{\beta-1} = s . \tag{62}$$

Equation (62) 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

$$\begin{aligned}\frac{v_{zz}}{v_z} + \frac{\theta}{z} &= (1-\beta)\,\frac{h_z}{h} \\ \frac{v_{tz}}{v_z} - \frac{\dot{s}}{s} &= (1-\beta)\,\frac{h_t}{h}.\end{aligned}$$

Next, use the Euler equation and the envelope condition after differentiating the HJB equation to get

$$\begin{aligned}
v_z\left(\rho+\delta_{ex}\right) &= \frac{\pi_z}{w} + v_{zz}\dot{z} + v_z\left(\theta z^{\theta-1}\zeta h^{\beta} - \left(\delta_I + g_F\right)\right) + \dot{v_z} \\
v_z\left(\rho+\delta_{ex}\right) &= \frac{\pi_z}{w} + v_{zz}\dot{z} + v_z\left(\theta z^{\theta-1}\zeta h^{\beta} - \theta\left(\delta_I + g_F\right) - (1-\theta)\left(\delta_I + g_F\right)\right) + v_z \\
v_z\left(\rho+\delta_{ex}+(1-\theta)\left(\delta_I+g_F\right)\right) &= \frac{\pi_z}{w} + v_{zz}\dot{z} + v_z\frac{\theta}{z}\left(z^{\theta}\zeta h^{\beta} - z\left(\delta_I + g_F\right)\right) + v_z\left\{(1-\beta)\,\right. \\
v_z\left(\rho+\delta_{ex}+(1-\theta)\left(\delta_I+g_F\right)\right) &= \frac{\pi_z}{w} + v_z\left(\frac{v_{zz}}{v_z}+\frac{\theta}{z}\right)\dot{z} + v_z\left\{(1-\beta)\,\frac{h_t}{h}+\frac{\dot{s}}{s}\right\} \\
\left(\rho+\delta_{ex}+(1-\theta)\left(\delta_I+g_F\right)\right) &= \frac{\pi_z}{w}\frac{1}{v_z} + \left(\frac{v_{zz}}{v_z}+\frac{\theta}{z}\right)\dot{z} + (1-\beta)\,\frac{h_t}{h}+\frac{\dot{s}}{s} \\
\left(\rho+\delta_{ex}+(1-\theta)\left(\delta_I+g_F\right)\right) &= \frac{\pi_z}{w}\frac{1}{v_z} + (1-\beta)\,\frac{\dot{h}}{h} + \\
\frac{1}{1-\beta}\left\{\left(\rho+\delta_{ex}+(1-\theta)\left(\delta_I+g_F\right)\right) - \left[\frac{\dot{s}}{s}+\frac{\pi_z}{w}\frac{1}{v_z}\right]\right\} &= \frac{\dot{h}}{h}
\end{aligned}$$

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

$$\begin{aligned}
\frac{\dot{h}}{h} &= \frac{1}{1-\beta}\left\{\left(\rho+\delta_{ex}+(1-\theta)\left(\delta_I+g_F\right)\right) - \frac{1}{v_z}\left[\frac{\pi}{w}\frac{(1-\alpha)(\sigma-1)}{z}\right] - \frac{\dot{s}}{s}\right\} \\
\frac{\dot{h}}{h} &= \frac{1}{1-\beta}\left\{\left(\rho+\delta_{ex}+(1-\theta)\left(\delta_I+g_F\right)\right) - \frac{\beta z^{\theta}\zeta h^{\beta-1}}{s}\left[\frac{\pi}{w}\frac{(1-\alpha)(\sigma-1)}{z}\right] - \frac{\dot{s}}{s}\right\}
\end{aligned}$$

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

$$\frac{\dot{z}}{z} = z^{\theta-1}\zeta h^{\beta} - \left(\delta_I + g_F\right).$$

In the steady state, we have that

$$h^{1-\beta} = \frac{1}{s}\frac{\beta(1-\alpha)(\sigma-1)(g_F+\delta_I)}{\rho+\delta_{ex}+(1-\theta)(\delta_I+g_F)}\left[\frac{\pi}{w}\right]\frac{z^{\theta-1}\zeta}{(g_F+\delta_I)} \tag{63}$$

$$z^{1-\theta} = \frac{\zeta h^{\beta}}{g_F+\delta_I} \tag{64}$$

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

$$hs = \frac{\beta(1-\alpha)(\sigma-1)(g_F+\delta_I)}{\rho+\delta_{ex}+(1-\theta)(\delta_I+g_F)}\left[\frac{\pi}{w}\right]\ .$$

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

$$\beta(1-\alpha)(\sigma-1) < \frac{\rho+\delta_{ex}}{g_F+\delta_I} + (1-\theta)\,.$$

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 ($(\sigma-1)(1-\alpha)$). The right hand side consist of the cost 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 frontier, 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.

Q-Theory:

Next I derive the same dynamics in the perhaps using the current value Hamiltonian and the familiar q-theory of investment approach, see for instance the textbook of Romer (2012). Instead of using the HJB, I can define the current value Hamiltonian,

$$H = \frac{\pi}{w} - sh + q_t\left[z^{\theta}\zeta h^{\beta} - (\delta_I + g_F) z\right]$$

The optimality conditions are standard and read

$$\begin{aligned} H_h &= 0 \\ &\Leftrightarrow \\ \beta q_t z^{\theta}\zeta h^{\beta-1} &= s \end{aligned}$$

and

$$\begin{aligned} H_z &= -\dot{q}_t + (\rho + \delta_{ex}) q_t \\ &\Leftrightarrow \\ \frac{\pi_z}{w} + q_t\left\{\theta\left(z^{\theta-1}\zeta h^{\beta} - (\delta_I + g_F)\right) - \left[(1-\theta)(\delta_I + g_F) + (\rho + \delta_{ex})\right]\right\} &= -\dot{q}_t \\ \frac{\pi_z}{w} + q_t\left\{\theta\left(\frac{\dot{z}}{z}\right) - \left[(1-\theta)(\delta_I + g_F) + (\rho + \delta_{ex})\right]\right\} &= -\dot{q}_t \end{aligned}$$

I can rewrite the previous equation, using $\exp(\theta\log z_t - \tilde{r}t)$ as integrating factor so that

$$\begin{aligned} -\exp(\theta\log z_t - \tilde{r}t)\frac{\pi_z}{w} &= \exp(\theta\log z_t - \tilde{r}t)\left\{\dot{q}_t + q_t\left\{\theta\left(\frac{\dot{z}}{z}\right) - \left[(1-\theta)(\delta_I + g_F) + (\rho + \delta_{ex})\right]\right\}\right\} \\ -\exp(\theta\log z_t - \tilde{r}t)\frac{\pi_z}{w} &= \frac{\partial q_t \exp(\theta\log z_t - \tilde{r}t)}{\partial dt} \end{aligned}$$

Now I can integrate this expression forward so that q_t indeed captures the marginal value of an extra unit of technology z where a transversality condition needs to hold to ensure that the expression is finite. This leads to

$$\begin{aligned} \underbrace{q_{\infty}\exp(\theta\log z_{\infty} - \tilde{r}\infty)}_{=0} - q_t\exp(\theta\log z_0) &= -\int_t^{\infty}\exp(\theta\log z_x - \tilde{r}x)\frac{\pi_z(x)}{w(x)}dx \\ &\Leftrightarrow \\ q_t &= \int_t^{\infty}\exp\left(\theta\log\left(\frac{z_x}{z_t}\right) - \tilde{r}x\right)\frac{\pi_z(x)}{w(x)}dx \\ q_t &= \int_t^{\infty}\exp(-(\rho+\delta_{ex})x)\exp\left(-\left[\theta\log\left(\frac{z_t}{z_x}\right) + (1-\theta)(g_F + \delta_I)x\right]\right)\frac{\pi_z(x)}{w(x)}dx \end{aligned}$$

In standard q-theory applications, θ equals one and π_z is falling in z.[83] The advantage of backwardness embodied in $1-\theta$ shows up in the firm problem and looks like an additional discount factor. The reader might note the strong resemblance to the neoclassical growth model where $\alpha k^{\alpha-1}=\rho+g+\delta$. Faster growth requires a higher return to capital, but the effect of this is attenuated for large θ as the diminishing returns in the accumulation of knowledge stock A disappear.

Another intuitive implication of the theory is that

$$q_t=\int_t^{\infty}\exp\left(-\left(\rho+\delta_{ex}\right)x\right)\exp\left(-\left[\left(1-\theta\right)\left(g_F+\delta_I\right)x\right]\right)\left(\frac{z_x}{z_t}\right)^{\theta}\frac{\pi_z\left(x\right)}{w\left(x\right)}dx$$

is relatively high when $z_t<z_x$. That is, when the current level of technology is low relative to the long-run steady state, the marginal product of an extra unit of technology is high. The extent to which this is the case is governed by θ, and the effect would disappear as θ approaches zero. One can then infer that a large θ will be helpful to produce long-lasting convergence dynamics. The previous equation also highlights that after a shock q converges back to its long-run value as long as $\frac{\pi_z(x)}{w(x)}$ is unchanged.

In the main text I consider a 10% percent increase in s and its effect on h. Here, it is clear that

$$\left\{\frac{\beta q_t z^{\theta}\zeta}{s}\right\}^{\frac{1}{1-\beta}} = h$$

holds and the increase in the price of skill must lead to an immediate jump down for both h and q. Over time, q recovers, and so does h but it will settle on a permanently lower level.

Just like in the standard q-theory problem, adjustment frictions disappear as

A.1.3 Dynamic Firm Problem and adoption Gap

It is worthwhile to clarify the relationship between general equilibrium and partial equilibrium. When deriving the aggregate dynamics of the economy, I obtain a well-behaved q-theory of investment in skilled labor. A crucial step in the derivation is to impose that the individual firm's productivity z_i is equal to average productivity z_{agg}. I know this must be true due to the homogeneous firm assumption, and it shows up in the first order condition of the firm as follows

$$\begin{aligned}
\frac{\partial\pi(z_i,z_{agg})}{\partial z_i} &= B_t\frac{\partial\left(\frac{z_{i,}}{z_{agg}}\right)^{(\sigma-1)(1-\alpha)}}{\partial z_i}\\
&= B_t\left(\sigma-1\right)\left(1-\alpha\right)\frac{\left(\frac{z_{i,}}{z_{agg}}\right)^{(\sigma-1)(1-\alpha)}}{z_i}\\
&= B_t\frac{(\sigma-1)(1-\alpha)}{z_i}\\
&= \frac{(\sigma-1)(1-\alpha)}{z_i}\pi
\end{aligned}$$

[83] Usually, z would be capital k and so the marginal profits would be equal to the marginal product of capital.

as used in the main text. Note that this general equilibrium effect is crucial to generate q-like dynamics as it leads to diminishing returns in z. This is only true because $\frac{z_i}{z_{agg}}$ cancels so that the convexity captured in $z^{(\sigma-1)(1-\alpha)}$ does not show up. This is absolutely crucial for the existence and uniqueness results in dynamic optimization, see Acemoglu (2009) chapter 7, and the inequality offered in proposition 1 that guarantees a solution is based on this general equilibrium structure whereby z_i and z_{agg} cancel in ways that ensure that profits hardly move along the transition path.

TO DO: to what extent do profits move? unclear, because we need to figure out what happens with entry, something must give, but note that in long and short run you end up on the

Phase Diagram:

Next, I show that a unique saddle-path stable equilibrium obtains where I keep s fixed. To show that first note that (63) establishes a negative link between z and h, while (64) establishes a positive link, implying a unique intersection given regularity conditions

$$\begin{aligned} \tfrac{d}{dz}\left(h_{ss}\right) &< 0 \\ \tfrac{d}{dh}\left(z_{ss}\right) &> 0. \end{aligned}$$

Next, I show the derivative of the differential equations

$$\begin{aligned} \frac{d}{dz}\frac{\dot{h}}{h}\left(1-\beta\right) &= -\frac{d}{dz}\frac{\beta z^{\theta}\zeta h^{\beta-1}}{s}\left[\frac{\pi}{w}\frac{\left(1-\alpha\right)\left(\sigma-1\right)}{z}\right] \\ &= \left(1-\theta\right)\frac{\beta z^{\theta-2}\zeta h^{\beta-1}}{s}\left[\frac{\pi}{w}\left(1-\alpha\right)\left(\sigma-1\right)\right] > 0 \end{aligned}$$

Second, consider the effect of an increase in h on z,

$$\tfrac{d}{dh}\tfrac{\dot{z}}{z} = \beta z^{\theta-1}\zeta h^{\beta} > 0$$

A.2 Entry with Partial Knowledge Spillovers

One strong assumption in the paper is the complete knowledge spillover from incumbents to entrants at entry. That is, after paying a fixed cost $f_e w$, the entrant is able to use the current level of know-how A^K. From then one, the entrant, like any other incumbent, hires skilled labor to adoption new frontier technology.

An alternative specification is one where the entrant only obtains a fraction $\iota A^K_{\max}$ where $\iota \in (0,1)$ and $A^K = \sup\left\{A^K_i : i \in \Omega_M\right\}$. This tweak turns the setting into a heterogeneous firm model similar to Luttmer (2007) and I assumed that the entrant learns from the most sophisticated entrants, but imperfectly so, hence $\iota < 1$.

A well-defined equilibrium is going to be characterized by a distribution $f\left(z\right)$ where the support is $z \in \left[\iota z_{\max}, z_{\max}\right]$. This leads to a normalized free entry condition of the type

$$f_e = v\left(\iota z_{\max}\right).$$

Note that the profit ratio of any two firms can be expressed as $\frac{\pi_i}{\pi_j} = \left(\frac{z_i}{z_j}\right)^{(1-\alpha)(\sigma-1)}$. Moreover, normalized profits for some firm i are given by $\frac{\pi(z_i)}{w} = \frac{(z_i)^{(1-\alpha)(\sigma-1)}}{\mathbb{E}\left[z^{(1-\alpha)(\sigma-1)}\right]}\frac{l^p}{m(\sigma-1)(1-\alpha)}$, see Melitz (2003) for the benchmark heterogeneous firm model.

