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

Florian Trouvain
University of Michigan

Job Market Paper

February 22, 2023

[most recent version here]

Abstract

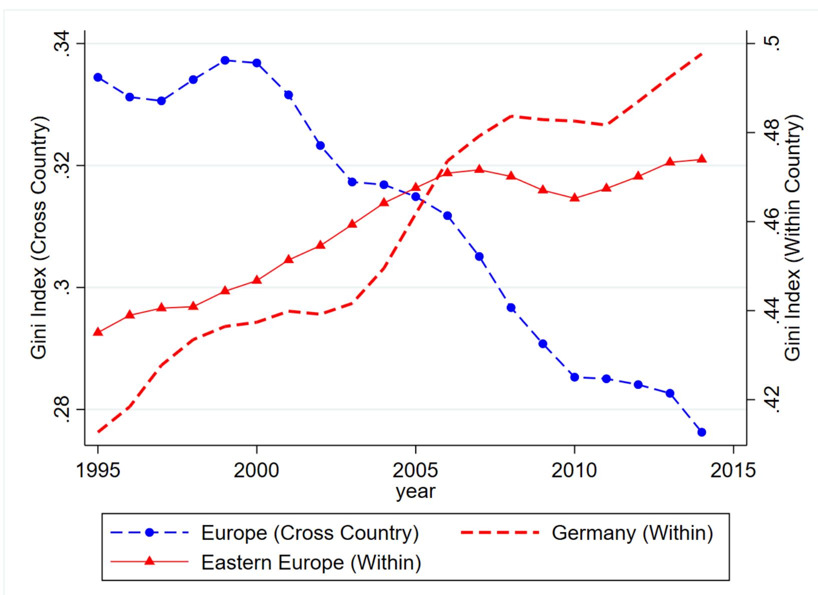
Economic Growth since the mid 1990s is characterized by i) declining cross-country inequality, ii) rising within-country inequality, and iii) overall weak growth in advanced economies. I provide a unifying explanation for these facts by developing a tractable general equilibrium model of long-run growth 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 within advanced economies. An increase in effective market size raises the returns to innovation. This pushes up the demand for skilled labor, which leads to a rising skill premium. As skilled labor is reallocated from domestic technology adoption toward global innovation, an innovation-adoption tradeoff emerges where the technological frontier expands while domestic technology adoption stalls. When pushing out the technological frontier is difficult, the growth drag from reduced adoption can dominate positive innovation effects, culminating in weak 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 five 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, Clement Imbert, 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, Mark Wright, and Fabrizio Zilibotti. 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 cross-country convergence and within-country divergence by plotting cross-country and within-country 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, see Milanovic (2016). While Eastern Europe experienced annual per capita growth of around 5% from 1995 to 2015, Western Europe fared less well. For example, Germany grew at an annual rate below 1%, which I single out here as it will be the focus of my empirical application.

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 build on Romer's benchmark endogenous growth model with two sectors, research and production, and two types of labor, high skilled and production labor. The research sector is standard and invents new technology using skilled labor. The production sector produces a final consumption good by combining idea embodying capital goods from the research sector with labor. My key departure is to introduce a technology adoption friction in the production sector, i.e. incorporating new ideas is a *costly and skill-intensive activity*. The rate at which new technology is adopted is determined endogenously by firms solving a dynamic problem. This setting leads 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.[1] Both this adoption gap, and the overall level of frontier technology, depend on the amount of skilled labor devoted to technology adoption and frontier innovation, respectively. The endogenous allocation of skilled labor across these two activities is the focus of this paper.[2]

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. Higher adoption effort in the production sector thus pushes up the net present value of innovation as the waiting time for a new idea to become profitable falls. 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, giving way to uneven economic growth where the innovation sector and skilled labor gain, while adoption activity and production worker wages in advanced economies stagnate. This happens in particular when advanced economies integrate with emerging markets where the former 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 describe 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 labor, the expansion of the innovation sector must come at the cost of reducing technology adoption

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

[2] I set aside the issue of the well-known innovation-production trade-off that is studied in Romer (1990) or Jones (1995) and instead focus on the allocation of skilled labor devoted to innovation relative to 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 main 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 emerging markets raises the returns to innovation, while *locally,* factor market competition leads to brain drain in the production sector within rich countries.

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. The model is thus able to confront and overturn the counterfactually strong pro-growth effects of market integration inherent to endogenous growth models. The pro-growth effects are directly tied to the non-rivalry of knowledge, which leads to increasing returns. The standard logic suggests that productivity growth is higher in the open economy as the advanced economy specializes in R&D, see Rivera-Batiz and Romer (1991). Weak productivity growth in the aftermath of globalization is thus puzzling. Introducing an endogenous adoption friction solves this puzzle as market integration can give rise to an innovation-adoption tradeoff that tames the strong pro-growth effects of the benchmark models of Romer (1990) and Jones (1995).

While the model maintains scale effects in innovation, a key departure is that there are constant-returns-to-scale in technology adoption. This feature generates an endogenous cross-country productivity distribution that is consistent with the data. The crucial determinant of cross-country productivity differences is the share of skilled labor relative to unskilled labor in the production sector and *not* the total amount of skilled labor, which follows form the constant-returns-to-scale property of technology adoption. The framework thus avoids counterfactual scale effects *across countries at a point in time*, i.e. country size is uncorrelated with productivity.[3]

A desirable feature of the framework is that divergence between research and production sector and rising inequality unfold only after integration between asymmetric countries where one party is the main supplier of innovation. This explains why globalization since the 1990s has had different effects compared to the process of trade integration among rich countries since WW2. In the case of symmetric countries, the benefit of exporting ideas exactly cancels with competition from abroad, leaving the returns to innovation and the skill premium unchanged. In contrast, the bias arises when emerging markets adopt technology while not contributing to the technological frontier with own innovation. Comparative advantage in innovation is thus crucial for the argument to hold. In addition, the framework offers a rationale for rising inequality in emerging markets. Suppressed technology adoption in the closed economy in the emerging market – perhaps due to government regulation – leads to suppressed demand for skilled labor. A lack of technology adoption by firms in the pre-reform

[3]See Jones (2005) and Ramondo, Rodríguez-Clare, and Saborío-Rodríguez (2016) for a discussion of scale effects in idea-based growth models and models of international trade.

period thus jointly explains low average income and low inequality.

The theory finds a direct empirical counterpart in the growth experience of advanced economies and emerging markets since the mid 1990s. I focus on Germany, which provides a useful case study as the country produces frontier technology and underwent a large and sudden integration shock with Eastern Europe after the fall of the Iron Curtain. I document empirically strong patenting activity and employment reallocation to "innovative" establishments from 1995 to 2015, in combination with weak aggregate growth, stagnant wages, and rising inequality. The co-existence of weak growth and strong innovative effort is puzzling through the lens of standard theory, but a calibrated version of the model can match both features, precisely because of the trade-off between innovation and adoption.

The calibration predicts a quantitatively large cumulative drop in real wages of production workers of 17%, relative to the counterfactual balanced growth path in autarky. This effect is to be understood as a level difference from one steady state to another as the long run rate of technological change is fixed. In contrast, integration leads to cumulative wage gains for skilled labor of 11%, adding up to an increase in the skill premium of 33%. Consistent with the data, employment in the innovation sector expands and boosts the development of frontier technology. This expansion comes at the cost of a rising domestic adoption gap. Skilled labor and emerging market as a whole are benefiting, while production workers in rich countries experience wage stagnation. Aggregating up worker income within advanced economies implies a cumulative growth drag of 10%, i.e. a temporary growth slowdown.

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 implies strong diminishing returns in research activity. Under this assumption, reallocating skilled labor into innovation has only modest positive effects on the technological frontier. Yet, the negative productivity effect of weakened technology adoption can be so large that net productivity declines. Consistent with this prediction is that my model economy is inefficient. The decentral equilibrium features too little adoption in the closed economy. This inefficiency is amplified in the open economy.