Now, consider the problem of some firm i using the HJB approach in the steady state (so that $v_t = 0$)

$$\begin{aligned}\left(\rho+\delta_{ex}\right)v\left(z_i\right) &= \pi\left(z_i\right) - sh_i + \partial_{z_i}v * \left[z_i^{\theta}\zeta h_i^{\beta} - \left(\delta_I + g_F\right)z_i\right] \\ h_i &= \left\{\frac{\left(\partial_{z_i}v\right)z_i^{\theta}\beta\zeta}{s}\right\}^{\frac{1}{1-\beta}}\end{aligned}$$

The same argument as before leads to a dynamic investment equation of the form

$$\frac{\dot{h}_i}{h_i} = \frac{1}{1-\beta}\left\{\left(\rho+\delta_{ex}+(1-\theta)\left(\delta_I+g_F\right)\right) - \frac{\beta z_i^{\theta}\zeta h_i^{\beta-1}}{s}\left[\frac{\pi}{w}\frac{(1-\alpha)(\sigma-1)}{z_i}\right]\right\}$$

$$\frac{\dot{h}_i}{h_i} = \frac{1}{1-\beta}\left\{\left(\rho+\delta_{ex}+(1-\theta)\left(\delta_I+g_F\right)\right) - \frac{\beta z_i^{\theta}\zeta h_i^{\beta-1}}{s}\left[\frac{z_i^{(1-\alpha)(\sigma-1)}}{m(1-\alpha)(\sigma-1)\,\mathbb{E}\left[z^{(1-\alpha)(\sigma-1)}\right]}\frac{(1-\alpha)(\sigma-1)}{z_i}\right]\right\}$$

$$\begin{aligned}\frac{\dot{h}_i}{h_i}(1-\beta) &= \left(\rho+\delta_{ex}+(1-\theta)\left(\delta_I+g_F\right)\right) \\ &\quad - \frac{\beta z_i^{\theta-1}\zeta h_i^{\beta-1}}{s}\left(\frac{z_i}{z_{\max}}\right)^{(1-\alpha)(\sigma-1)}(1-\alpha)(\sigma-1)\underbrace{\left[\frac{1}{m(1-\alpha)(\sigma-1)\,\mathbb{E}\left[\left(\frac{z}{z_{\max}}\right)^{(1-\alpha)(\sigma-1)}\right]}\right]}_{=:\Lambda_{z_{\max}}}\end{aligned}$$

$$\frac{\dot{h}_i}{h_i} = \frac{1}{1-\beta}\left\{\left(\rho+\delta_{ex}+(1-\theta)\left(\delta_I+g_F\right)\right) - \frac{(1-\alpha)(\sigma-1)\beta z_i^{\theta-1}\zeta h_i^{\beta-1}}{s}\left(\frac{z_i}{z_{\max}}\right)^{(1-\alpha)(\sigma-1)}\Lambda_{z_{\max}}\right\}$$

In the steady state $z_i = z_{\max}$. So then not much changes except for an additional constant $\Lambda_{z_{\max}} > 1$, which is greater one since entering firms make below-average profits until they converge. In a free entry equilibrium, this is compensated for by higher profits on average.

Now if I can show that this dynamic investment condition delivers well-behaved transition dynamics and leads to convergence toward $z_{\max}$, then indeed this equilibrium is very similar to the one considered in the main part of the paper with complete knowledge spillover.

First, I show that the second order derivative is positive so that $\partial_{z_i}v\left(z_i\right)$ is strictly increasing in

z_i. Using the q-theory derivation, one can rewrite

$$\partial_{z_i} v = \int_t^{\infty} \partial_z \pi\left(z_i; x\right) \exp\left(\theta \log \frac{z_x}{z_t} - \left\{\left[1-\theta\right]\left(g_F + \delta_I\right) + \rho + \delta_{ex}\right\}\right) dx$$
$$= \left(\partial_z \pi\left(z_{\max}\right)\right) * \int_t^{\infty} \left(\frac{z\left(x\right)}{z_{\max}}\right)^{(1-\alpha)(\sigma-1)-1} \exp\left(\theta \log\left(\frac{z_x}{z_t}\right) - \left\{\left[1-\theta\right]\left(g_F + \delta_I\right) + \rho + \delta_{ex}\right\} x\right) dx.$$

In the steady state $z_{\max}$ is fixed, and the partial derivative of the firm with the maximum productivity level is fixed as well. It is easy to see that the right hand side is increasing in z_t, assuming that $z\left(x\right) \geq z_t$ and $\left(\sigma - 1\right)\left(1 - \alpha\right) \geq 1$. This means that the marginal benefit of investing in adoption is increasing for an entrant, reflecting the knowledge spillover in θ as well as the fact that relative profits are rising as the firm catches up with the local state of the art. State of the art is not to be confused with the frontier. Since $\partial_{z_i} v$ is rising in z_i this means that the equilibrium adoption effort is increasing over time.

This is very different from the q-like investment dynamics of the homogenous firm model that is explained in the previous section. The fundamental difference here is that the aggregate productivity level does not change while the entrant is converging to the frontier. In contrast, in the homogenous firm model, productivity improvements are undone by other firms making the exact same choices so profits do not move and are always proportional to $\frac{Y}{M}$, precisely because changes in z_i and z_{agg} cancel. This is not the case here because indeed $z_i \neq z_{agg}$. This also means that the problem is not concave and might raise concerns about whether this firm will incur ever increasing levels of investment. This is not the case because eventually the firm reaches the level $z_{\max}$ at which the dynamics change, precisely because now the aggregate price index would move against it if it were to continue adoption more technology. The convergence to the steady state is thus one up to level $z_{\max}$ and the problem is not well defined for $z_i > z_{\max}$.

Now I perform a similar phase diagram analysis as before, but note the very different dynamics: both the resource constraint ($\dot{z} = 0$) and the dynamics investment equation ($\dot{h} = 0$) in the steady state are upward sloping since $z_i^{\theta - 1 + (1-\alpha)(\sigma-1)}$ is raised by a positive component, again because $\mathbb{E}\left[z^{(1-\alpha)(\sigma-1)}\right]$ is fixed relative to the homogenous firm case where the average productivity moves precisely the same as aggregate productivity. Deriving the link between z_i and h_i in the steady state now means that a new loci emerges that starts at zero, and is upward sloping so as to hit the steady state equilibrium from below when $z_i = z_{\max}$. Note that $\frac{d\dot{h}}{dz_i} < 0$ while $\frac{d\dot{z}_i}{dh_i} > 0$ so that the only solution that converges to the steady state from below is one where h_i jumps up at entry, and then keeps increasing incrementally up until $z_{\max}$ is reached. The speed of this convergence is going to pin down the distribution $f\left(z\right)$.

I don't solve for this stationary distribution. This is of interest in its own, and something that I focus on in follow up work. Here, I just want to highlight that a model with this layer of heterogeneity looks identical to a homogeneous firm model in the steady state where

$$f_e = v(\lambda z_{\max})$$

so that the problem of the cohort of firms that has adopted the most amount of technology reads

$$\frac{\dot{h}_i}{h_i} = \frac{1}{1-\beta}\left\{(\rho + \delta_{ex} + (1-\theta)(\delta_I + g_F)) - \frac{\beta(1-\alpha)(\sigma-1)\, z_{\max}^{\theta-1}\zeta h_i^{\beta-1}}{s}\pi(z_{\max})\right\}.$$

Adoption effort by this firm has the same relationship to the relative price of skill. Since the distribution is a scaled down version of the adoption of the firms that has the maximum level of adoption, the same partial elasticity of z with respect to s obtains. To see this, note that a higher price of skill leaves the relative problem of each firm unchanged ($\frac{z_i}{z_{\max}}$the same), but pushes overall demand down ($z_{\max}$ down as well). How so? Recall

$$\left\{\frac{\beta q_i z_i^{\theta}\zeta}{s}\right\}^{\frac{1}{1-\beta}} = h_i$$
$$\left\{\frac{\beta q z_{\max}^{\theta}\left(\frac{z}{z_{\max}}\right)^{\theta}\zeta}{s}\right\}^{\frac{1}{1-\beta}} = h_{\max}\frac{h}{h_{\max}}$$

Note that

Aggregate demand for skilled labor for adoption purposes in the steady state can be derived as

$$h^D = m\int_{\lambda z_{\max}}^{z_{\max}} f(z)\, h(z)\, dz$$

where $\lambda \to 1$ brings back the benchmark case in the main part of the text with a Dirac measure at 1.

An implication of this heterogeneous firm model is that profits must be higher for the most productive firms. To see this, note that the partial knowledge spillover effectively makes entry more expensive since initial profits are very low. To be compensated for this, average profits in the long run have to be higher. This adjustment occurs through endogenous firm entry which leads to an overall smaller measure of firms in equilibrium. Of course, this matters for the overall demand of skilled labor in production. The point is, however, that one can effectively simulate this feature by simply raising the overall fixed cost.

This is the sense in which the simplification in the main part is very useful without leading to fundamentally difference economics in a model of heterogeneous firms. The main benefit of this simplification is twofold. First, it allows me to focus on cross country inequality. This is the main point, and the knowledge spillover is to be understood as one that plays out on the country level for medium run dynamics. Second, it drastically simplifies the problem of an innovator. Note that an innovator needs to anticipate exactly when its ideas are adopted. In the heterogenous firm model, this means integrating over a distribution. This is quite feasible in the steady state for a stationary distribution. Off the steady state, transition dynamics are already very complex already due to endogenous entry, multiple state variables, and an endogenous waiting time τ. Abstracting away form additional firm heterogeneity is thus very helpful.

that a higher fixed cost leads to

this

While solving for these transition dynamics requires simulating the transition of an individual firm that just entered, and then solving for a stationary distribution, aggregating this up over different cohorts so as to generate a stationary long-run distribution, it is important to note that this is irrelevant for the qualitative behavior of the learning dynamics of the

MORE STUFF BELOW<NEEDS MORE WORK AND THE PLOT TOO

For a unique equilibrium to exist one needs the loci to intersect exactly once.

One final thing to show is that a fresh entrant optimally picks $\dot{z} > 0$. Consider the problem of an entrant and note

$$\begin{aligned}
\partial_{z_i} v &= \int_t^{\infty} \pi_{z_i}(x) \exp\left(\theta \log \frac{z_x}{z_t} - \{[1-\theta](g_F + \delta_I) + \rho + \delta_{ex}\}\right). \\
&= (\partial_z \pi(z_{\max})) * \int_t^{\infty} (\lambda)^{(1-\alpha)(\sigma-1)-1} \exp\left(\theta \log\left(\frac{z_x}{\lambda z_{\max}}\right) - \{[1-\theta](g_F + \delta_I) + \rho + \delta_{ex}\}\right)
\end{aligned}$$

What needs to be shown is that the newest entrants have an incentive to adopt technology, i.e. $\frac{(1-\alpha)(\sigma-1)\beta(\lambda z_{\max})_i^{\theta-1}\zeta h_i^{\beta-1}}{s}(\lambda)^{(1-\alpha)(\sigma-1)}\Lambda_{z_{\max}} > \rho + \delta_{ex} + (1-\theta)(\delta_I + g_F)$

but h is endogenous so what do you do?

consider an h_i where the firm stays put.

$$\begin{aligned}
\zeta h_i^{\beta} &= (g_F + \delta_I)(\lambda z_{\max})^{1-\theta} \\
h_i &= \left(\frac{g_F + \delta_I}{\zeta}\right)^{\frac{1}{\beta}} (\lambda z_{\max})^{\frac{1-\theta}{\beta}}
\end{aligned}$$

now plug that into differential equation

$$(\rho + \delta_{ex} + (1-\theta)(\delta_I + g_F)) - \frac{(1-\alpha)(\sigma-1)\beta(\lambda z_{\max})^{-\left(\frac{1-\theta}{\beta}\right)}\zeta\left(\left(\frac{g_F+\delta_I}{\zeta}\right)^{\frac{1}{\beta}}\right)^{\beta-1}}{s}(\lambda)^{(1-\alpha)(\sigma-1)}\Lambda_{z_{\max}}$$

In fact, this version of the model leads to more realistic convergence dynamics that avoid the initial "jump" that happens in the q-type dynamics derived before. The reason is that

Off the steady state, consider the initial thingy with the worst productivity. Need to show its only going up.

what about smooth pasting and in particular

$$
\begin{aligned}
v\left(z_{\max}-\epsilon\right) &\approx v\left(z_{\max}\right) \\
v'\left(z_{\max}-\epsilon\right) &\approx v'\left(z_{\max}\right)
\end{aligned}
$$

even though strangely the frontier has very different dynamics. This is worth exploring more but I won't do so in this paper.

A.3 Innovation problem and market clearing conditions for high skilled labor

A.3.1 Innovation

To solve for the demand of human capital in the innovation sector I first need to compute the present discounted value of an innovation. Computing the integral over all instantaneous profits π^I (19) in the future by taking account of the waiting time δ leads to

$$
V_{I,t} = \int_{t+\tau}^{\infty} \exp\left(-\left(r+\delta_I\right)(u-t)\right) L_u^P w_u \alpha \left(\tfrac{1}{A_u}\right) du \quad .
$$

Note that both the wage rate and production labor L^P grow at a constant rate, and so does the overall level of technology A, which allows me to solve the integral

$$
\begin{aligned}
&= L_t^P w_{L,t} \alpha \left(\tfrac{1}{A_t}\right) \int_{t+\tau}^{\infty} \exp\left(-\left(r-g_w-g_L+g_A+\delta_I\right)(u-t)\right) du \\
&= \left(\tfrac{L_t^P w_{L,t}\alpha}{A_{F,t} z}\right) \left(\tfrac{1}{r-g_w-g_L+g_A+\delta_I}\right) \exp\left(\left(\tfrac{\rho+g_A-g_L+\delta_I}{\delta_I+g_A}\right) \log z\right) \\
&= \left(\tfrac{L_t^P w_{L,t}\alpha}{A_{F,t}}\right) \left(\tfrac{1}{\rho-g_L+g_A+\delta_I}\right) z^{\frac{\rho-g_L}{g_A+\delta_I}} ,
\end{aligned}
$$

where the second line follows by using $\tau = -\frac{\log z}{g_A+\delta_I}$ in the steady state.

A.3.2 Innovation on and off the balanced growth path

Note that $V^I = \int_{t+\tau}^{\infty} \exp\left(-\int_t^u \left(r+\delta_I\right) dx\right) \pi_u du$, and differentiating this expression leads to the HJB representation

$$
\left(r+\delta_I\right) V_t^I - \dot{V}_t^I = \exp\left(-\int_t^{t+\tau} \left(r+\delta_I\right) dx\right) \frac{\alpha L_{t+\tau}^P w_{t+\tau}}{A_{F,t+\tau} z_{t+\tau}} \left[1+\tau'\right] . \tag{65}
$$

Note that as long as the free entry condition is binding, it must be that the time derivative of the value function is consistent with rising entry cost, i.e.

$$\begin{aligned} f_R w_H A_F^{-\phi} &= V^I \\ &\Leftrightarrow \\ g_{H_F}(1-\lambda) + g_{w_H} - \phi g_{A_F} &= \frac{\dot{V}^I}{V^I} \end{aligned}$$

where I used the fact that $f_R = \frac{H_F^{1-\lambda}}{\gamma_R}$. Plugging this back into (65) leads to

$$V_t^I = \frac{\exp\left(-\int_t^{t+\tau}(r+\delta_I)dx\right)}{r+\delta_I - g_{H_F}(1-\lambda) - g_{w_H} + \phi g_{A_F}} \frac{\alpha L_{t+\tau}^P w_{t+\tau}}{A_{F,t+\tau} z_{t+\tau}} \left[1+\tau'\right]$$

and with the free entry condition the following arbitrage condition holds on and off the balanced growth path,

$$\begin{aligned} \frac{w_{H,t}A_{F,t}^{-\phi}H_{F,t}^{1-\lambda}}{\gamma_R} &= \frac{\exp\left(-\int_t^{t+\tau}(r+\delta_I)dx\right)}{r+\delta_I - g_{H_F}(1-\lambda) - g_{w_H} + \phi g_{A_F}} \frac{\alpha L_{t+\tau}^P w_{t+\tau}}{A_{F,t+\tau} z_{t+\tau}} \left[1+\tau'\right] \\ &\Leftrightarrow \\ \frac{w_{H,t}A_{F,t}^{-\phi}H_{F,t}^{1-\lambda}}{\gamma_R} &= \frac{\exp\left(-\int_t^{t+\tau}\left(r+\delta_I - g_w - g_{L^P} + g_A + g_z\right)dx\right)}{r+\delta_I - g_{H_F}(1-\lambda) - g_{w_H} + \phi g_{A_F}} \frac{\alpha L_t^P w_t}{A_{F,t} z_t} \left[1+\tau'\right]. \end{aligned}$$

Now in the steady state it is easy to see that

$$\begin{aligned} \frac{w_{H,t}A_{F,t}^{-\phi}H_{F,t}^{1-\lambda}}{\gamma_R} &= \frac{\exp\left(-\int_t^{t+\tau}\left(r+\delta_I - g_w - g_L + g_A\right)dx\right)}{r+\delta_I - g_{H_F}(1-\lambda) - g_{w_H} + \phi g_{A_F}} \frac{\alpha L_t^P w_t}{A_{F,t} z_t} \\ &= \frac{\exp\left(-\tau\left(r+\delta_I - g_w - g_L + g_A\right)\right)}{r+\delta_I - g_{H_F}(1-\lambda) - g_{w_H} + \phi g_{A_F}} \frac{\alpha L_t^P w_t}{A_{F,t} z_t} \\ &= \frac{\exp\left(\log z\left(\frac{r+\delta_I - g_w - g_L + g_A}{g_A+\delta_I}\right)\right)}{r+\delta_I - g_{H_F}(1-\lambda) - g_{w_H} + \phi g_{A_F}} \frac{\alpha L_t^P w_t}{A_{F,t} z_t} \\ &= \frac{z^{\frac{r-g_w-g_L}{g_A+\delta_I}+1}}{r+\delta_I - g_{H_F}(1-\lambda) - g_{w_H} + \phi g_{A_F}} \frac{\alpha L_t^P w_t}{A_{F,t} z_t} \\ \frac{w_{H,t}A_{F,t}^{-\phi}H_{F,t}^{1-\lambda}}{\gamma_R} &= \frac{(z)^{\frac{\rho-g_L}{g_A+\delta_I}}}{r+\delta_I - g_{H_F}(1-\lambda) - g_{w_H} + \phi g_{A_F}} \frac{\alpha L_t^P w_t}{A_{F,t}} \\ \frac{w_{H,t}A_{F,t}^{-\phi}H_{F,t}^{1-\lambda}}{\gamma_R} &= \frac{(z)^{\frac{\rho-g_L}{g_A+\delta_I}}}{\rho - g_L + \delta_I + g_L(\lambda) + \frac{\phi}{1-\phi}\lambda g_L} \frac{\alpha L_t^P w_t}{A_{F,t}} \\ \frac{w_{H,t}A_{F,t}^{-\phi}H_{F,t}^{1-\lambda}}{\gamma_R} &= \frac{(z)^{\frac{\tilde{\rho}}{g_A+\delta_I}}}{\tilde{\rho}+\delta_I+g_A} \frac{\alpha L_t^P w_t}{A_{F,t}} \end{aligned}$$

where I used that in the steady state $g_A = \frac{\lambda}{1-\phi} g_L$ and $\tilde{\rho} = \rho - g_L$.

Further, note that I can write the free entry condition as a function of the time invariant piece of the fixed cost, γ, and ratios that are stable in the steady state. Define $h_{F,t} := \frac{H_F}{L}$ and $a_{F,t} = \frac{A_{F,t}^{1-\phi}}{L_t^{\lambda}}$, then

$$\begin{aligned}
\frac{w_{H,t}A_{F,t}^{-\phi}H_{F,t}^{1-\lambda}}{\gamma_R} &= \frac{\exp\left(-\int_t^{t+\tau}\left(r+\delta_I-g_w-g_{L^P}+g_A+g_z\right)dx\right)}{r+\delta_I-g_{H_F}(1-\lambda)-g_{w_H}+\phi g_{A_F}}\frac{\alpha L_t^P w_t}{A_{F,t}z_t}\left[1+\tau'\right] \\
&\Leftrightarrow \\
\frac{1}{\gamma_R} &= \frac{\exp\left(-\int_t^{t+\tau}(r+\delta_I)dx\right)}{r+\delta_I-g_{H_F}(1-\lambda)-g_{w_H}+\phi g_{A_F}}\frac{\alpha L_{t+\tau}^P w_{t+\tau}}{A_{F,t+\tau}z_{t+\tau}}\frac{1}{w_{H,t}}A_{F,t}^{\phi}H_{F,t}^{\lambda-1}\left[1+\tau'\right] \\
&= \frac{\exp\left(-\int_t^{t+\tau}(r+\delta_I)dx\right)}{r+\delta_I-g_{H_F}(1-\lambda)-g_{w_H}+\phi g_{A_F}}\frac{\alpha L_{t+\tau}^P/L_t^P w_{t+\tau}/w_t}{(A_{F,t+\tau}/A_{F,t})z_{t+\tau}}\frac{L_t^P}{s_t}A_{F,t}^{\phi-1}H_{F,t}^{\lambda-1}\left[1+\tau'\right] \\
&= \frac{\exp\left(-\int_t^{t+\tau}\left(r+\delta_I-g_w-g_{L^P}+g_{A_F}\right)dx\right)}{r+\delta_I-g_{H_F}(1-\lambda)-g_{w_H}+\phi g_{A_F}}\frac{\alpha l_t^P}{z_{t+\tau}}\frac{1}{s_t}\frac{L_t}{H_{F,t}}\frac{H_{F,t}^{\lambda}}{A_{F,t}^{1-\phi}}\left[1+\tau'\right] \\
\frac{1}{\gamma_R} &= \frac{\exp\left(-\int_t^{t+\tau}\left(r+\delta_I-g_w-g_{L^P}+g_{A_F}\right)dx\right)}{r+\delta_I-g_{H_F}(1-\lambda)-g_{w_H}+\phi g_{A_F}}\frac{\alpha l_t^P}{z_{t+\tau}}\frac{1}{s_t}\frac{1}{h_{F,t}}\frac{h_{F,t}^{\lambda}}{a_{F,t}}\left[1+\tau'\right]
\end{aligned}$$

and in the steady state the demand for skilled labor in research can be derived combining the free entry condition with the resource constraint. First, normalize the resource constraint

$$\begin{aligned}
\dot{A_F} &= \gamma_R A_F^{\phi}H_F^{\lambda}-\delta_I A_F \\
(g_{A_F}+\delta_I) &= \frac{\gamma_R h_F^{\lambda}}{a_F}
\end{aligned}$$

and now combine the two to get

$$h_F = \frac{g_A+\delta_I}{\tilde{\rho}+\delta_I+g_A} * \left(\frac{\alpha l_t^P}{s}\right) * (z)^{\frac{\tilde{\rho}}{g_A+\delta_I}} .$$

ss

Normalizing V^I by the cost of entry into innovation, $A_{F,t}^{-\phi}w_{H,t}f_R$, leads to the following normalized HJB equation

$$\left(r+\delta_I-g_{w_H}+\phi g_{A_F}\right)v_t^I-\dot{v}_t^I = \frac{\exp\left(-\int_t^{t+\tau}\left(r+\delta_I-g_{w_H}+\frac{\phi}{1-\phi}\left[g_{a_F}+g_L\right]\right)dx\right)\frac{\alpha l_{t+\tau}^P}{s_{t+\tau}a_{F,t+\tau}}\frac{1+\tau'}{z_{t+\tau}}}{r+\delta_I-g_{w_H}+\phi g_{A_F}} .$$

As long as the entry condition is strictly binding, this leads to a simplified representation because $v_t^I = f_R$ and hence $\dot{v}_t^I = 0$. This implies that the value function equals

$$v_t^I = \exp\left(-\int_t^{t+\tau}\left(r+\delta_I-g_{w_H}+\frac{\phi}{1-\phi}\left[g_{a_F}+g_L\right]\right)dx\right)\frac{\frac{\alpha l_{t+\tau}^P}{s_{t+\tau}a_{F,t+\tau}z_{t+\tau}}}{r+\delta_I-g_{w_H}+\phi g_{A_F}}\left(1+\tau'\right).$$

One can rewrite this expression again in terms of the original value function so that

$$
\begin{aligned}
V_t^I &= \exp\left(-\int_t^{t+\tau}\left(r+\delta_I-g_{w_H}+\frac{\phi}{1-\phi}\left[g_{a_F}+g_L\right]\right)dx\right)\frac{\left(\frac{A_{F,t}}{A_{F,t+\tau}}\right)^{-\phi}\frac{w_{H,t}}{w_{H,t+\tau}}\frac{\alpha w_{t+\tau}L_{t+\tau}^P}{A_{F,t+\tau}z_{t+\tau}}}{r+\delta_I-g_{w_H}+\phi g_{A_F}}\left(1+\tau'\right) \\
V_t^I &= \exp\left(-\tau\left[r+\delta_I\right]\right)\frac{1}{r+\delta_I-g_{w_H}+\phi g_{A_F}}\frac{\alpha w_{t+\tau}L_{t+\tau}^P}{A_{t+\tau}}\left(1+\tau'\right)
\end{aligned}
$$

In the steady state is thus follows that the normalized value function equals

$$
v_t^I \quad = \quad \exp\left(-\tau\left(\rho+\delta_I-g_L+\tfrac{g_L}{1-\phi}\right)\right)\tfrac{\alpha}{\rho-g_L+\delta_I+\frac{g_L}{1-\phi}}\tfrac{l^P}{sa_F z}
$$

and can be rewritten in terms of the actual value function

$$
\begin{aligned}
V^I &= \exp\left(-\tau\left(\rho+\delta_I-g_L+\frac{g_L}{1-\phi}\right)\right)\frac{\alpha}{\rho-g_L+\delta_I+\frac{g_L}{1-\phi}}\frac{l^P}{sa_F z}w_H A_F^{-\phi} \\
&= \exp\left(\frac{\log z}{g_F+\delta_I}\left(\rho+\delta_I-g_L+\frac{g_L}{1-\phi}\right)\right)\frac{\alpha}{\rho-g_L+\delta_I+\frac{g_L}{1-\phi}}\frac{L^P w}{a_F z}\frac{A_F^{-\phi}}{L} \\
&= z^{\left(\frac{\rho-g_L}{\frac{g_L}{1-\phi}+\delta_I}+1\right)}\frac{\alpha}{\rho-g_L+\delta_I+\frac{g_L}{1-\phi}}\frac{L^P w}{\frac{A_F^{1-\phi}}{L}z}\frac{A_F^{-\phi}}{L} \\
&= z^{\left(\frac{\rho-g_L}{\frac{g_L}{1-\phi}+\delta_I}\right)}\frac{\alpha}{\rho-g_L+\delta_I+\frac{g_L}{1-\phi}}\frac{L^P w}{A_F}
\end{aligned}
$$

as desired.