In a final empirical exercise I leverage regional specialization in innovation vs. production across local labor markets within Germany, together with the fall of the Iron Curtain, to test the main predictions of the theory. The empirical exercise confirms that growth is biased towards innovative, high-income regions in Germany. These regions experience relatively higher growth in average real wages, skilled employment, and total population after market integration with Eastern Europe. In contrast, before 1994 in the pre-integration equilibrium, wage growth and skilled labor growth was fastest in laggard regions, consistent with adoption-driven growth. The empirical evidence thus corroborates the main point of the theory: market integration between advanced economies and emerging markets shaped the rate and distribution of economic growth across workers, regions, and countries. A model with an endogenous adoption gap and two types of labor is well-suited to capture these patterns in a parsimonious way.

Relationship to the literature: This paper relates to four different streams of the literature. First, the paper builds on the large literature on endogenous 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 addition, since innovation and adoption activity compete for skilled labor in general equilibrium, a crucial factor market rivalry emerges. This allows me to match empirical growth patterns 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. Moreover, technology adoption as the main driver of cross-country income differences is the central hypothesis of Parente and Prescott (1994). I extend this line of work by considering how adoption in emerging markets impacts the return to innovation in advanced ones. Recent work on directed technological change (Acemoglu, 2003) in combination with offshoring as in Acemoglu, Gancia, and Zilibotti (2015) is closely related and shares key predictions regarding the uneven effects of market integration. An important difference is that I introduce an endogenous technology adoption gap which allows for the coexistence of strong innovative activity and weak productivity growth. Moreover, the model highlights a novel innovation-adoption tradeoff in advanced economies, in particular when emerging markets are catching up.

Second, a number of papers have studied the recent productivity slowdown.[4] One strand of this literature focuses on the negative effect of declining population growth on productivity growth and business dynamics (Peters and Walsh, 2019; Jones, 2020; Hopenhayn, Neira, and Singhania, 2018; Engbom et al., 2019). While I agree that this is a central force, my theory highlights a new channel of weak technology adoption. This adoption margin provides a micro-foundation for empirical work that finds weak technology diffusion to be an important driver of slow productivity growth, see Andrews, Criscuolo, and Gal (2015) and Akcigit and Ates (2019). In addition, the model explains the rising share of innovative activity in the economy. Note that in the benchmark model of Jones (1995), falling

[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.

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 can resolve this tension. An alternative explanation for the productivity slowdown marries models of Schumpeterian growth (Aghion and Howitt, 1990; Grossman and Helpman, 1991b; Klette and Kortum, 2004) with biased technology shocks that favor large incumbents and suppress competition, see for instance De Ridder (2019), Rempel (2021), 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 analyzes how openness and comparative advantage shape sectoral specialization and economic growth, see Feenstra (2015) for a textbook introduction. On balance, the literature finds that market integration raises per capita growth (Rivera-Batiz and Romer (1991), Grossman and Helpman (1991a), Grossman and Helpman (2018), Sampson (2016), Hsieh, Klenow, and Nath (2019), Buera and Oberfield (2020), or Perla, Tonetti, and Waugh (2021)). My theory is consistent with this work in that integration is pro-innovation in advanced economies, but it may not always be pro-growth. For emerging markets, integration is always growth-enhancing as access to technology improves, which is consistent with the importance of technology adoption for cross-country income differences (Comin and Hobijn, 2010a; Comin and Mestieri, 2014) and the large literature on development and trade, see Irwin (2019) for a review.[6]

Most of the literature in international trade takes technology as given and studies the impact of trade on wage inequality and welfare, see Wood (1994), Leamer (1994), Feenstra and Hanson (1996) and more recently Adao et al. (2020).[7] Moreover, quantitative work has found offshoring and international trade to be relatively unimportant for rising wage inequality, see Arkolakis et al. (2018) or Galle, Rodríguez-Clare, and Yi (2017). In addition, increasing integration leads to welfare gains from trade (Costinot and Rodríguez-Clare, 2014) so weak growth remains puzzling. My approach abstracts away from import competition or offshoring and makes a novel point about missing domestic technology adoption. The key channel works through the reallocation of skilled labor from domestic adoption toward global innovation, which is a consequence of fast technology adoption in emerging markets.

Fourth, this paper relates to a large literature in labor economics that studies the wage inequality and the skill premium. Katz and Murphy (1992), Bound and Johnson (1992), and Krueger (1993) are seminal papers that focus on the recent rise in the skill premium in the US. Goldin and Katz (2010) study the evolution of inequality in the US over the long-run. In contrast to this literature, the skill

[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 in the long-run so that share of resources devoted to innovation asymptotes to zero.

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

[7] A related literature in international trade studies the impact of globalization on inequality in heterogenous firm models. See Helpman, Itskhoki, and Redding (2010), Liu and Trefler (2008), Sampson (2014), or Burstein and Vogel (2017).

premium not only matters as distributional accounting device in my theory 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 square with Katz and Murphy (1992)'s benchmark model of skill-biased technological change.[8] A related literature has focused on the task content of work and automation (Autor, Levy, and Murnane, 2003; Acemoglu and Restrepo, 2018a), and I incorporate this feature among other extensions into 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. The model is also related to Caselli (1999) and Beaudry, Doms, and Lewis (2010) which highlight the importance of skilled labor in adopting technology.[9]

The rest of the paper proceeds as follows: 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 empirical evidence to support the central mechanism. Section 6 concludes.

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.[10] Without loss of generality, I assume that the measure of capitalists is equal to L. Dynastic capitalists solve a forward-looking consumption-saving problem

[8] Note that in the benchmark model of skill-biased technological change, biased productivity growth towards skilled labor raises wages *for all workers* due to the strong complementarity between low-skilled and high-skilled workers, albeit for some more than others. This is inconsistent with observed wage stagnation for non-college workers, see Acemoglu and Autor (2011) for a discussion.

[9] The predictions of the theory are also consistent with recent work of Imbert et al. (2022) which finds that unskilled migration within China stalls TFP growth and innovation. In my model, unskilled immigration would raise the local skill premium, which would lead to weak technology adoption and receding innovation.

[10] 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 a state variable.

$$\max_{\{c,B\}} \int_0^\infty e^{-(\rho-g_L)t} \log c_t \, dt$$
$$\text{s.t. } \dot{B} = rB - C. \tag{1}$$

Total assets in the economy are denoted as B, which includes both physical capital and shares in firms, and I drop t subscripts for readability. Changes in total assets $\dot{B}$ 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

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

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

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

where the elasticity of substitution between differentiated intermediate goods equals σ. $L^{-\delta_Y}$ is an additional productivity shifter. Note that the market structure in the production sector is one of monopolistic competition, so population growth leads to additional productivity growth in the production sector, above and beyond research-driven technological change. I take this effect out by assuming $\delta_Y = \frac{1}{\sigma-1}$ but none of the qualitative insights hinge on this adjustment.[11]

The final good serves as the numeraire. It can be used for consumption or turned into physical capital one for one. Denoting aggregate consumption as $\tilde{C}$, i.e. the sum of capitalist and worker consumption, the usual law of motion of capital 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 compete monopolistically. The problem of an intermediate goods firm can be split into a static profit

[11] 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. A micro-foundation for this ad-hoc adjustment could be provided by adding a fixed factor, say land, into a constant-returns-to-scale aggregate production function.

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,

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

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.[12] 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 $\overline{x} = x_j \, \forall j \in \Omega_{A_i}$ and $A_i \overline{x} = \tilde{x}_i$.[13] Equal spending across capital goods is 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 each period.[14]

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

$$\max_{p_i, q_i, \{x_{ji}\}, l_i} \pi_i = p_i q_i - c(q_i)$$

$$s.t.$$

$$q_i = Y p_i^{-\sigma}$$

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

The solution concept is one of monopolistic competition where the firm takes factor prices p_x and w as well as aggregate variables as given. This static problem is well-known, and leads to a constant

[12] 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.

[13] In words, $\tilde{x}$ is the total quantity of capital goods on the firm level, $\overline{x}$ is the quantity of each individual capital good on the firm level, and so the total number of capital goods times the quantity of an individual capital good equals the total quantity of capital goods $A_i \overline{x} = \tilde{x}$, given symmetry. The aggregate quantity then follows by integrating over all firms, i.e. $\int_{\Omega_i} \tilde{x}_i di = X$.

[14] Wether firms own capital, or households own capital is not consequential, just like in the neoclassical growth model. Importantly, capital is combined with intellectual property to create a differentiated capital good as I detail below.

markup over marginal cost. The marginal cost is a weighted geometric average where the weights are given by the Cobb-Douglas output elasticities,

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

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 within each individual capital good variety. An increase in A_{iK}, for a fixed amount of capital spending $p_x\tilde{x}_i = \int p_{xj}x_{ij}dj$, makes the firm more productive and pushes down marginal cost.[15] The price of a differentiated intermediate good reads

$$p_i = \frac{\sigma}{\sigma-1}mc_i, \tag{7}$$

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

$$\begin{aligned} wl_i &= \tilde{r}_i\frac{\sigma-1}{\sigma}(1-\alpha) \\ p_x\tilde{x}_i &= \tilde{r}_i\frac{\sigma-1}{\sigma}\alpha, \end{aligned} \tag{8}$$

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

$$\pi_i^o = \frac{\tilde{r}_i}{\sigma}. \tag{9}$$

Dynamic adoption problem: The adoption of new capital goods is a costly process 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. 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

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

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. Adopting new capital varieties requires skilled labor, so adoption-driven productivity improvements only occur as long as the firm hires skilled labor $h_i > 0$. 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

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

discuss below. Note that (10) implies that the 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 $A_{iK} = A_F$.[16]

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,

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

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.[17]

Free entry: I close the production sector by assuming free entry after paying a fixed entry cost in terms of production labor.[18] I assume that entrants reach the know-how of incumbents as they enter, which captures a knowledge spillover within each country in anticipation of the open economy setting later on. The smaller the fixed cost, the larger is this spillover. This spillover implies that I can drop the i subscript since all incumbents 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. The benefit of the symmetric firm model is that it simplifies the innovator problem since innovators only need to keep track of one adoption gap instead of an entire distribution, which would be the case in a model with heterogeneous firms. Abstracting away from this layer of heterogeneity allows me to develop a

[16] 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.

[17] 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.

[18] 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.

tractable theory of cross-country and skill-type inequality. An alternative heterogeneous firm setting where entry features an imperfect knowledge spillover is considered in the appendix.[19] In the appendix A.6 I consider such a setting.[20]

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.[21] 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 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

[19] See Luttmer (2007), Lucas (2009b), Sampson (2016) and Buera and Oberfield (2020) for models that focus on knowledge spillovers.

[20] In this setting 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. 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.

[21] A justification for this congestion force is the possibility of useless duplication, i.e. two researchers coming up with the same idea.

profits

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

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 time 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 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 problem. From the point of view of the patent owner, this gives rise to a static pricing problem

$$\begin{aligned}
\max_{p_{xj}, X_j} \pi_{Ij} &= p_{xj} X_j - c\left(X_j\right) \\
s.t.& \\
X_j &= R_X \left(\frac{p_{xj}}{P_x}\right)^{-\frac{1}{1-\alpha}} \\
c\left(X_j\right) &= X_j \left(r + \delta_k\right)
\end{aligned}$$

where R_X is aggregate spending on capital goods and $\int X_j dj = A\overline{x}M = X$ is aggregate demand for capital goods.[22] The cost function $c(.)$ is linear, and P_x is an aggregate capital goods price index. 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 price of any capital good equals

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

which is a constant markup over marginal cost.[23] After solving for the endogenous price index and aggregating over all intermediate goods firms, the flow profits are equal to a constant share of total revenue divided by the total measure of active ideas, which in turn is proportional to the wage bill in

[22]Note that $X_j = x_{ji}M = \overline{x}M$ due to symmetry, and integrating over all adopted capital goods varieties delivers the demand for capital in production, $A\overline{x}M$. In equilibrium, this needs to match physical capital accumulation on the household side, $X = K$.

[23]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. $y = \left(\frac{\left(\int x^\rho dj\right)^{\frac{1}{\rho}}}{\alpha}\right)^{\alpha} \left(\frac{l}{1-\alpha}\right)^{1-\alpha}$ so that the markup is related to ρ while the capital share is still a function of α.

the economy due to Cobb-Douglas production[24]

$$\pi_I \quad = \quad \frac{\alpha L_P w}{A} \tag{19}$$

Now I turn to the endogenous waiting time. I partition the set of capital goods Ω_{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.[25] 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.[26] 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 derived 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.[27] Second, a flow $A_t(\delta_I + g_A)\,dt$ is adopted every instant, which could be negative or positive.[28] 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. 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

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

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 $\phi <$

[24]Formally, $\pi_I \;=\; \int_i r_i^x \left[\frac{p_x^{-\frac{1}{1-\alpha}}}{P_x^{1-\frac{1}{1-\alpha}}}\right]\left[(p_x) - (r + \delta_k)\right] di \;=\; R_X \int_i \left[\frac{p_x^{-\frac{1}{1-\alpha}}}{P_x^{1-\frac{1}{1-\alpha}}}\right]\left[p_x(1-\alpha)\right] di \;=\; \frac{R_X}{A}(1-\alpha) \;=\; \frac{\alpha}{1-\alpha}\frac{L_P w}{A}(1-\alpha) = \frac{\alpha L_P w}{A}$.

[25]This assumption is useful to generate churn among innovators in the absence of population growth but is otherwise inconsequential.

[26]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.

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

[28]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.

1, consist of a sequence of prices $\{w_t, w_{Ht}, r_t, p_{xt}, p_{it}, V_t, V_{It}\}$ and allocations $\{L_{Pt}, L_{Et}, H_{Dt}, H_{Ft}, X_t, K_t, M_t, A_t, A_{Ft}, C_t\}$ 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.
- There is 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 using production labor as normalizing factor 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 $\Gamma = -\log z$. I next derive key results of the model, while a detailed derivation can be found in the appendix.

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

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

$$\text{s.t.} \tag{22}$$

$$\dot{z} = \zeta z^\theta h^\beta - \left(g_F + \delta_I\right)z, \tag{23}$$

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

$$\left(\partial_z v\right)\beta\zeta z^\theta h^{\beta-1} \quad = \quad s. \tag{24}$$

Equation (24) 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,[29] leading to the following law of motion for skilled labor growth for an individual firm

$$\frac{\dot{h}}{h} = \frac{1}{1-\beta}\left\{\underbrace{\rho+\delta_{ex}}_{\text{effective discounting}} + \underbrace{(1-\theta)(g_F+\delta_I)}_{\text{effective depreciation}} - \underbrace{\left\{\frac{\beta z^{\theta}\zeta h^{\beta-1}}{s}\left[\frac{\pi_t}{w}\frac{(1-\alpha)(\sigma-1)}{z}\right]+\frac{\dot{s}}{s}\right\}}_{\text{marginal benefit of extra unit of skilled labor}}\right\}. \tag{25}$$

Equation (25) 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 investment literature, firms make forward-looking decisions, which depend on the current stock z (capital in the investment literature) and lead to an optimal level of investment. As long as $\beta < 1$, investment in the form of hiring skilled labor to adopt new technology (h) runs into diminishing returns. The curvature captured in β as well as the advantage of backwardness $(1-\theta)$ shape the speed of adjustment, 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

$$h = \frac{1}{s}\frac{\beta(1-\alpha)(\sigma-1)(g_F+\delta_I)}{\rho+\delta_{ex}+(1-\theta)(g_F+\delta_I)}\left[\frac{\pi}{w}\right]. \tag{26}$$