A.3.3 Innovation profits in the open economy

$$
\begin{aligned}
V_{I,t} \quad &= \quad \textstyle\int_{t+\delta}^{\infty}\exp\left(-\left(r+\delta_I\right)\left(u-t\right)\right)L_u^P w_{L,u}\alpha\left(\frac{1}{A_u}\right)du \\
&+ \quad \textstyle\int_{t+\delta^*}^{\infty}\exp\left(-\left(r+\delta_I\right)\left(u-t\right)\right)L_u^{P*}w_{L,u}^*\alpha\left(\frac{1}{A_u}\right)du \\
&= \quad \left(\tfrac{\alpha}{r-g_w-g_L+g_A+\delta_I}\right)\left\{\left(\tfrac{L_t^P w_{L,t}}{A_t}\right)\exp\left(-\left(r-g_w-g_L+g_A+\delta_I\right)\delta\right)+\left(\tfrac{L_t^{P*}w_{L,t}^*}{A_t^*}\right)\exp\left(-\left(r-g_w-g_L+g_A+\delta_I\right)\right.\right. \\
&= \quad \left(\tfrac{\alpha}{r-g_w-g_L+g_A+\delta_I}\right)\tfrac{L_t^P w_{L,t}}{A^F}\left\{\left(\tfrac{1}{z}\right)\exp\left(-\left(r-g_w-g_L+g_A+\delta_I\right)\delta\right)+\tfrac{L_t^{P*}w_{L,t}^*}{L_t^P w_{L,t}}\left(\tfrac{1}{z^*}\right)\exp\left(-\left(r-g_w-g_L+g_A+\delta_I\right.\right.\right. \\
&= \quad \left(\tfrac{\alpha}{\rho-g_L+g_A+\delta_I}\right)\tfrac{L_t^P w_{L,t}}{A^F}z^{\frac{\rho-g_L}{g_A+\delta_I}}\left\{1+\tfrac{L_t^{P*}w_{L,t}^*}{L_t^P w_{L,t}}\left(\tfrac{z^*}{z}\right)^{\frac{\rho-g_L}{g_A+\delta_I}}\right\}
\end{aligned}
$$

A.3.4 Waiting time for innovator

The waiting time for an innovator can be derived as follows. Recall equation (20). Use an integrating factor and note that on the balanced growth path with a constant adoption gap, g_A must equal g_F. Then,

$$\begin{aligned}
\dot{W}_t &= -\delta_I W - A_t\left(\delta_I + g_A\right) \\
\int_t^{t+\tau} \frac{\partial \exp(\delta_I u) W_u}{\partial u} &= -\int_t^{t+\tau} \exp\left(\delta_I u\right) A_u \left(\delta_I + g_A\right) du \\
\exp\left(\delta_I\left[t+\tau\right]\right) W_{t+\tau} - \exp\left(\delta_I t\right) W_t &= -A_0\left[\exp\left(\left[g_F + \delta_I\right]\left(t+\tau\right)\right) - \exp\left(\left[g_A + \delta_I\right] t\right)\right] \\
W_{\tau+t} &= \exp\left(-\delta_I \tau\right) X_t - A_0\left[\exp\left(\left[g_F\right]\left(t+\tau\right)\right) - \exp\left(-\delta_I\left[\tau\right]\right)\exp\left(\left[g_A\right] t\right)\right] \\
W_{\tau+t} &= \exp\left(-\delta_I \tau\right)\left[A_{F,t} - A_t\right] - A_{t+\tau}\left[1 - \exp\left(-\left[\delta_I + g_F\right]\left[\tau\right]\right)\right] \\
W_{\tau+t} &= \exp\left(-\delta_I \tau\right)\left[A_{F,t} - A_t\right] - \left[A_{t+\tau} - A_t \exp\left(-\delta_I \tau\right)\right] \\
W_{\tau+t} &= \exp\left(-\delta_I \tau\right)\left[A_{F,t}\right] - A_{t+\tau} \\
W_{\tau+t} &= \exp\left(-\delta_I \tau\right)\left[A_{F,t}\right] - \exp\left(g_A \tau\right) A_t
\end{aligned}$$

Now set $W\left(t, t+\tau\right) = 0$ so that

$$\begin{aligned}
\frac{A_t}{A_{F,t}} &= \exp\left(-\left[g_A + \delta_I\right]\tau\right) \\
&\Leftrightarrow \\
-\frac{\log z}{\delta_I + g_A} &= \tau.
\end{aligned}$$

The same argument applies to the case for no growth ($g_F = g_A = 0$) with the only difference that $A_t = A$. Moreover, the same argument applies to a more general version that implicitly defines the waiting time off the steady state:

$$\begin{aligned}
\dot{W}_t &= -\delta_I W - A_t\left(\delta_I + g_A\right) \\
\int_t^{t+\tau} \frac{\partial \exp(\delta_I u) W_u}{\partial u} &= -\int_t^{t+\tau} \exp\left(\delta_I u\right) A_u \left(\delta_I + g_A\right) du \\
\exp\left(\delta_I\left[t+\tau\right]\right) W_{t+\tau} - \exp\left(\delta_I t\right) W_t &= -\left[\exp\left(\delta_I\left[t+\tau\right]\right) A_{t+\tau} - \exp\left(\delta_I t\right) A_t\right] \\
W_{t+\tau} - W_t \exp\left(-\delta_I \tau\right) &= -A_{t+\tau} + \exp\left(-\delta_I \tau\right) A_t \\
W_{t+\tau} &= \left[A_{F,t} - A_t\right]\exp\left(-\delta_I \tau\right) - A_{t+\tau} + \exp\left(-\delta_I \tau\right) A_t
\end{aligned}$$

Now impose that $W\left(t, t+\tau\right) = 0$ so

$$\begin{aligned}
0 &= A_{F,t}\exp\left(-\delta_I \tau\right) - A_{t+\tau} \\
\frac{A_{F,t}}{A_t} &= \exp\left(\delta_I \tau\right)\frac{A_{t+\tau}}{A_t} \\
-\log z_t &= \tau\left[\delta_I + \frac{\int_t^{t+\tau} g_A(x) dx}{\tau}\right]
\end{aligned}$$

which generalizes and nests the steady state result.

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

$$\begin{aligned}\log A_{F,t} - \log A_t &= \delta_I \tau + \log A_{t+\tau} - \log A_t \\ \frac{\log A_{F,t} - \log A_{t+\tau}}{\delta_I} &= \tau.\end{aligned}$$

I totally differentiate this expression to obtain

$$\begin{aligned}g_{A_F} dt &= \delta_I d\tau + g_A(t+\tau)\, d\tau + g_A(t+\tau)\, dt \\ &\Leftrightarrow \\ \frac{d\tau}{dt} &= \frac{g_{A_{F,t}} - g_A(t+\tau)}{\delta_I + g_A(t+\tau)}\end{aligned}$$

One concern mentioned in the main text is to ensure that $1+\tau' > 0$, i.e. $\tau' > -1$. To see that this concern does not materialize, take note of the following inequality,

$$\begin{aligned}\frac{d\tau}{dt} &\geq -1 \\ &\Leftrightarrow \\ \frac{g_{A_{F,t}} - g_A(t+\tau)}{\delta_I + g_A(t+\tau)} &\geq -1 \\ &\Leftrightarrow \\ g_{A_{F,t}} - g_A(t+\tau) + \delta_I + g_A(t+\tau) &\geq 0 \\ g_{A_{F,t}} + \delta_I &\geq 0\end{aligned}$$

which shows that the derivative can never become too negative so that the flow profits are multiplied by a negative number. Note that I implicitly used the fact that $g_A > -\delta_I$. Note that under no learning effort whatsoever, $g_A = -\delta_I$ emerges as the knife-edge case which makes the derivative τ' explode. But, as long as $\beta \in (0,1)$, the firm will always pick an interior solution and invest at least a small amount in learning so that indeed $g_A > -\delta_I$. Thus this knife-edge case can be ruled out and generically $1+\tau' > 0$ holds.

A.3.5 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 ideas that are adopted at each instant. The probability density is then simply the product of the two, given statistical independence. Just as in the baseline case, a z close to unity makes the adoption friction vanish as the Poisson arrival rate of being adopted explodes. A z close to zero pushes the net present value of an innovation to zero as the probability of being adopted also approaches zero.

One can compute the expected present discounted value of a patent to obtain the research arbitrage condition. The different functional form leads to a different net present value but the insight that adoption and innovation are complementary on the market for ideas does not change and the reader might just as well use this stochastic formulation. The benefit is that stochastic adoption is more realistic in the sense that most certainly most innovators do not know when, if ever, their idea becomes profitable. The downside is that the market clearing condition is slightly more complicated.

A.4 Nesting Jones (1995)

Proposition 6. *Suppose $\delta_{ex} = \delta_I = 0$ and there is a sequence $k \in \{1, 2, 3, ...\}$, such that β_k and $f_{e,k}$ converges to zero from above, while σ_k is strictly increasing in k and unbounded, together with $\lim \beta_k (\sigma_k - 1) = 0$, $\lim \theta_k = 1$, and $\lim \sigma_k f_{e,k} \rho = b \in R^{++}$. Moreover, suppose that production labor and high-skilled labor are perfect substitutes so that $s = 1$ leading to a labor market clearing condition of the form $L = L^R + L^P$ for labor devoted to research or production, respectively. Then, the model is identical to Jones (1995).*

Intuitively, proposition 6 argues that there exists a sequence of parameters that lets the model converge to a competitive production side with no adoption gap at all. That sequence requires the adoption effort to decline ($\beta \to 0$) while the markup disappears ($\frac{\sigma}{\sigma-1} \to 1$), the spillover ($\theta \to 0$) disappears, and the fixed cost f_e goes to zero allowing for a competitive equilibrium.[84]

A.5 GDP Accounting

I decompose GDP into it's different components in the simple closed economy version of the model which helps clarify how to map the structure of the model to national accounts data.

$$\begin{aligned} gdp = & \quad Y + \dot{M} V_M + \dot{A_F} V_I \\ = & \quad Y + w L^E + w_H H^F \\ = & \quad C + \underbrace{Y - C}_{I_X} + \underbrace{w L^E}_{I_M} + \underbrace{w_H H^F}_{I_{A_F}} \end{aligned}$$

[84] For this limit to be well defined I need to make sure that convergence happens at the right rate so that the measure of firms M converges to some positive constant b. The measure of firms in the competitive equilibrium is usually not pinned down since constant-returns-to-scale in a perfectly competitive economy imply that firm size is irrelevant.

And the law of motion of capital, coming from the household budget constraint and the income side of the economy, reads

$$\begin{aligned}
\dot{X} &= rX + w\left(L^E + L^P\right) + w_H\left(H^D + H^F\right) + \Pi_P + \Pi_F - V_F\dot{A}_F - V_M\dot{M} - C \\
&= rX + w\left(L^E + L^P\right) + w_H\left(H^D + H^F\right) + \left(\frac{Y}{\sigma} - w_H H^D\right) + \Pi_F - V_F\dot{A}_F - V_M\dot{M} - C \\
&= Y + w\left(L^E\right) + w_H H^F - V_F\dot{A}_F - V_M\dot{M} - C \\
&= Y + w\left(L^E\right) - V_M\dot{M} - C \\
&= Y - C
\end{aligned}$$

which intuitively follows from total output minus total consumption of the final good.[85]

A.6 Open Economy Analytical Results

Proof that an increase in the fundamental research productivity of the home economy raises the skill premium at home and lowers the skill premium abroad.

First, note that market clearing can be rewritten as

$$\begin{aligned}
\left\{\frac{\chi}{z}\Lambda^{FO}\left((z)^{\frac{\tilde{\rho}}{g_A+\delta_I}+1} + (z^*)^{\frac{\tilde{\rho}}{g_A+\delta_I}+1}\right)\right\} &= sh^{tot} - \Lambda^D \\
\left\{\frac{\chi^*}{z^*}\Lambda^{FO}\left((z)^{\frac{\tilde{\rho}}{g_A+\delta_I}+1} + (z^*)^{\frac{\tilde{\rho}}{g_A+\delta_I}+1}\right)\right\} &= s^*h^{tot,*} - \Lambda^D
\end{aligned}$$

It follows that

$$\frac{\frac{\chi}{z}}{\frac{\chi^*}{z^*}} = \frac{sh^{tot} - \Lambda^D}{s^*h^{tot,*} - \Lambda^D}.$$

[85] Note that even though the human capital devoted to the adoption of new ideas is an investment activity from the firm's point of view, it won't show up that way in the national accounts data as this adoption related activity is not separated out from labor devoted to production.

Recall that $\left(\frac{\chi}{1-\chi}\right) = \left(\frac{\gamma}{\gamma^*}\right)^{\frac{1}{1-\lambda}} \left(\frac{s}{s^*}\frac{z}{z^*}\right)^{-\frac{\lambda}{1-\lambda}}$, and combining this with the previous equation yields

$$\begin{aligned}
\frac{\chi}{1-\chi} &= \left(\frac{z}{z^*}\right)\frac{sh^{tot}-\Lambda^D}{s^*h^{tot,*}-\Lambda^D}\\
\left(\frac{\gamma}{\gamma^*}\right)^{\frac{1}{1-\lambda}}\left(\frac{s}{s^*}\frac{z}{z^*}\right)^{-\frac{\lambda}{1-\lambda}} &= \left(\frac{z}{z^*}\right)\frac{sh^{tot}-\Lambda^D}{s^*h^{tot,*}-\Lambda^D}\\
&\Leftrightarrow\\
\left(\frac{\gamma}{\gamma^*}\right)^{\frac{1}{1-\lambda}}\left(\frac{s}{s^*}\frac{z}{z^*}\right)^{-\frac{\lambda}{1-\lambda}} &= \left(\frac{z}{z^*}\right)\frac{sh^{tot}-\Lambda^D}{s^*h^{tot,*}-\Lambda^D}\\
\left(\frac{\gamma}{\gamma^*}\right)^{\frac{1}{1-\lambda}}\left(\frac{s}{s^*}\left(\frac{s}{s^*}\right)^{-\frac{\beta}{1-\theta}}\right)^{-\frac{\lambda}{1-\lambda}} &= \left(\frac{s}{s^*}\right)^{-\frac{\beta}{1-\theta}}\frac{sh^{tot}-\Lambda^D}{s^*h^{tot,*}-\Lambda^D}\\
\left(\frac{\gamma}{\gamma^*}\right)\left(\left(\frac{s}{s^*}\right)^{\frac{1-\theta-\beta}{1-\theta}}\right)^{-\lambda} &= \left(\frac{s}{s^*}\right)^{-\frac{(1-\lambda)\beta}{1-\theta}}\left(\frac{sh^{tot}-\Lambda^D}{s^*h^{tot,*}-\Lambda^D}\right)^{1-\lambda}\\
\left(\frac{\gamma}{\gamma^*}\right) &= \left(\frac{s}{s^*}\right)^{\frac{(1-\theta-\beta)\lambda-\beta(1-\lambda)}{1-\theta}}\left(\frac{sh^{tot}-\Lambda^D}{s^*h^{tot,*}-\Lambda^D}\right)^{1-\lambda}\\
\left(\frac{\gamma}{\gamma^*}\right) &= \left(\frac{s}{s^*}\right)^{\frac{(1-\theta-\beta)\lambda-\beta(1-\lambda)}{1-\theta}}\left(\frac{sh^{tot}-\Lambda^D}{s^*h^{tot,*}-\Lambda^D}\right)^{1-\lambda}\\
\left(\frac{\gamma}{\gamma^*}\right) &= \left(\frac{s}{s^*}\right)^{\frac{(1-\theta-\beta)\lambda-\beta(1-\lambda)-(1-\lambda)(1-\theta)}{1-\theta}}\left(\frac{s}{s^*}\right)^{1-\lambda}\left(\frac{sh^{tot}-\Lambda^D}{s^*h^{tot,*}-\Lambda^D}\right)^{1-\lambda}\\
\left(\frac{\gamma}{\gamma^*}\right) &= \left(\frac{s}{s^*}\right)^{\frac{(1-\theta-\beta)\lambda+(1-\lambda)(1-\theta-\beta)}{1-\theta}}\left(\frac{h^{tot}-\frac{\Lambda^D}{s}}{h^{tot,*}-\frac{\Lambda^D}{s^*}}\right)^{1-\lambda}\\
\left(\frac{\gamma}{\gamma^*}\right) &= \left(\frac{s}{s^*}\right)^{\frac{1-\theta-\beta}{1-\theta}}\left(\frac{h^{tot}-\frac{\Lambda^D}{s}}{h^{tot,*}-\frac{\Lambda^D}{s^*}}\right)^{1-\lambda} \qquad (66)
\end{aligned}$$

Assumption that $\theta+\beta<1$ is important because you want that skilled labor becomes more expensive in real terms when demand goes up. If not, the real wage of skilled labor would be higher in places with a lower skill premium. If that is desired the reader can flip the inequality but care must be taken that the relevant computational inequalities, especially 1, is still respected.Within advanced economies, skilled labor seems to flock to high-skill premium as observed in knowledge flows to the US. Across the world as a whole, this relationship is less clear.