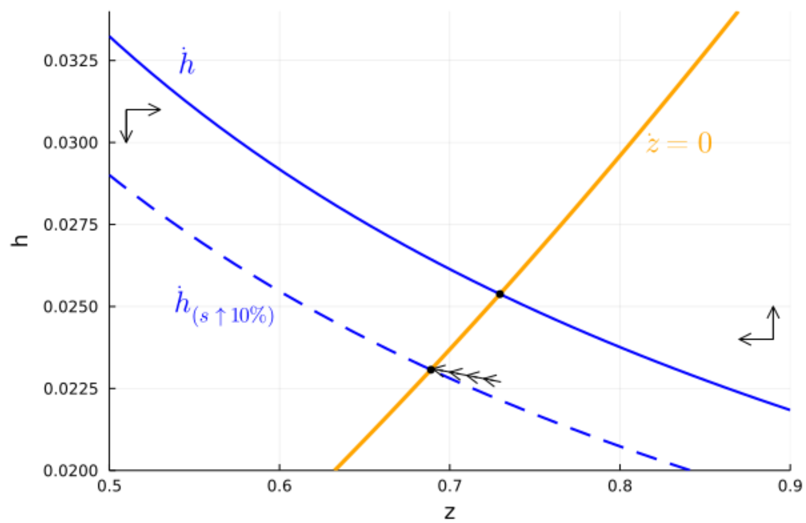
The demand for skilled labor is proportional to normalized profits, and falling in the skill premium. Moreover, it is positively related to the sensitivity of profits with respect to productivity, $(\sigma-1)(1-\alpha)$, as a large demand elasticity will make firms benefit more from a technological improvements, ceteris paribus.[30]

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

[30] 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.

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 (25) 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 2. 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.

Figure 2. Phase Diagram



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

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

Following the steps in Peters and Walsh (2019), one can show that the normalized value function, $v = \frac{V}{w}$, equals

$$v = \frac{\frac{\pi}{w} - sh}{r + \delta_{ex} - g_w} \tag{27}$$

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.[31]

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 (27) and (26) the normalized operating profits on the balanced growth path are pinned down

$$f_e\left(\rho+\delta_{ex}\right)\kappa_2 = \frac{\pi}{w}, \tag{28}$$

where I used $r = \rho + g_w$. Equation (28) is directly related to the flow cost of entry, $f_e(\rho+\delta_{ex})$, but it features an extra term $\kappa_2 > 1$,[32] 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

$$f_e\left(\rho+\delta_{ex}\right)\kappa_2 \quad = \quad \frac{l_P}{m}\frac{1}{(1-\alpha)(\sigma-1)}. \tag{29}$$

Together with the normalized firm entry resource constraint

$$\dot{m} \quad = \quad \frac{l_E}{f_e} - \left(\delta_{ex}+g_L\right)m, \tag{30}$$

the steady state normalized measure of firms reads

$$m \quad = \quad \frac{1}{f_e\left[(\rho+\delta_{ex})(1-\alpha)(\sigma-1)\kappa_2+g_L+\delta_{ex}\right]}. \tag{31}$$

Note that out of steady state, the endogenous firm measure is not constant.[33]

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 (26). Taking logs leads to

$$\log z \quad = \quad -\frac{\beta}{1-\theta}\log s + \frac{1}{1-\theta}\log\left(\frac{\zeta}{(g_F+\delta_I)}\left(\frac{\pi}{w}\kappa_1\right)^{\beta}\right). \tag{32}$$

Expression (32) 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, diminishing returns in adoption (β) together with the advantage of backwardness $(1-\theta)$

[31] 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.

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

[33] This is an implication of (27) 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.

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 quickly. In addition, the effect is strong when θ is large, which parameterizes how important current knowledge is to adopt new ideas.

In this model, 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. A rising skill premium simply makes adoption more expansive, and thus hurts technology adoption. The strength of this effect directly depends on the ratio $\frac{\beta}{1-\theta}$.

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

$$V_I \quad = \quad \int_{t+\tau}^{\infty} \exp\left(-\int_t^u (r_x+\delta_I)\,dx\right)\pi_{Iu}du \tag{33}$$

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

$$V_I = \exp\left(-\int_t^{t+\tau}[r_u+\delta_I]\,du\right)(1+\tau')\frac{\pi_{I,t+\tau}}{r+\delta_I-g_{w_H}-(1-\lambda)\,g_{H_F}+\phi g_F} \tag{34}$$

which holds on and off the balanced growth path.[34] 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.

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* $(g_A+\delta_I)$

$$\tau \quad = \quad -\frac{\log z}{g_A+\delta_I}. \tag{35}$$

Proof in A.2.3.

Intuitively, equation (35) 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. This is similar to the well-known relationship between distance, speed, and travel time in physics. Note, however, that the adoption gap is endogenous. For example, as z approaches unity, the waiting time shrinks to zero and vice versa, the waiting time shoots off to infinity when $z \to 0$.

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

In the steady state, the present value of an innovation thus simplifies to[35]

$$V_I \quad = \quad \frac{1}{\tilde{\rho}+g_F+\delta_I}\left(\frac{\alpha L_P w}{A_F}\right) z^{\frac{\tilde{\rho}}{g_A+\delta_I}} \tag{36}$$

where $\widetilde{\rho} := \rho - g_L > 0$ is the effective discount factor of the dynastic household and I substituted out τ using (35). The 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 (36) 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,

$$\frac{1}{\gamma} \quad = \quad \frac{1}{s}\frac{\alpha l_P}{\tilde{\rho}+g_F+\delta_I}\left(\frac{h_F^{\lambda-1}}{a_F}\right) z^{\frac{\tilde{\rho}}{g_A+\delta_I}}\,. \tag{37}$$

Two important economic mechanisms are captured in (37). 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

$$\begin{aligned} h_D \quad &= \quad m h_i \\ &= \quad \frac{1}{s}\Lambda^D \end{aligned} \tag{38}$$

where Λ^D collects elements that are constants in the steady state.[36] Using a normalized version of the law of motion of ideas (14), I get

$$\frac{\gamma h_F^{\lambda}}{(g_F+\delta_I)} \quad = \quad a_F. \tag{39}$$

Combining (37) with (39) leads to the research sector's normalized demand for skilled labor

$$\begin{aligned} h_F &= \frac{1}{s}\left(\frac{g_F+\delta_I}{\tilde{\rho}+g_F+\delta_I}\right)\alpha l_P\,(z)^{\frac{\tilde{\rho}}{\delta_I+g_F}} \\ &= \frac{1}{s}\,(z)^{\frac{\tilde{\rho}}{\delta_I+g_F}}\,\Lambda^F. \end{aligned} \tag{40}$$

[35]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 $\exp\left(-\int_t^{t+\tau}[r_u+\delta_I]\,du\right)\frac{L_{P,t+\tau}}{L_{Pt}}L_{Pt} = \left(-\int_t^{t+\tau}[r_u+\delta_I-g_{L_P}]\,du\right)L_{Pt}$. Moreover, I use the fact that $A = A_F z$ to substitute out A.

[36]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 state.

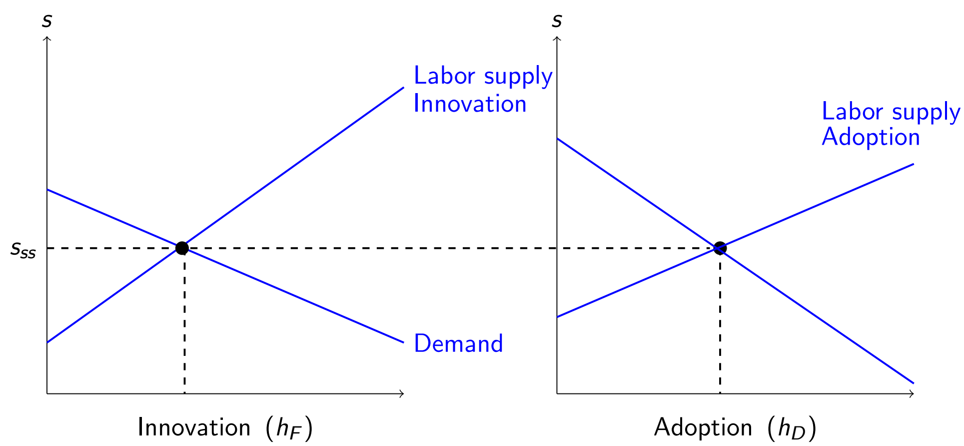
Adding up (38) and (40) and imposing market clearing, I obtain the following equation that implicitly defines the relative price of skill

$$\left\{ \frac{1}{s} \left(z\right)^{\frac{\tilde{\rho}}{\delta_I + g_F}} \Lambda^F + \frac{1}{s} \Lambda^D \right\} \quad = h_{tot} \tag{41}$$

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 plot in figure 3 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.