Now consider an increase in $\Delta\gamma>0$. I proof by contradiction that improving a country's comparative advantage in research will raise the skill premium in the home economy, while the skill premium falls in the foreign economy. To make this point, I consider a number of cases and show that they lead to contradictions.

1. $\frac{\Delta s}{s} > \frac{\Delta s^*}{s} > 0$.

In this case 66 may be consistent but it turns out that such a shift is not consistent with market clearing. Recall that foreign market clearing requires

$$\frac{\chi^*}{z^*}\Lambda^{FO}\left((z)^{\frac{\tilde{\rho}}{g_A+\delta_I}+1}+(z^*)^{\frac{\tilde{\rho}}{g_A+\delta_I}+1}\right)=s^*h^{tot,*}-\Lambda^D$$
$$\chi^*\Lambda^{FO}\left(\left(\frac{z}{z^*}\right)(z)^{\frac{\tilde{\rho}}{g_A+\delta_I}}+(z^*)^{\frac{\tilde{\rho}}{g_A+\delta_I}}\right)=s^*h^{tot,*}-\Lambda^D$$

Since the price of skill goes up everywhere, so the term in parentheses is going to decline. Note that z goes down and z/z^* goes down because $\frac{\Delta s}{s}>\frac{\Delta s^*}{s^*}$. Since the price of skill goes up everywhere, the right hand side is increasing and the only way that this market clearing condition holds is thus for $\chi^*$to increase. This implies that χ has to decline, which means that h^F must decline, which in turn means that market clearing does not hold in the home economy. Intuitively, how can the skill price rise if you do less research than before.

2.$\frac{\Delta s^*}{s^*}>\frac{\Delta s}{s}>0$.

This case is not consistent with an increase in γ,check equation 66.

3 & 4 & 5. One can rule out declining skill prices as well, using a similar argument. And having the foreign price of skill go up and the domestic price of skill decline can be ruled out as well.

6.$\frac{\Delta s}{s}>0>\frac{\Delta s^*}{s^*}$. This case is intuitive and in fact that only solution to an increase in the home economy's fundamental research productivity. Intuitively, improved comparative advantage means that the home economy specializes more in research. Since research is skill intensive, this drives up the price of skill in the home economy. The opposite happens in the foreign economy which specializes on producing final output. This releases skilled labor which pushes down the skill premium in the foreign economy.

A.7 Transitional Dynamics for aggregate economy model

A.7.1 Firm value function along the transition

Suppose that free entry into innovation and production holds. In that case, it must be that $f_e=v_t(t,z)$. Now the value function solves the HJB

$$(r_t+\delta_{ex}-g_w)\,v \quad = \quad \max_h \dot{v}+\tfrac{\pi_t(z)}{w}-s_th+v_z\dot{z}$$

Following similar steps as in the partial equilibrium derivation with fixed interest rate, this differential equation leads to a forward looking law of motion

$$\frac{\dot{h}}{h} \quad = \quad \frac{1}{1-\beta}\left\{\left(r_t+\delta_{ex}-g_w-\frac{\dot{s}}{s}+(1-\theta)(\delta_I+g_F)\right)-\frac{\beta z^\theta\zeta h^{\beta-1}}{s}\left[\frac{\pi}{w}\frac{(1-\alpha)(\sigma-1)}{z}\right]\right\}$$

that governs adoption effort out of steady state. This dynamic equation is tied to free entry through its dependence on profits. Moreover, note that the free entry condition implies $v_z=-\dot{v}$. I can use this relationship to obtain $(r_t + \delta_{ex} - g_w)\, v = \frac{\pi_t(z)}{w} - s_t h$ where it must be understood that h solves the dynamic adoption problem. Rearranging yields

$$v \;=\; \frac{\frac{\pi_t(z)}{w} - s_t h}{r_t + \delta_{ex} - g_w}$$

where I did not assume anything about the stationarity of any of the variables. If this is a solution, then it must be consistent with the free entry condition. Consider

$$\begin{aligned}
V &= \max \int_t^{\infty} \exp\left(-\int_t^u (r_k + \delta_{ex})\, dk\right) [\pi_u - w_{H,u} h_u]\, du \\
f_e &= \max \int_t^{\infty} \exp\left(-\int_t^u (r_k + \delta_{ex})\, dk\right) \left[\frac{\pi_u}{w_u}\frac{w_u}{w_t} - \frac{w_{H,u}}{w_u}\frac{w_u}{w_t} h_u\right] du \\
f_e &= \max \int_t^{\infty} \exp\left(-\int_t^u (r_k + \delta_{ex} - g_{w,u})\, dk\right) [(r + \delta_{ex} - g_w)\, f_e]\, du \\
f_e &= f_e
\end{aligned}$$

and note that indeed the guess solves the value function. This simplicity is due to the fact that the free entry condition at any point disciplines the profits that an incumbent firm can earn. 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. Further simplifying the adoption decision yields

$$\begin{aligned}
\frac{\dot h}{h}(1-\beta) &= \left\{\left(r_t + \delta_{ex} - g_w - \frac{\dot s}{s} + (1-\theta)(\delta_I + g_F)\right) - \frac{\beta(1-\alpha)(\sigma-1) z^{\theta} \zeta h^{\beta}}{zsh}\left[\frac{1}{w}\frac{Y}{M}\frac{1}{\sigma}\right]\right\} \\
\frac{\dot h}{h}(1-\beta) &= \left\{\left(r_t + \delta_{ex} - g_w - \frac{\dot s}{s} + (1-\theta)(\delta_I + g_F)\right) - \frac{\beta(1-\alpha)(\sigma-1) z^{\theta} \zeta h^{\beta}}{zsh}\left[\frac{L^P}{1-\alpha}\left(\frac{\sigma}{\sigma-1}\right)\frac{1}{M}\frac{1}{\sigma}\right]\right\} \\
\frac{\dot h}{h}(1-\beta) &= \left\{\left(r_t + \delta_{ex} - g_w - \frac{\dot s}{s} + (1-\theta)(\delta_I + g_F)\right) - \frac{\beta z^{\theta} \zeta h^{\beta}}{zsh}\left[\frac{L^P}{M}\right]\right\} \\
\frac{\dot h}{h}(1-\beta) &= \left\{\left(r_t + \delta_{ex} - g_w - \frac{\dot s}{s} + (1-\theta)(\delta_I + g_F)\right) - \frac{\beta z^{\theta} \zeta h^{\beta}}{zsh}\left[\frac{l^P}{m}\right]\right\}
\end{aligned}$$

where $l^P = L^P/L$ and $m = M/L$ are normalized variables that are constant in the steady state but not along the transition path. Moreover, define $a = \frac{A^{1-\phi}}{L}$, $a_F = \frac{A_F^{1-\phi}}{L}$ and $\tilde h = H/L$, and note that $g_a = (1-\phi)\, g_A - g_L \Leftrightarrow g_A = \frac{g_a + g_L}{1-\phi}$ and similarly, $g_{A_F} = \frac{g_{a_F} + g_L}{1-\phi}$. Then I can rewrite the law of motion of adoption to get

$$\begin{aligned}
\frac{\dot h}{h}(1-\beta) &= \left(r_t + \delta_{ex} - g_w - \frac{\dot s}{s} + (1-\theta)\left(\delta_I + \frac{g_{a_F} + g_L}{1-\phi}\right)\right) - \frac{\beta z^{\theta-1} \zeta h^{\beta-1}}{s}\left[\frac{l^P}{m}\right] \\
\frac{\dot h}{h}(1-\beta) &= \left(r_t + \delta_{ex} - g_w - \frac{\dot s}{s} + (1-\theta)\left(\delta_I + \frac{g_{a_F} + g_L}{1-\phi}\right)\right) - \frac{\beta \left(\frac{a_F}{a}\right)^{\frac{1-\theta}{1-\phi}} \zeta h^{\beta-1}}{s}\left[\frac{l^P}{m}\right]
\end{aligned}$$

A.7.2 Innovation Problem along the transition path

Differentiating the HJB equation of an innovator reads[86]

$$\left(r+\delta_I\right) V_t^I-\dot{V_t^I}=\exp\left(-\int_t^{t+\tau}\left(r+\delta_I\right) dx\right) \frac{\alpha L_{t+\tau}^P w_{t+\tau}}{A_{F,t+\tau} z_{t+\tau}}\left[1+\tau'\right].$$

Now you normalized by $A_{F,t}^{-\phi} w_{H,t}$ and obtain

$$\begin{aligned}
\left(r+\delta_I-g_{w_H}+\phi g_{A_F}\right) v_t^I-\dot{v_t^I} &= \exp\left(-\int_t^{t+\tau}\left(r+\delta_I\right) dx\right) \frac{\alpha L_{t+\tau}^P w_{t+\tau}}{A_{F,t}^{-\phi} w_{H,t} A_{F,t+\tau} z_{t+\tau}}\left[1+\tau'\right] \\
\left(r+\delta_I-g_{w_H}+\phi g_{A_F}\right) v_t^I-\dot{v_t^I} &= \exp\left(-\int_t^{t+\tau}\left(r+\delta_I\right) dx\right) \frac{\alpha l_{t+\tau}^P}{s_{t+\tau}} \frac{1}{a_{F,t+\tau} z_{t+\tau}}\left(\frac{w_{H,t+\tau}}{w_{H,t}}\right)\left(\frac{A_{F,t}}{A_{F,t+\tau}}\right)^\phi\left[1+\tau'\right] \\
\left(r+\delta_I-g_{w_H}+\phi g_{A_F}\right) v_t^I-\dot{v_t^I} &= \exp\left(-\int_t^{t+\tau}\left(r+\delta_I\right) dx\right) \frac{\alpha l_{t+\tau}^P}{s_{t+\tau}} \frac{1}{a_{F,t+\tau} z_{t+\tau}}\left(\frac{w_{H,t+\tau}}{w_{H,t}}\right)\left(\frac{a_{F,t}}{a_{F,t+\tau}} \frac{L_t}{L_{t+\tau}}\right)^{\frac{\phi}{1-\phi}}\left[1+\tau'\right] \\
\left(r+\delta_I-g_{w_H}+\phi g_{A_F}\right) v_t^I-\dot{v_t^I} &= \exp\left(-\int_t^{t+\tau}\left(r+\delta_I-g_{w_H}+\frac{\phi}{1-\phi}\left[g_{a_F}+g_L\right]\right) dx\right) \frac{\alpha l_{t+\tau}^P}{s_{t+\tau} a_{F,t+\tau}} \frac{1+\tau'}{z_{t+\tau}}
\end{aligned}$$

and the HJB for selling to the foreign market reads

$$\left(r+\delta_I-g_{w_H}+\phi g_{A_F}\right) v_t^{I*}-\dot{v_t^{I*}}=\exp\left(-\int_t^{t+\tau^*}\left(r+\delta_I-g_{w_H}+\frac{\phi}{1-\phi}\left[g_{a_F}+g_L\right]\right) dx\right) \frac{\alpha l_{t+\tau^*}^{P*}}{s_{t+\tau^*} a_{F,t+\tau^*}} \frac{z_{t+\tau^*}^*}{z_{t+\tau^*}} \frac{1+\tau'^*}{z_{t+\tau^*}^*}$$

where I used the fact that $\frac{z}{z^*}=\frac{w}{w^*}$. Next, note that by free entry

$$v^I+v^{I*}=f_R$$

and thus as long as the free entry condition is binding it must be that $\dot{v^I}+\dot{v^{I*}}=0$. This condition is quite crucial as it simplifies the problem: it ties the future profit flow to the current cost of entry, taking into account the appropriate discounting. This discounting rate depends on standard parameters such as the effective discount rate $\left(r+\delta_I\right)$ and the change in the entry cost, but is also features changes in the waiting time. Adding up the two HJB equations leads to a free entry condition that has to hold on and off the balanced growth path as long as the free entry condition is binding:

$$\begin{aligned}
\left(r+\delta_I-g_{w_H}+\phi g_{A_F}\right)\left(v_t^I+v_t^{I*}\right) &= \frac{\alpha}{s_{t+\tau} a_{F,t+\tau}} \frac{1}{z_{t+\tau}} \exp\left(-\int_t^{t+\tau}\left(r+\delta_I-g_{w_H}+\frac{\phi}{1-\phi}\left[g_{a_F}+g_L\right]\right) dx\right) l_{t+\tau}^P\left(1+\right. \\
&+ \frac{\alpha}{s_{t+\tau^*} a_{F,t+\tau^*}} \frac{1}{z_{t+\tau^*}^*} \exp\left(-\int_t^{t+\tau^*}\left(r+\delta_I-g_{w_H}+\frac{\phi}{1-\phi}\left[g_{a_F}+g_L\right]\right) dx\right) l_{t+\tau^*}^{P*}\left(1+\tau\right. \\
\left(r+\delta_I-g_{w_H}+\phi g_{A_F}\right)\left(f_R\right) &= \frac{\alpha}{s_t a_{F,t+\tau}} \frac{1}{z_{t+\tau}} \exp\left(-\int_t^{t+\tau}\left(r+\delta_I-g_w+\frac{\phi}{1-\phi}\left[g_{a_F}+g_L\right]\right) dx\right) l_{t+\tau}^P\left(1+\tau'\right. \\
&+ \frac{\alpha}{s_t a_{F,t+\tau^*}} \frac{1}{z_{t+\tau^*}^*} \exp\left(-\int_t^{t+\tau^*}\left(r+\delta_I-g_w+\frac{\phi}{1-\phi}\left[g_{a_F}+g_L\right]\right) dx\right) l_{t+\tau^*}^{P*}\left(1+\tau'^*\right) \\
\left(r+\delta_I-g_{w_H}+\phi g_{A_F}\right)\left(\frac{f_R s_t}{\alpha}\right) a_{F,t} &= \frac{1}{z_{t+\tau}}\left(\frac{a_{F,t}}{a_{F,t+\tau}}\right)^{\frac{1}{1-\phi}} \exp\left(-\int_t^{t+\tau}\left(r+\delta_I-g_w+\frac{\phi}{1-\phi} g_L\right) dx\right) l_{t+\tau}^P\left(1+\tau'\right) \\
&+ \frac{1}{z_{t+\tau^*}^*}\left(\frac{a_{F,t}}{a_{F,t+\tau^*}}\right)^{\frac{1}{1-\phi}} \exp\left(-\int_t^{t+\tau^*}\left(r+\delta_I-g_w+\frac{\phi}{1-\phi} g_L\right) dx\right) l_{t+\tau^*}^{P*}\left(1+\tau'^*\right)
\end{aligned}$$

[86] Note that $V^I=\int_{t+\tau}^{\infty} \exp\left(-\int_t^u\left(r+\delta_I\right) dx\right) \pi_u du$,differentiating this expression leads to the HJB.

I need to know τ to proceed. Using (**??**) I know that

$$-\frac{\log z_t}{\left[\delta_I + \frac{\int_t^{t+\tau} g_A dx}{\tau}\right]} \quad = \quad \tau_t$$

which subsumes the steady state result. Define $\mathbb{E}_{\tau(t)}\left[g_A\right] := \frac{\int_t^{t+\tau} g_A dx}{\tau}$. Moreover, the derivative of the time gap τ reads

$$\frac{d\tau}{dt} \quad = \quad \frac{g_{A_{F,t}} - g_A(t+\tau)}{\delta_I + g_A(t+\tau)} \quad .$$

Combining this with the free entry condition leads to the following expression

$$\begin{aligned}
\left(r + \delta_I - g_{w_H} + \phi g_{A_F}\right)\left(\frac{f_R s_t}{\alpha}\right) a_{F,t} \quad &= \quad \frac{1}{z_{t+\tau}}\left(\frac{a_{F,t}}{a_{F,t+\tau}}\right)^{\frac{1}{1-\phi}} \exp\left(\frac{\log z_t}{\left[\delta_I + \frac{\int_t^{t+\tau} g_A dx}{\tau}\right]}\left(r + \delta_I - g_w + \frac{\phi}{1-\phi} g_L\right) dx\right) l_{t+\tau}^P (1 - \\
&+ \quad \frac{1}{z_{t+\tau^*}^*}\left(\frac{a_{F,t}}{a_{F,t+\tau^*}}\right)^{\frac{1}{1-\phi}} \exp\left(\frac{\log z_t^*}{\left[\delta_I + \frac{\int_t^{t+\tau^*} g_{A^*} dx}{\tau^*}\right]}\left(r + \delta_I - g_w + \frac{\phi}{1-\phi} g_L\right) dx\right) l_{t+\tau^*}^{P^*} (1 + \\
\left(r + \delta_I - g_{w_H} + \phi g_{A_F}\right)\left(\frac{f_R s_t}{\alpha}\right) a_{F,t} \quad &= \quad \frac{1}{z_{t+\tau}}\left(\frac{a_{F,t}}{a_{F,t+\tau}}\right)^{\frac{1}{1-\phi}} \exp\left(\frac{\log z_t}{\left[\delta_I + \frac{\int_t^{t+\tau} g_A dx}{\tau}\right]}\left(r + \delta_I - g_w + \frac{\phi}{1-\phi} g_L\right) dx\right) l_{t+\tau}^P (1 - \\
&+ \quad \frac{1}{z_{t+\tau^*}^*}\left(\frac{a_{F,t}}{a_{F,t+\tau^*}}\right)^{\frac{1}{1-\phi}} \exp\left(\frac{\log z_t^*}{\left[\delta_I + \frac{\int_t^{t+\tau^*} g_{A^*} dx}{\tau^*}\right]}\left(r + \delta_I - g_w + \frac{\phi}{1-\phi} g_L\right) dx\right) l_{t+\tau^*}^{P^*} (1 + \\
\left(r + \delta_I - g_{w_H} + \phi g_{A_F}\right)\left(\frac{f_R s_t}{\alpha}\right) a_{F,t} \quad &= \quad \frac{1}{z_{t+\tau}}\left(\frac{a_{F,t}}{a_{F,t+\tau}}\right)^{\frac{1}{1-\phi}} z_t^{\left(\frac{r+\delta_I - \mathbb{E}_{\tau(t)}[g_w] + \frac{\phi}{1-\phi} g_L}{\delta_I + \mathbb{E}_{\tau(t)}\left[g_A\right]}\right)} l_{t+\tau}^P \left(1+\tau'\right) \\
&+ \quad \frac{1}{z_{t+\tau^*}^*}\left(\frac{a_{F,t}}{a_{F,t+\tau^*}}\right)^{\frac{1}{1-\phi}} z_t^{*\left(\frac{r+\delta_I - \mathbb{E}_{\tau^*(t)}[g_w] + \frac{\phi}{1-\phi} g_L}{\delta_I + \mathbb{E}_{\tau^*(t)}\left[g_{A^*}\right]}\right)} l_{t+\tau^*}^{P^*} \left(1+\tau'^*\right) \frac{z_{t+\tau^*}^*}{z_{t+\tau^*}} \\
\left(r + \delta_I - g_{w_H} + \phi g_{A_F}\right)\left(\frac{f_R s_t}{\alpha}\right) a_{F,t} \quad &= \quad \frac{z_t}{z_{t+\tau}}\left(\frac{a_{F,t}}{a_{F,t+\tau}}\right)^{\frac{1}{1-\phi}} z_t^{\left(\frac{r - \mathbb{E}_{\tau(t)}[g_w] - g_L + \frac{g_L}{1-\phi} - \mathbb{E}_{\tau(t)}\left[g_A\right]}{\delta_I + \mathbb{E}_{\tau(t)}\left[g_A\right]}\right)} l_{t+\tau}^P \left(1+\tau'\right) \\
&+ \quad \frac{z_t^*}{z_{t+\tau^*}^*}\left(\frac{a_{F,t}}{a_{F,t+\tau^*}}\right)^{\frac{1}{1-\phi}} z_t^{*\left(\frac{r - \mathbb{E}_{\tau^*(t)}[g_w] - g_L + \frac{g_L}{1-\phi} - \mathbb{E}_{\tau^*(t)}\left[g_{A^*}\right]}{\delta_I + \mathbb{E}_{\tau^*(t)}\left[g_{A^*}\right]}\right)} l_{t+\tau^*}^{P^*} \left(1+\tau'^*\right) \frac{z_{t+\tau}^*}{z_{t+\tau}} \\
\left(r + \delta_I - g_{w_H} + \phi g_{A_F}\right)\left(\frac{f_R s_t}{\alpha}\right) a_{F,t} \quad &= \quad \left(\frac{a_t}{a_{t+\tau}}\right)^{\frac{1}{1-\phi}} z_t^{\left(\frac{r - \mathbb{E}_{\tau(t)}[g_w] - g_L + \frac{g_L}{1-\phi} - \mathbb{E}_{\tau(t)}\left[g_A\right]}{\delta_I + \mathbb{E}_{\tau(t)}\left[g_A\right]}\right)} l_{t+\tau}^P \left(1+\tau'\right) \\
&+ \quad \left(\frac{a_t^*}{a_{t+\tau}^*}\right)^{\frac{1}{1-\phi}} z_t^{*\left(\frac{r - \mathbb{E}_{\tau^*(t)}[g_w] - g_L + \frac{g_L}{1-\phi} - \mathbb{E}_{\tau^*(t)}\left[g_{A^*}\right]}{\delta_I + \mathbb{E}_{\tau^*(t)}\left[g_{A^*}\right]}\right)} \frac{z_{t+\tau^*}^*}{z_{t+\tau^*}} l_{t+\tau^*}^{P^*} \left(1+\tau'^*\right)
\end{aligned}$$

The free entry condition takes the simple form

$$\begin{aligned}\left(r+\delta_I-g_{w_H}+\tfrac{\phi}{1-\phi}\left(g_{a_F}+g_L\right)\right)\left(\tfrac{f_R s_t}{\alpha}\right)a_{F,t} \quad &= \left(\frac{a_t}{a_{t+\tau}}\right)^{\frac{1}{1-\phi}} z_t^{\left(\frac{r-\mathbb{E}_{\tau(t)}[g_w]-g_L+\frac{g_L}{1-\phi}-\mathbb{E}_{\tau(t)}[g_A]}{\delta_I+\mathbb{E}_{\tau(t)}[g_A]}\right)} l^P_{t+\tau}\left(1+\tau'\right) \\ &+ \left(\frac{a^*_t}{a^*_{t+\tau}}\right)^{\frac{1}{1-\phi}} z_t^{*\left(\frac{r-\mathbb{E}_{\tau^*(t)}[g_w]-g_L+\frac{g_L}{1-\phi}-\mathbb{E}_{\tau^*(t)}[g_{A^*}]}{\delta_I+\mathbb{E}_{\tau^*(t)}[g_{A^*}]}\right)} \frac{z^*_{t+\tau^*}}{z_{t+\tau^*}} l^{P^*}_{t+\tau^*}\left(1+\tau'^*\right).\end{aligned}$$

A.8 Household Problem and Law of Motion of Capital

I simplify the transition dynamics by focusing on the case where only capitalists make forward-looking consumption-saving choices, similar to Moll (2014), Kleinman, Liu, and Redding (2021), and Caliendo and Parro (2019), building on Angeletos (2007).

The euler equation together with the per capita budget constraint implies that

$$c_t = \rho \tilde{B}_t$$

where $\tilde{B}_t = \frac{B_t}{L_t}$ are per capita assets. Given log utility I can directly focus on the physical capital accumulation resource constraint since $C = (r-(\rho-g_L))\left(K+MV+\int V_I(x)\,dx\right)$, which implies that a fraction $(\rho-g_L)K$ will be consumed, while physical capital reproduces itself at rate rK, which already takes into account depreciation

$$\begin{aligned}\dot{K} &= rK-(\rho-g_L)K \\ \dot{K} &= (r+\delta_k)K-(\rho+\delta_k-g_L)K \\ \dot{K} &= \hat{\alpha}Y-(\rho-g_L+\delta_k)K\end{aligned}$$

with $\hat{\alpha} = \alpha^2 * \frac{\sigma-1}{\sigma}$. Note how both markups in production and innovation are encoded in this expression, which comes from the first order condition of cost minimization of the intermediate goods producer with respect to the capital good. Normalizing by effective units of labor, i.e. $k = \frac{K}{L^P A_F z}$, leads to a law of motion of effective units of capital

$$\begin{aligned}\frac{\dot{k}}{k} &= \frac{\dot{K}}{K}-g_{L^P}-g_F-g_z \\ &= \hat{\alpha}\frac{Y}{K}-(\rho-g_L+\delta_k)-g_{L^P}-g_F-g_z \\ &= \underbrace{\hat{\alpha}\frac{y}{k}-(\rho+g_F+\delta_k)}_{=0\text{ in steady state}}-\underbrace{(g_{L^P}-g_L)-g_z}_{=0\text{ in steady state}}\end{aligned}$$

Two final remarks are in order. First, note that the interest rate always is equal to $(r+\delta_k) = \alpha^2\frac{\sigma-1}{\sigma}\left(\frac{1}{\alpha}\right)^{\alpha}\left(\frac{1}{1-\alpha}\right)^{1-\alpha}\left(\frac{A_F z L^P}{K}\right)^{1-\alpha} = \alpha^2\frac{\sigma-1}{\sigma}\left(\frac{1}{\alpha}\right)^{\alpha}\left(\frac{1}{1-\alpha}\right)^{1-\alpha}(k)^{\alpha-1}$due to static demand for capital in the production sector. Second, note that $\frac{y}{k} = \left(\frac{1}{\alpha}\right)^{\alpha} \left(\frac{1}{1-\alpha}\right)^{1-\alpha} (k)^{\alpha}$ which is a resource constraint. Imposing $r = g_F + \rho$, one can solve for the steady state, and now the interest rate can be computed backwards using this law of motion for capital.

One thing that is left to prove is that the fact that there is asset accumulation in assets in the research sector as someone has to own the firm. Note, however, that firm entry requires only labor so my conjecture is that it has no implications for capital accumulation beyond the effects that are captured in changing A_F, z, g_{L^P}. Proof outstanding. Worst case, imagine there are two types of capitalists, one hold capital and the other ones invest in research and production sector firms, in which case the argument goes through for sure.

B Extensions

B.1 Immigration

INCOMPLETE AND CONJECTURE.

A fully integrated equilibrium behaves differently from the baseline model. Note that the factor price equalization theorem does not hold precisely because countries have different research productivities so goods market trade is no substitute for immigration. World output would be maximized by moving all workers from the emerging market to the advanced economy. If the skill ratio of the foreign economy is the same or higher, integration also improves welfare for each worker group. The welfare implications for the scenario where the foreign economy has a lower skill ratio are ambiguous.