Figure 3. Market Clearing for Skilled Labor



Aggregation: This economy behaves similar to a neoclassical economy where a rising real wage and a constant real rate characterize the balanced growth path. Unlike the neoclassical model, long run growth is endogenous and depends on the interaction of population growth and innovation.

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* $g_K = g_L + g_A$*. Moreover, the ratio of labor devoted to innovation relative to adoption,* $\frac{H_F}{H_D}$*, is constant and so is skilled labor demand for an individual firm,* $h_{it} = h_i$*.*

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,

$$Y = \left(\frac{K}{\alpha}\right)^{\alpha} \left(\frac{z A_F L^P}{1-\alpha}\right)^{1-\alpha}. \tag{42}$$

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

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

A constant capital-effective labor ratio on the balanced growth path follows

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

where $k := \frac{K}{z A_F L_P}$ and $\Lambda_\alpha = \left(\frac{1}{\alpha}\right)^{\alpha} \left(\frac{1}{1-\alpha}\right)^{1-\alpha}$. The real wage for production and high skilled workers reads

$$\begin{aligned} w &= (1-\alpha)\frac{\sigma-1}{\sigma}\Lambda_\alpha z A_F k_{ss}^{\alpha} \\ w_H &= sw. \end{aligned} \tag{43}$$

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.3. 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. Crucially, it also allows for the possibility of a growth slowdown in the face of rising innovative effort and frontier technology growth, a case where 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 (37), the partial equilibrium elasticity of the total measure of ideas A_F with respect to the adoption gap in the steady state equals[37]

$$\frac{\partial \log A_F}{\partial \log z} = \frac{\lambda}{1-\phi} \frac{\tilde{\rho}}{g_A + \delta_I}. \tag{44}$$

As production firms raise their adoption effort and push up z, the present discounted value of an innovation increases due to a reduction in the 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 are rivalrous as both activities compete 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. To see this, I consider how innovation and adoption respond to different fundamental shocks in the model, showing that the two activities move in lock-step, even in general equilibrium.

First, consider an increase in γ, the research productivity. From equation (39), 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 (38) and (40)

$$\frac{H_F}{H_D} = \left(\frac{g_F+\delta_I}{\tilde{\rho}+g_F+\delta_I}\right) \frac{\alpha(\sigma-1)(1-\alpha)}{\kappa_1} (z)^{\frac{\tilde{\rho}}{\delta_I+g_A}} . \tag{45}$$

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

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

declines.

Finally, suppose that the relative supply of skilled labor shrinks. This is a reduced form way to consider a changing task content of work that makes skill effectively more scarce, see 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 (45). Since z is falling due to a rising skill premium, innovation is hurt twice. First, a direct input cost effect hurts innovation as skilled labor has become more expensive, and second, a rising adoption gap further hurts innovation by pushing down the net present value of an innovation. Again, innovation and adoption move in the same direction, and innovation responds even stronger than adoption in the face of a negative skill supply shock.

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 focus on the implications of my model in a simple two-country open economy setting.[38] Countries have different fundamental research productivity γ and different relative skill endowments h_{tot} but are otherwise identical. In particular, preferences and non-research related 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.[39] 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,[40] and balanced trade shuts down inter-temporal trade motives.[41] Importantly, even if a domestic capital good is produced abroad for foreign use, the domestic inventor still receives a royalty. The model of trade will thus be one where emerging markets 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

[38] Most of the results generalize to a multi-country setting as I point out along the way.

[39] 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 large 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.

[40] 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.

[41] See Obstfeld and Rogoff (1995)'s inter-temporal approach to the current account, and Aristizabal-Ramirez, Leahy, and Tesar (2022) for a recent contribution.

foreign variables of the emerging market, while the advanced economy represents the home 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 advanced economy and the emerging market, $\frac{A^*}{A} = \frac{z^* A_F^W}{z A_F^W} = \frac{z^*}{z}$, which directly pins down the relative wages of production workers

$$\frac{w^*}{w} = \frac{z^*}{z}. \tag{46}$$

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)). This feature can be easily generalized into a multi-country setting where a country's adoption effort measured in terms of skill-to-production labor ratios pin down an economy's position on the global productivity distribution. Differences in specialization in innovation complicate the mapping from adoption to GDP slightly, and lead to country-specific skill 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

$$z_c \propto h_c^{\frac{\beta}{1-\theta}}. \tag{47}$$

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 that long-run growth requires idea-based technological change (Jones, 2005).[42]

[42] 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 pick parameter values that the real income of skilled labor is highest in places where it is most scarce, despite weak technology adoption. That is to say, skill scarcity dominates the negative effect of weak adoption within each country. 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 while real income of low income households is relatively low compared to other advanced economies.

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: I assume for simplicity that the knowledge spillover A_F^{ϕ} is global,[43] which leads to the following law of motion of ideas in the advanced economy.

$$\dot{A}_F = \frac{\left(A_F^W\right)^{\phi} H_F}{f_R} - \delta_I A_F. \tag{48}$$

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

$$\frac{h_F}{h_F^*} = \left(\frac{\frac{\gamma}{w_H}}{\frac{\gamma^*}{w_H^*}}\right)^{\frac{1}{1-\lambda}} \tag{49}$$

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 (48), it follows that[44]

$$\left(\frac{\chi}{\chi^*}\right) = \frac{\gamma}{\gamma^*}\left(\frac{h_F}{h_F^*}\right)^{\lambda}. \tag{50}$$

[43]See Grossman and Helpman (1991a) for an in-depth discussion of this issue. Global knowledge spillovers seem a natural assumption in a model of long-run growth.

[44]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 emerging market economy and compute the ratio.

Combining this expression with (49) and noting that $\frac{w_H}{w_H^*} = \frac{s}{s^*}\frac{z}{z^*}$ leads to

$$\left(\frac{\chi}{1-\chi}\right) = \left(\frac{\gamma}{\gamma^*}\right)^{\frac{1}{1-\lambda}} \left(\frac{s}{s^*}\frac{z}{z^*}\right)^{-\frac{\lambda}{1-\lambda}} . \tag{51}$$

Equation (51) 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* skilled wage ratio $\frac{w_H}{w_H^*}$ (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.[45]

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 advanced economies and emerging markets reads

$$\begin{aligned} \left\{\frac{\chi}{z}\Lambda_{FO}\left((z)^{\frac{\tilde{\rho}}{g_A+\delta_I}+1} + (z^*)^{\frac{\tilde{\rho}}{g_A+\delta_I}+1}\right) + \Lambda_D\right\} &= sh_{tot} \\ \left\{\frac{\chi^*}{z^*}\Lambda_{FO}\left((z)^{\frac{\tilde{\rho}}{g_A+\delta_I}+1} + (z^*)^{\frac{\tilde{\rho}}{g_A+\delta_I}+1}\right) + \Lambda_D^*\right\} &= s^*h_{tot}^* \end{aligned} \tag{52}$$

where $\Lambda_{FO} = l_P\alpha\frac{g_F+\delta_I}{\tilde{\rho}+g_F+\delta_I}$. Note that both χ and z are functions of s and s^* so an equilibrium involves a set of skill premia that solve (52).

Balanced trade implies

$$\underbrace{\{Y+Y^*\}\frac{\sigma-1}{\sigma}\alpha(1-\alpha)(\chi-\chi^*)}_{\text{Innovator Profits/Royalty}} = \underbrace{Y^* - \tilde{C}^* - I^*}_{\text{Final good exports}} \tag{53}$$

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 (53).[46] 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.

To see this, in proposition 4 I consider what happens if the research productivity γ of the home economy increases.

Proposition 4. *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*

[45] 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.

[46] 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 spending on capital goods reads $P_xX = \frac{\sigma-1}{\sigma}\alpha Y$. 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.

the foreign economy falls. Proof see appendix A.5.

Proposition 4 is in contrast to the closed economy result from the previous section where 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 leads to a larger share of world research 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. In particular, it ensures that skilled labor cannot be worse off in real terms after an increase in the skill premium.[47] This inequality is also respected when matching data moments in the quantitative application.

Real Income in the Open Economy: The setting leads to a simple formula to compute the real wage effects of market integration for each skill group, similar to Arkolakis, Costinot, and Rodríguez-Clare (2012). Specifically, the real wage of production workers in the open economy relative to the real wage in autarky (closed) is summarized by the following sufficient statistic

$$\frac{w^{open}}{w^{closed}} = \underbrace{\left(\frac{h_F^{open}}{h_F^{closed}}\right)^{\frac{\lambda}{1-\phi}} \left(\frac{1}{\chi}\right)^{\frac{1}{1-\phi}}}_{\text{Gains from frontier innovation}} \underbrace{\left(\frac{s^{open}}{s^{closed}}\right)^{-\frac{\beta}{1-\theta}}}_{\text{Loss from missing adoption}}$$

and

$$\frac{w_H^{open}}{w_H^{closed}} = \underbrace{\left(\frac{h_F^{open}}{h_F^{closed}}\right)^{\frac{\lambda}{1-\phi}} \left(\frac{1}{\chi}\right)^{\frac{1}{1-\phi}}}_{\text{Gains from frontier innovation}} \underbrace{\left(\frac{s^{open}}{s^{closed}}\right)^{\frac{1-(\beta+\theta)}{1-\theta}}}_{\text{Gains from rising skill premium}}.$$

The benefits from market integration are captured in i) increasing innovative effort in the advanced economy $\left(\frac{h_F^{open}}{h_F^{closed}}\right)^{\frac{\lambda}{1-\phi}}$, and ii) gains from specialization in research $\left(\frac{1}{\chi}\right)^{\frac{1}{1-\phi}}$ that depend on a constant scale elasticity $\frac{1}{1-\phi}$ as well as the idea trade share χ (which could be expressed as import share $\chi = 1-\chi^*$) . The novel feature is the endogenous adoption margin which shows up in the skill price ratio $\left(\frac{s^{open}}{s^{closed}}\right)^{-\frac{\beta}{1-\theta}}$. While an increase in the relative price of skill raises the real wage of high skilled workers since I assume $1>\beta+\theta$, it clearly hurts production workers. The reason is that a rising skill premium leads to less domestic technology adoption. While weak adoption in principal hurts skilled labor as well, the negative effect of weak adoption is dominated by direct wage gains due to a rising skill premium, i.e. $-\frac{\beta}{1-\theta}+1>0$. The framework allows for a richer response of market integration on growth, whereby gains from rising innovative effort are counteracted by receding technology adoption.

[47] In principal, adoption could fall to an extent that skilled labor loses even though their relative price increases. Note that the inequality is a sufficient condition to ensure skilled labor sees real wage improvements after an increase in the skill premia. There are many configurations where $\beta+\theta>1$ and skilled labor still gains in real terms after an increase in the skill premium, but this prediction now depends on other parameters as well.

This latter effect is parsimoniously captured in changes in the skill premium raised by a constant elasticity $\frac{\beta}{1-\theta}$.

It is worth noting that integration between symmetric countries delivers the standard variety gains from trade as in Krugman (1980) without any negative distributional effects, and hence leaves equilibrium adoption unchanged as well.

Proposition 5. *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 more capital goods available at a factor $2^{\frac{1}{1-\phi}}$,[48] 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.[49] 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.[50] Heterogeneity across countries in terms of their fundamental research productivity or skill endowments changes this result.

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

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 as long as the advanced economy has a lower skill premium than the emerging market. Clearly, the skill premium is endogenous and the central assumption is that $h_{tot} > h^*_{tot}$ is sufficiently different so that even after the advanced economy specializes into research activity, which pushes up the skill premium, it remains true that $s < s^*$. Given the effect of the skill premium on factory floor productivity, this is an inequality that any reasonable equilibrium should obey, i.e. if that was not the case, the emerging market would adopt more technology than the advanced economy. When λ is close to unity, relatively small differences in the skill premium give rise to large differences in cross-country specialization in innovation. For similar

[48] If on the other hand N new countries join the world economy and hav the same skill endowment and research productivity, the scale effects will be of the order $N^{\frac{1}{1-\phi}}$.

[49] 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.

[50] This result requires similar research productivity (No Ricardian comparative advantage) and similar factor ratios (No neoclassical factor bias) so that trade integration does not change the returns to any factor. Yet, unlike Ricardian or neoclassical trade models, trade integration still generates gains due to increasing returns in the research sector, i.e. intra-industry gains from trade as in Krugman (1980).

countries, this leads to factor-price equalization where a country specializes in innovation up until the skill premia are equal. For very different countries, a corner solution emerges where all innovation takes place in the advanced economy and $s < s^*$ still holds true in the integrated equilibrium. I call this case *asymmetric integration.* While the emerging market 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.[51]

To see how market integration raises the skill premium in the advanced economy, consider the modified present discounted value of an innovation in the open economy. 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

$$V_I = \underbrace{\left(\frac{\alpha}{\tilde{\rho} + g_F + \delta_I}\right) \frac{L_P w}{A^F} z^{\frac{\tilde{\rho}}{g_A + \delta_I}}}_{\text{same as closed economy}} \left\{ 1 + \underbrace{\frac{L_P^* w^*}{L_P w} \left(\frac{z^*}{z}\right)^{\frac{\tilde{\rho}}{g_A + \delta_I}}}_{\text{additional market size effect}} \right\}. \tag{54}$$

Equation (54) reveals that the strength of the idea demand shock depends on i) the adoption gap $(z^*)^{\frac{\tilde{\rho}}{g_A + \delta_I}}$ in the emerging market, as well as ii) GDP summarized in $L_P^* w^*$ *relative* to variables in the advanced economy. 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 (52) when using $\frac{w^*}{w} = \frac{z^*}{z}$. The model can easily accommodate countries of different size as equation (54) shows, and a larger foreign labor force exerts more pull on innovation in the advanced economy.

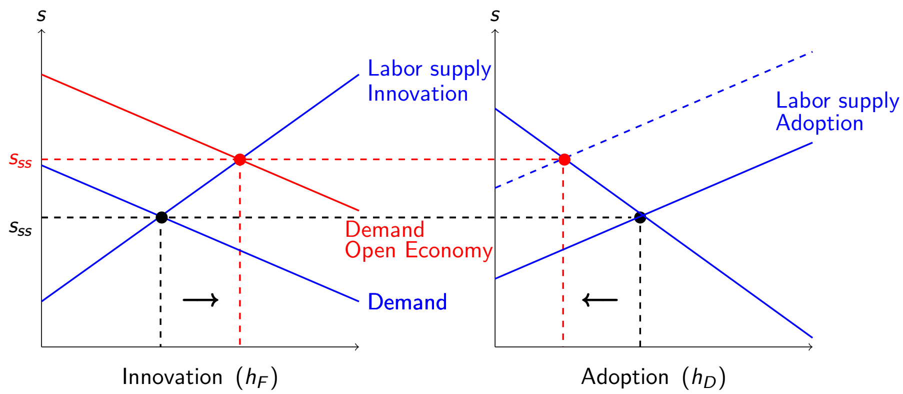
Note that market integration directly increases the market size of innovators, which raises profits that are arbitraged away by increasing entry into innovation. Importantly, convergence in the emerging market further raises the returns to innovation. Note that a rising wage rate ($w^* \uparrow$) and 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 to 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

[51] 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. For more recent years, this assumption may be less appropriate as China is starting to contribute to the global technological frontier. 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.

that is capturing the benefits from market integration in advanced economies is skilled labor. Figure 4 summarizes the main argument of this paper in a simple supply-demand plot.