Production workers in the home economy are losing as their factor becomes more abundant. Skilled labor is exposed to two different shocks. The production labor supply shock raises the skill premium unambiguously as can be seen by the market clearing condition (40). This suggests gains for skilled labor through a simple scarcity effect. Note, however, that a larger share of skilled labor is devoted to technology adoption since the production sector is expanding faster than the research sector. If the total amount of skilled labor devoted to research declines, which depends on the whole set of parameters and the difference in the skill-ratios, the real wage effects for skilled labor are ambiguous as rising adoption gap and declining overall research stock may reduce their real wage.[87]

Production workers are better off in the scenario where only skilled labor from the emerging markets are allowed to move. This has two effects. First, it pushes down the skill premium, boosting both innovation and adoption and raising real wage growth of production workers in the advanced economy. Second, there would be devastating consequences for the emerging market since skilled labor is the engine of development their economy would stop adopting new technology.

[87] A sufficient condition for skilled labor to strictly improve is to ensure that the total amount of skilled labor devoted to innovation does not decline and $\beta + \theta < 1$, the latter bounding the response of the adoption gap on skill prices.

B.1.1 Emerging Market Contributing to the World Technological Frontier

The scenario considered here is arguably too bleak, and the most benevolent development would be one where the emerging market eventually contributes to the technological frontier. To formalize this scenario, suppose that $\gamma = \gamma^*$ and $h = h^*$ but $z > z^*$ i.e. the emerging market starts out of steady state but is otherwise identical to the advanced economy. I know the steady state solution provides productivity gains to both economies according to the constant elasticity $d\log w = \frac{1}{1-\phi} d\log L$, so a doubling of market size raises wages relative to trend by $2^{1-\phi} - 1 \approx 40\%$ for $\phi = -1$.

Initially, research takes a backseat in economy that is out of steady state, since returns to adoption are higher. In the long run symmetric equilibrium with same amount of research. Transition dynamics to be completed soon.

B.1.2 Different Sectoral Factor Intensity and Endogenous Labor Supply

In the baseline model I assume that production only requires capital and production labor, while adoption and innovation only requires skilled labor. This should be viewed as a simplified limiting case of a model where innovation requires a composite labor input $G_I(H, L)$ that is produced according to a constant returns to scale production function. Differentiating the cost function that pertains to G_I with respect to H leads to the amount of skilled labor needed to produce one unit of the composite good, denoted by b_I, see Feenstra (2015)'s introduction to Heckscher-Ohlin theory of international trade. Assuming that $b_I > b_D > b_P$ is a useful generalization of the benchmark model so that each activity, innovation, adoption, and production, requires a mix of different labor types. I impose a strict ranking in terms of their factor intensity. Note that Heckscher-Ohlin theory and in particular the Rybczynski theorem would suggest an even stronger contraction in the production sector, but the gains from trade will be more broadly shared across worker types. Intuitively, this setting allows low skilled workers to benefit from gains in specialization in innovation.

Similar to the adjustment patterns in the model with composite labor goods, one can allow for an endogenous labor supply that will increase reallocation into innovation and ease the pressure on the skill premium. It would be easy, however, to extend the model by allowing workers to choose their education. One can incorporate this effortlessly into the market clearing condition for high-skilled labor (40) simply by letting the relative supply of skilled labor h^{tot} be a function of the skill premium $h^{tot} = h(s)$ s.t. $h'(s) > 0, h''(s) \geq 0$, and $h(1) = 0$.[88] Again, such a model offers more scope for production labor to gain from market integration.

[88] Micro-foundations to obtain an upward-sloping relative supply of skilled labor are plentiful, see for instance Acemoglu et al. (2018).

Part I

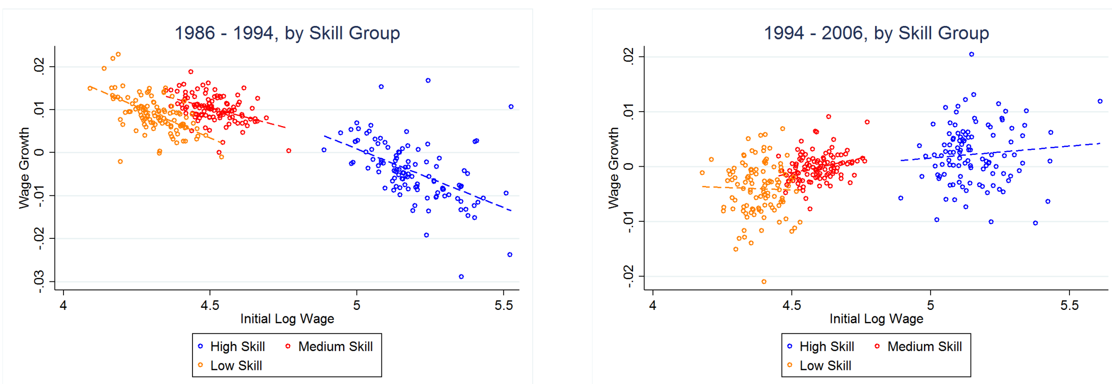
Empirical Appendix

C IAB DATA

The data I use is provided by the IAB and comprises an establishment panel (BHP) that constitutes a 50% random sample of establishments in Germany. The data contains the county in which the establishment is located, as well as sectoral information, and the number and composition of workers, including detailed information on educational attainment. I use Kosfeld and Werner (2012)'s definition of local labor markets (excluding Berlin) which leaves me with 108 regions.

D Changing Convergence Dynamics

Figure 12. Convergence by Skill Group

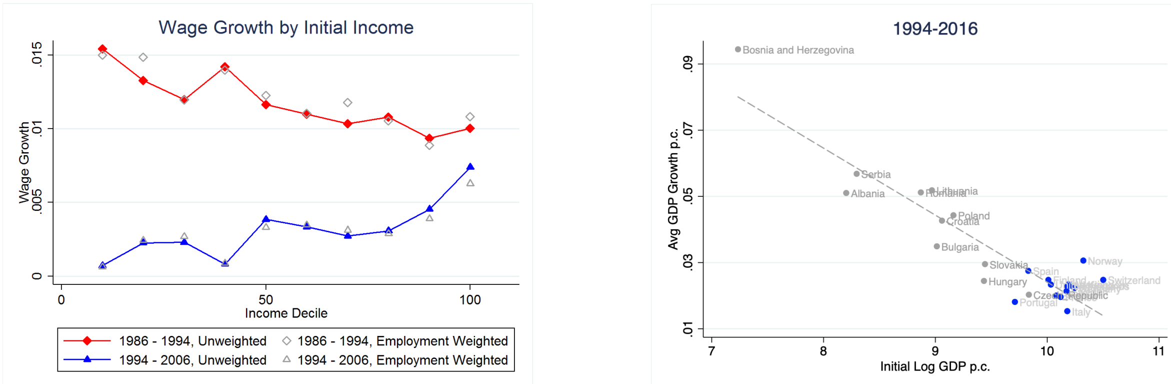


IAB BHP data. My plots.

The first point of changing convergence dynamics can best be illustrated by a classic Barro growth regression with log of initial GDP per capita on the x-axis and the average geometric per capita growth rate on the y-axis in figure 16 for regions in West Germany. While the negative relationship between initial income and growth shows up clearly in the pre 1990s, it has disappeared in the post 1990 economy.

A simple and naive exercise that illustrates the potential of "missing convergence" to account for the

Figure 13. Regional Divergence and Global Convergence



The left panel is based on the BHP dataset of the IAB. Regions are defined as local labor markets following Kosfeld and Werner (2012) which implies that there are 109 local labor markets in West Germany, each of which is assigned to a wage decile based on the average wage in the base period. The right hand side panel uses data from the Penn World Tables 9.0, see Feenstra, Inklaar, and Timmer (2016). Country income is measured in PPP.

aggregate productivity slowdown is to use the slope of the convergence relationship in 1960 – 1989 and predict counterfactual growth rates for the post 1990 period, where regions are properly aggregated to produce aggregate output per capita. To do this, I need to take a stance on the constant that I use for the prediction. I use the simple average of the top 5 regions in terms of initial income in the sample, represented by the solid black dot. Table **??** reports the naive counterfactual in the very last row on the right column: if convergence had continued, aggregate per capita GDP growth would have been 2.55 % instead of 2.18 %.

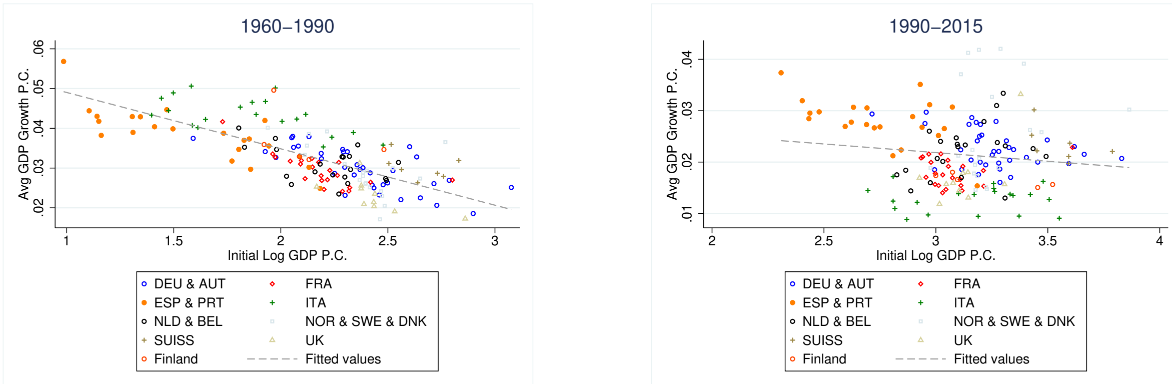
Note: The table reports results from a naive prediction exercise to quantify the lack of convergence for aggregate growth based on figure 16. The top 5 regions grew even slightly faster than the aggregate in 1990 -- 2015 but were omitted due to rounding.

Table 5. p.c. growth rates in West Germany

	1960 – 1990	1990 – 2015
observed agg growth	2.70	2.18
observed top 5 region growth	2.35	2.18
counterfactual agg growth	2.67	2.55

Table **??** performs the same exercise as table **??** but starting a decade later. The reason is that the micro data only starts in 1975. Note that growth has remarkably slowed but convergence is still a powerful engine of growth as the aggregate growth rate is substantially higher than the growth rate of the top 5 regions. The fact that convergence growth in Germany operates until the mid 90s is important since a major contribution of this paper is to offer new evidence on the establishment dynamics of regional catch-up growth in a large advanced economy. Of course, this simple exercise

Figure 14. Regional Convergence in Europe



The data is based on Rosés and Wolf (2018). I group small countries with very few internal regions such as Portugal or Austria to their larger neighbors, Spain and Germany respectively. I only consider West-Germany, all East German regions are dropped from the analysis, to make the sample comparable with the micro data and avoid the episode of state socialism in the former DDR.

is hard to discipline and the rest of the paper leverages the micro data to offer more evidence on the changing convergence dynamics.

E Additional Results from Barro Catch-up Regression for Regions in Germany

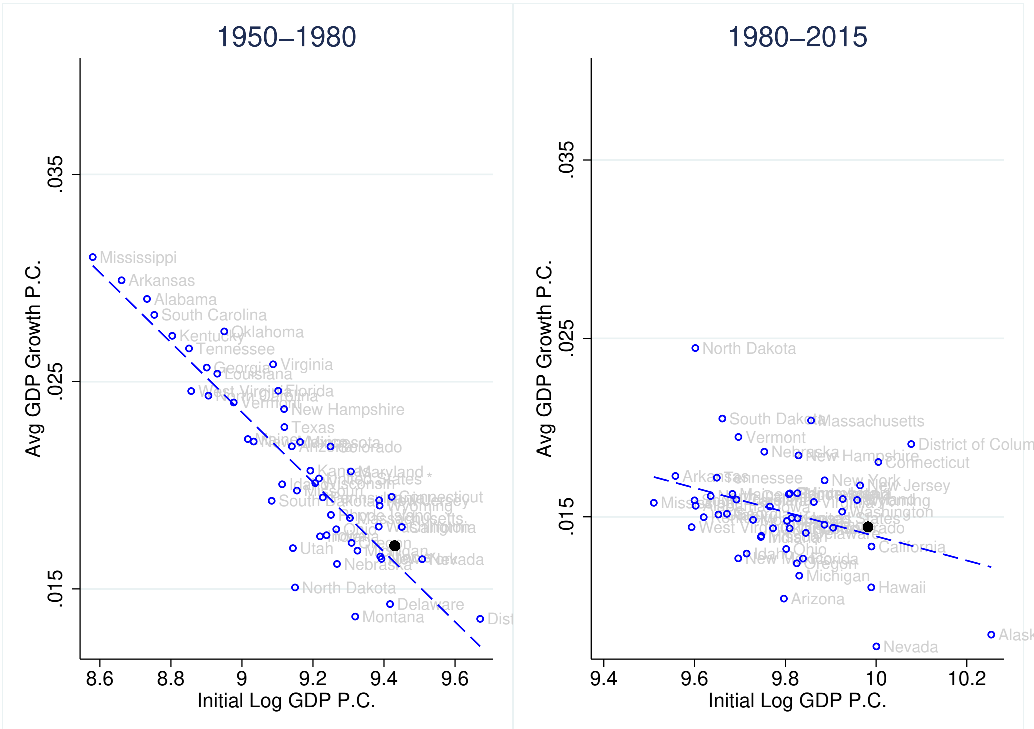
Employment, High-Skill Firms, and Professional Occupations

Note that measures of employment only considers full-time employees. When computing the number of high-skill establishments, I count every establishment as high-skill whenever strictly more than 33% of the full-time employees have a college degree. Professional occupations includes the following: technicians (az_bf_tec), semi professionals (az_bf_semi), engineers (az_bf_ing), professionals (az_bf_prof), and managers (az_bf_man). The definitions follow the Blossfeld occupational classification.

for International Trade

Building on the work of Autor, Dorn, and Hanson (2013) and Dauth, Findeisen, and Suedekum (2014), I use a shift-share approach that interacts the rise in trade with Eastern Europe as well as China with initial industry employment shares, to control for the effect of rising exports and imports over the sample period. Specifically, I use the following measure of import exposure, $\Delta\,(Import\,exp)_{j,t}^{East} = \sum_j \frac{E_{j,i,t}}{E_{i,t}} \frac{\Delta Im_{i,t}^{D\leftarrow East}}{E_{j,t}}$, where $\Delta Im_{i,t}^{D\leftarrow East}$ is the total increase in real imports (total value deflated in 1998 Euros) from the East, here including both China as well as Eastern Europe and Eurasia. This choice is informed by the fact that the rise in the German trade-to-GDP ratio is largely attributable to the rise of China and the fall of the Iron Curtain (Dauth, Findeisen, and Suedekum, 2014). The relevant

Figure 15. Regional Convergence USA



This figure plots initial log per capita income against real per capita growth. The data is from the Bureau of Economic Analysis (https://www.bea.gov/data/gdp). The black dot indicates the average per capita growth of the top five regions in each period.

time interval to measure the increase in trade is chosen from 1996 to 2005. The initial employment shares are measured in 1994 using full-time workers only. While I don't instrument for trade flows as in Autor, Dorn, and Hanson (2013), I do use lagged initial shares in 1994 while the trade flows are measured from 1996 onwards. Measures for export exposure are analogous. I don't instrument for trade flows because I do not try to estimate causal effects. Instead, controlling for "endogenous" trade flows is a more challenging robustness test in this context precisely because it might pick up local demand and productivity shocks. Lagging the shares by two periods relative to the ADH approach is due to the fact that I do not have the data for 1995. The trade data are from BACI and the OECD trade in services statistics, and the sectoral classification used are WZ93 3-digit for manufacturing and WZ93 2-digit for services. I follow Becker et al. (2019) in mapping BACI and OECD industries to the German industry classification, see their paper for details.