Figure 4. Market Clearing for Skilled Labor in Open Economy



While innovation and adoption were characterized by a strong complementarity in the closed economy, factor market rivalry and competition for skilled labor dominates the relationship between innovation and adoption within advanced economies in the integrated equilibrium. Note that innovation is still responding to adoption, but it is responding to *foreign adoption*.

To summarize, in the advanced economy innovation takes off, adoption recedes, and inequality increase after market 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 more they adopt, the stronger is the pull on innovation in the advanced economy. Growth abroad and inequality in the advanced economy are thus linked. These are qualitative insights that hold in general in this type of model given asymmetric integration.

In order to compute the aggregate growth effects and the exact increase in the skill premium, I have to pin down relevant parameters and simulate the model. Before I turn to this quantitative application, I conclude the theoretical section by considering extensions and I contrast the theory to recent work on skill-biased technological change and the effect of declining population growth on productivity.

Emerging Market in Autarky: One issue that arises is how to model the emerging market in the closed economy. I specify the closed economy as follows. First, I introduce a technology adoption friction similar to Parente and Prescott (1994). Specifically, suppose that there is a market-share reallocation friction parameterized by some $\mu < 1$ so that the marginal product of technology adoption is suppressed relative to the market equilibrium,

$$\mu\beta\left(\sigma-1\right)\left(1-\alpha\right)\propto V_{A}^{*}<V_{A}$$

which in turn leads to depressed demand for skilled labor. The parameter μ stands in for the many frictions that prevent firms from gaining new market share in the heavily restricted economic environment that was common in the Soviet Union. Importantly, whenever technology adoption is a skill-intensive activity, such frictions depress the skill premium as demand for skilled labor is artificially low. Algebraically, this shows up in a lower $\Lambda^{D,*}$. Market reforms lead to frictionless technology adoption, which means that both productivity and inequality should go up in the emerging market. Productivity increases due to the standard channel of technology adoption. In addition, since adoption is a skill-intensive activity, the demand for skilled labor increases as returns to adoption rise, which pushes up the skill premium in the emerging market.

For completeness, I assume that in autarky innovators in emerging markets copy ideas from advanced economies without compensating the original inventor using the law of motion

$$\dot{A}_{F}^{closed,*}=\gamma^{copy}A_{F}^{\phi}\left(H_{F}^{*}\right)-\delta_{I}A_{F}.$$

Copying is easier than invention and I assume $\gamma^{copy}>\gamma^{*}$. In an integrated post-reform market equilibrium, copying ideas is not tolerated and stolen technology loses all its value. I need this type of technology stealing to avoid a scenario where the emerging market is counterfactually poor in the closed economy, which would be the case in my calibration later on as they are skill scarce.

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 consider 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.[52] I abstract away from this secular trend to show what my approach can contribute to this established 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

[52] 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 20 in the appendix.

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

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.[53] 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, however, is a more promising approach 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).[54] The model remains mostly unchanged except for an additional term $\tilde{\Lambda}_{\tilde{\beta}}$ in the labor market clearing condition,

$$\frac{1}{s}\left(\tilde{\Lambda}_F z^{\frac{\tilde{\rho}}{g_F+\delta_I}} + \tilde{\Lambda}_D + \tilde{\Lambda}_{\tilde{\beta}}\right) = h_{tot}. \tag{55}$$

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

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. The effect of globalization on the returns to innovation will resolve this tension and help make sense of rising innovative activity in advanced economies.

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

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

[54]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.[55] 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

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

which creates a weaker response of the skill premium on technology adoption since $\mu < 1$. Note, however, that there would be an additional negative effect on firm entry, i.e. $\frac{\partial \log m}{\partial \log s} < 0$. Since entry costs partially depend on high-skilled wages, a rising skill premium raises entry cost. The free entry condition then implies a relatively smaller number of firms in equilibrium so that rising profits make up for higher entry costs. I abstract away from this margin for simplicity. However, missing firm entry and slowing firm dynamics have been documented by Decker et al. (2017), Decker et al. (2020), or Karahan, Pugsley, and Şahin (2019), and are likely to be related to weak aggregate growth. 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.[56]

In appendix 3.1, I consider 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.

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

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

4 Quantification

4.1 Calibration of the Model

To calibrate the model I need to pin down a number of parameters. I focus on the German economy as a stand-in for advanced economies more broadly. This is useful because Germany produces frontier technology, experiences a major market integration shock with Eastern Europe in the mid 90s, and offers rich worker and establishment data to test key predictions of the model.

Growth $\left\{g_L, \phi, \frac{H}{L}\right\}$: 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 build on recent work of Bloom et al. (2020) which find that ideas are getting harder to find in the sense that ϕ is quite negative. I set ϕ equal to -1 which is a lower bound on the negative dynamic externality they measure.[57] 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, based on the Barro and Lee (2013) 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 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 pick the skill-intensity $\frac{H}{L}$ to be .15, which is slightly above the relative share of the population over the period 1980 – 2015 that has a college education, based on data from Barro and Lee (2013),[58] and very similar to the high skill-low-skill ratio in Acemoglu et al. (2018) of .16.

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

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

so that given $g_F = 1\%$ and $\delta_I = 4\%$, a reasonable estimate for θ is thus 0.6 which ensures that

[57] The negative results of integration on growth and inequality are amplified as the negative ϕ increases in absolute value.

[58] When only requiring some tertiary education, the ratio goes up to 20%, which is still low compared to other advanced economies, partly due to the apprenticeship system in the labor market. The ratio in 1980 is substantially lower than in 2015, and 15% is an average.

$\hat{\beta}_B = -.02$. This leads to slow convergence dynamics relative to a neoclassical model.[59] 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.

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

$$z_c = \left(\frac{\zeta h_c^{\beta}}{g_F + \delta_I}\right)^{\frac{1}{1-\theta}} \tag{56}$$

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 $\{\theta, \beta\}$ translate this initial distribution into observed cross country inequality. A small β leads to small cross country income differences. Taking logs of (56) and adding a measurement error u allows me to back out β by running the following regression

$$\log z_{ct} = \alpha + \delta_t + \frac{\beta}{1-\theta} \log h_{ct} + u_{ct}. \tag{57}$$

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.[60]

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.[61] I obtain a coefficient (robust standard error) of .9 (.06) with an R-squared of 65%, as can be seen in figure 5. 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.

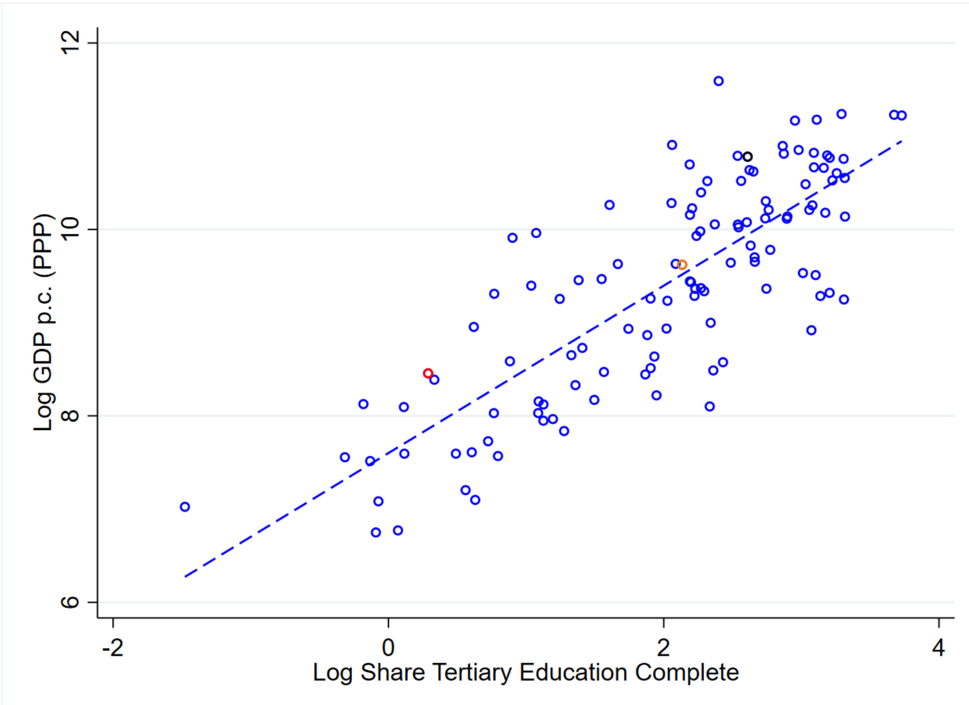
Moreover, I set α to be equal to .5, which is the capital share in production. Once one takes

[59] 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.