Figure 16. Regional Convergence in Germany

Regional GDP data from Rosés and Wolf (2018). The black circles indicates simple averages of the top 5 regions in the respective base periods. In the right panel, this is the point which I force the projection of the counterfactual growth rates based on the 1960-1990 convergence relationship to go through, i.e. the grey line. The grey line and the blue line in the left panel, thus, have the same slope.

F Additional Information on Patent Data

The data is provided by Crios-Patstat Coffano and Tarasconi (2014) and contains patent data from the European Patent Office (EPO) from 1977 - 2014. I use the following files to build up the dataset:

- "priorities.txt", this file is important to take account of the priority date in order to get the timing of the paten counts right, as well as which year to assign a patent to.
- "applicants.txt", this file has information on inventors, and importantly on the location on the nuts3 level.

G Additional Information on Historical Databases

The historical databases are Rosés and Wolf (2018) for regional GDP in europe as well as Federico and Tena Junguito (2016) for long run exports shares. Note that the world trade historical database of Federico and Tena Junguito (2016) ends in 2014, so I supplement the sample with nominal export shares from the world bank. Note that while the world bank export share is highly correlated with the historical data (rho=.96) the levels are quite different for the period from 2000 - 2014. I therefore run a bivariat regression where I predict the measure of openness based on a regression for nominal export shares on world bank nominal export shares for the period 2000 - 2014 and the world bank export shares for 2015 - 2017. Alternatively one can use the 2014 measure to proxy for the 2015 moving average export share that I use in the main text. I do the same thing for individual countries, but because there is much more heterogeneity now, I only use the period from 2010 to 2014 to estimate the slope coefficient, and then project up until 2017 using slope coefficient and individual fixed effect. This ensures that there is no large jump at 2015, which would have been the case if using data from 2000 onward.

H Additional Information on Aggregate Wages and Employment

I work with the more detailed 5 industry code classification which helps me build a measure of innovative activity.

I work with az_ges when simply computing sector shares and employment shares because the data is more readily available. I do work with az_vz when computing average wages etc because the wages are averages over full time workers only.

H.1 Shifting Employment Patterns

Here is a list of sectors and which I classify as innovation vs. production, and I also compare how these employment patterns look when I include ICT and finance industries which are not part of the baseline plot.

H.2 Wage Stagnation in Germany

I plot average daily wages using the BHP data from the IAB over a long horizon. I plot the aggregate average wage, i.e. total labor income divided by total employment. As observed in a number of studies(Card, Heining, and Kline, 2013; Doepke and Gaetani, 2020) the skill premium does not respond as strong in the micro data than it does when using aggregate accounts from the KLEMS data. Wage stagnation, though, seems to be a trend that both series agree on.

Doepke and Gaetani (2020) argue that the skill-premium rose less in German due to specific features of the labor market. Note, however, that there is no disagreement of the overall rise in inequality since the 1990s. An alternative explanation is that first mis-measurement due to top coding (IAB data) and underreporting (SOEP data) leads to an understatement of the skill premium. And second, much of the inequality should play out among workers who are able to work in "innovative" industries relative to production-focused industries through the lens of my model. A worker's education is correlated with this, but not perfectly so. When I plot average wages across establishments in innovation and production in figure 19, a gap emerges just as it does in the KLEMS data, consistent with the main story in this paper and the overall rise in inequality.

H.3 Convergence Regressions in Germany

Note that in the period from 1986 - 1994, $\hat{\beta}_{Barro}$ equals -0.16. In contrast, the sign reverses in the period from 1994 - 2005, reading +0.16. Note that this constitutes a fundamental shift in the distribution of growth − from laggard regions to the most advanced. The estimates are robust to controlling for a host of variables measured in the base period as reported in table 6. Both a shift-share based measure of exporting following Autor, Dorn, and Hanson (2013) and average establishment size help span some of the growth of high-income regions. This is consistent with exports having a positive impact on wages, and in the model of Melitz (2003), larger firms benefit more from market integration. Importantly, note that a measure of import competition, using the same shift share approach does not help at all to understand changing growth dynamics. While the convergence coefficient changes little, the coefficient on imports is positive and has a p-value < 0.001, suggesting that importing intermediate goods helped a region to become more productive. Taken together, laggard regions grew very poorly not because they were directly exposed to import competition. It looks like they were left behind because they were untouched by globalization. This is precisely how the model works where production-centric regions stagnate because of a reallocation of skilled labor towards more innovative regions. Globalization matters, but indirectly through the rivalry on factor markets that leads to weak adoption in the hinterlands.[89]

I Patents and Growth

[89] Autor, Dorn, and Hanson (2013) focus on the employment margin of the China shock, and do not find strong wage effects. In the simple cross-sectional setting I use here, and without using their instrument, wage growth is positively related to both import and export exposure.

Table 6. Barro Coefficient with controls

	Controls in base period	$\hat{\beta}_{Barro}^{1986-1994}$		$\hat{\beta}_{Barro}^{1994-2006}$		obs
		Coeff.	SE	Coeff.	SE	
1.	-	**-0.0160**	.00434	**0.0183**	.00349	109
2.	avg. establishment size	**-.02114**	.00558	**0.01096**	.00425	109
3.	college share	**-0.0227**	.00658	**0.0309**	.00871	109
4.	manufacturing share	**-0.0152**	.00456	**0.0204**	.00306	109
5.	share of professional occupations	**-0.0120**	.00564	**0.0265**	.00499	109
6.	share of engineers and scientists	**-0.0260**	.00552	**0.01785**	.00750	109
7.	import exposure (shift share)	**NA**	NA	**0.01543**	.00312	109
8.	export exposure (shift share)	**NA**	NA	**0.01208**	.00385	109

This table reports the catch-up coefficient after controlling for the respective variable in logs. Standard errors are clustered at the regional level. The share of professional occupations includes the following occupation codes in the IAB: az_bf_tec, az_bf_semi, az_bf_ing, az_bf_prof, az_bf_man (technical, semi professional, engineers, professional, managers). See the IAB codebook for additional details (http://doku.iab.de/fdz/reporte/2016/DR_03-16_EN.pdf).

Table 7. Patents and Market Size

	(1) lg_d_patent_yoy_ws	(2) lg_d_patent_yoy_ws
lg_d_exports_all_GER_yoy_ws	0.00785***	
	(5.55)	
lg_d_exports_east_GER_yoy_ws		0.0858***
		(5.21)
_cons	9.279	12.52*
	(1.76)	(2.30)
N	119	119

t statistics in parentheses

* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

Figure 17. Sectoral Classification Innovation vs. Production

Industry Digit	variable_n	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	Effektenver 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 Zentralbanker	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	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 verbun	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 Nota	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 un	0	0
5-Steller	(w03_5)	74124	Praxen von Steuerbevollmaechtigte	0	0
5-Steller	(w03_5)	74125	Buchfuehrung (ohne Datenverarbe	0	0

Industry Di	variable_n	sector cod	label	baseline	baseline plus Fi
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 Hol	1	1
5-Steller	(w03_5)	74152	Managementtaetigkeiten von sor	1	1
5-Steller	(w03_5)	74153	Geschlossene Immobilienfonds m	0	0
5-Steller	(w03_5)	74154	Geschlossene Immobilienfonds m	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 Beratungsleistung	0	1
5-Steller	(w08_5)	62030	Betrieb von Datenverarbeitungse	0	1
5-Steller	(w08_5)	62090	Erbringung von sonstigen Dienstl	0	1
5-Steller	(w08_5)	63110	Datenverarbeitung, Hosting und	0	1
5-Steller	(w08_5)	63120	Webportale	0	0
5-Steller	(w08_5)	63910	Korrespondenz- und Nachrichtenl	0	0
5-Steller	(w08_5)	63990	Erbringung von sonstigen Informa	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 Genossenscha	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 u	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 Versicherungsmakl	0	0
5-Steller	(w08_5)	66290	Sonstige mit Versicherungsdienst	0	0
5-Steller	(w08_5)	66300	Fondsmanagement	0	1
5-Steller	(w08_5)	69101	Rechtsanwaltskanzleien mit Nota	0	0
5-Steller	(w08_5)	69102	Rechtsanwaltskanzleien ohne Not	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 Wirtschaftsprueferinr	0	0
5-Steller	(w08_5)	69202	Praxen von vereidigten Buchprue	0	0
5-Steller	(w08_5)	69203	Praxen von Steuerbevollmaechtig	0	0
5-Steller	(w08_5)	69204	Buchfuehrung (ohne Datenverarb	0	0
5-Steller	(w08_5)	70101	Managementtaetigkeiten von Hol	1	1
5-Steller	(w08_5)	70109	Sonstige Verwaltung und Fuehrur	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 Entwicklu	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 vo	0	0
5-Steller	(w08_5)	73200	Markt- und Meinungsforschung	1	1

Figure 18. Wage Stagnation in Germany over the long run

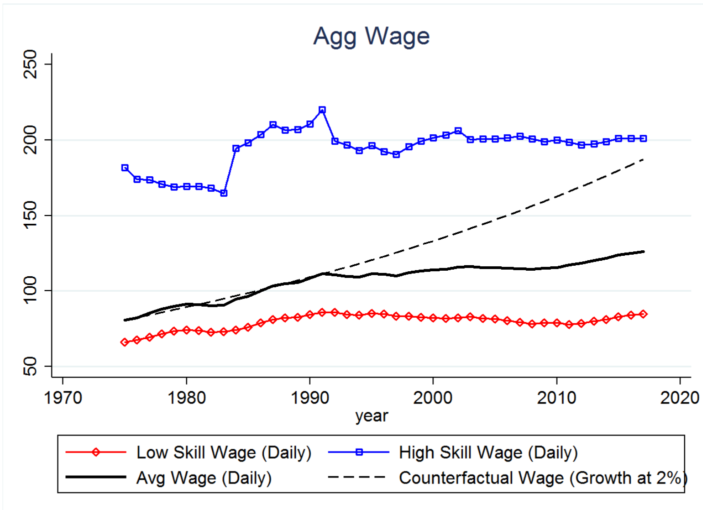
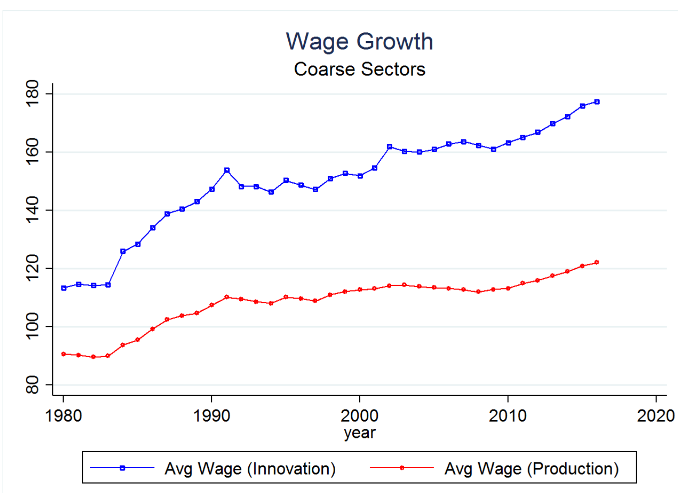
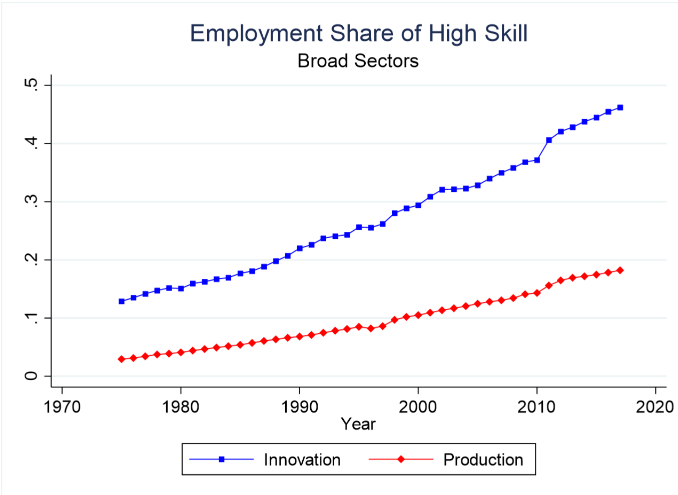


Figure 19. Wages in Innovation and Production



IAB BHP data. Average refers to the total wage bill of each group divided by the total number of employees.

Figure 20. Skill-Share across Sectors



IAB BHP data. Measure divides full time skilled labor in each sector-group by total full time employment. Note the divergence that sets in since the 1990s.

Figure 21. GDP & Patents Across Regions in West Germany

Nuts2 level of regional aggregation across West Germany. Plot log GDP per capita against log Patent, and note a stronger positive relationship between the two in the later period. The slope is different at 8% statistical significance level.