[60] 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.

[61] 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.

Figure 5. 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.

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%.[62] In the standard neoclassical model, this parameter shapes the convergence dynamics. In the model at hand, long-run convergence is instead a function of θ which leads to slow convergence due to a (weak) advantage of backwardness. An important difference to the neoclassical model is that convergence here happens in terms of TFP, not just capital-labor ratios.[63]

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.[64]

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

[63] The dominant variable that shapes the speed of convergence is pinned down by the smallest eigenvalue of the linearized system, which in my case would indeed be related to technology adoption. See Buera et al. (2020) for recent work on this issue in long-run models of structural change.

[64] The reader may wonder whether this leads to a very large profit share, compared for instance to the estimate of Basu and Fernald (1997). This is not the case since 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 parameters, 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 as to avoid "too much" adoption.

Firm Entry, Exit, Fixed Cost, Constant in Adoption $\{\delta_I, \delta_{ex}, f_e, f_R, \zeta\}$:

I classify establishment into production and research sector, and measure average employment-weighted establishment age in the micro data for each sector separately. On average, firms are roughly 25 years old in both sectors, which leads to a Poisson arrival rate of death of .04. I thus implicitly assume that an idea is equal to an establishment.[65] I set the fixed entry cost into research and production equal to unity. While the size of the fixed cost does not matter much in the research sector since the allocation of skilled labor is independent of the fixed cost of entry, the size of the fixed cost in the production sector matters in the following way. If the fixed cost is very large, the normalized equilibrium measure of firms is low. For a fixed amount of skilled labor devoted to adoption, every production firm individual has more skilled labor devoted to technology adoption since $h_i = \frac{h_D}{m}$. In the limiting case where $m \to 0$, the firm will hit a boundary so that $z = 1$ and the first order condition does not apply. One can avoid this, even for very large fixed costs, by letting ζ shrink as f_e grows. The reader thus should think of setting ζ and f_e jointly. I pick a combination where the firm choses an interior solution that leads to a reasonable skill premium in the closed economy. Given $f_e = 1$ and aiming for a skill premium of 2 in 1994, I set ζ equal to .23.

Foreign Economy $\{z^*, L^*, s_0^*\}$**:** 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 abroad. 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 the US and Japan in the 2000s. I proceed by assuming that market integration happens between two equal sized countries in the sense that $L = L^*$ and the foreign relative technology level shifts from $z_{1995}^* = .2$ to $z_{2015}^* = .4$. This development story is consistent with a relative skill share of .04 in the emerging market. I pick the parameters μ (initial adoption friction) and γ^{copy} (imitation of ideas) so that they are consistent with $z_{1995}^* = .2$ and an autarky skill-premium of 1.7.[66]

4.2 Quantitative Results

Steady State Results: First, I compare steady state differences. The initial equilibrium is in autarky while the new steady state is an integrated equilibrium as discussed in the previous section. I then compute the percent difference between a counterfactual autarky wage and the wage in the long run steady state in the integrated equilibrium. In this semi-endogenous growth model, the long-run

[65] This is not perfect since some firms have multiple ideas, and some ideas are produced in multiple establishments but it is a first step to pin down this parameter for a lack of a tighter mapping from endogenous growth theory to data.

[66] I pick a skill premium that is 15% lower than in advanced economies, following arguments in Milanovic (2016) about a relatively lower skill premium in the Eastern Block.

growth rate is fixed and the comparison is based on level differences in the long run. I discuss transition dynamics below.

Table 1 and table summarize the cumulative wage effects for production and skilled labor, comparing the closed economy to the open economy. The most remarkable result is that in the new steady state real wages for production workers in Germany are 17% lower compared to the balanced growth path under autarky. While real wages of production workers stagnate, high skilled wages gain 11% in real terms. The negative effects on production worker wages are driven by weak domestic technology adoption, which dominates the gains from increased innovation. This means that the skill premium in Germany rises by 33%. The model does a good job at matching stagnant wage growth for production workers, and is qualitatively consistent with a rising skill premium in Germany, see figure 8 based on KLEMS data that I display in the next section, but the skill premium moves only around 10%. There is an ongoing debate to what extent the skill premium has increased in Germany.[67] The Gini index in the model increases by 6pp, which accounts for 75% of the observed increase in the data of roughly 8pp. Clearly, the model is stylized in that there are only two types of workers. The rising skill premium in my framework is mean to capture the rising returns to workers that are able to develop new ideas (innovation) or implement new ideas in a new context (adoption) in an increasingly global market for technology. The skill premium measured in the data is likely missing some of these effects as not everyone with a college degree will fall into this category.[68]

Table 1. Long Run Wage Effects

	Production Worker	High Skilled	GDP p.c.
Germany	-17%	+11%	-10%
Poland	+100%	+218%	+114%

I pick as a stand-in for Emerging Markets Poland, which is in the middle of the pack among Eastern European countries, poorer than the Czech Republic but richer than Romania. GDP growth in Poland is strong in my model and reflects broadly wage gains of production workers in Poland from 1995 to 2015. The overall wage gains are driven by technology adoption. The steady state skill premium in Poland implied by the growth spurt from $z_0 = .2$ to $z_\infty = .4$ is $s^{POL} = 3.7$, which implies an increase in the skill premium of 218%. This is an overstatement, and an implication of the model that wants a poorer country to have a larger skill premium. The implied skill-to-labor ratio needed through the lens of the model to generate the growth spurt is lower than the actual share in Poland, which is highly

[67] See Dustmann, Ludsteck, and Schönberg (2009) who find a rise in the skill premium, while Doepke and Gaetani (2020) find none.

[68] See the recent work of Smith et al. (2019) highlighting the importance of human capital for top income, which often comes in the form of profits and misleadingly characterized as capital income. In my model, capitalists doe not gain from integration in the long run, but skilled labor, which is the key input into innovation.

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 to deliver a wide adoption gap and lower real wages.[69] The skill premium in Poland has increased by 12% from 1995 to 2015, although again it seems likely that this measure misses large gains for workers that are able to adopt modern technology.[70]

Table 2. Long Run Effects for Symmetric Integration

	GDP p.c.	Skill Premium
baseline model	+41%	0%
Jones (1995)	+41%	NA

The focus of the analysis rests on the advanced economy, to draw out how foreign adoption gives rise to a domestic innovation-adoption tradeoff. It would be easy to match the evolution of growth and inequality of the Polish economy by allowing for additional heterogeneity. I prefer this simple version of the model that delivers a number of unique quantitative predictions with heterogeneity in the relative skill share as single driving force.

Symmetric Integration and Comparison to Jones (1995): Table 2 shows the effect in the case of symmetric integration, i.e. the two identical advanced economies integrate. GDP per capita, and thus real wages, increase by 41% relative to autarky. Importantly, symmetric integration leaves the skill-premium unchanged and the gains from market integration are shared evenly. In Jones (1995), there is only one type of labor and no adoption margin so the skill premium plays no role. Yet, the gains from integration are identical, essentially because the adoption margin does not respond in the case of symmetric integration and the scale effects are the same as in Jones's bench mark model where scale effects are parameterized by the constant elasticity of the real wage to the labor force $d \log w = \frac{1}{1-\phi} d \log L$.

The discrepancy between the model's disappointing growth effects in advanced economies and the

[69] 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, the model needs Germany to have a much larger college share than Poland to produce innovation and have a higher share of skilled labor devoted to adoption. 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).

[70] 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 in real terms. 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 now 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. I apply a CPI to compute real wage growth and use the series POLCPIALLMINMEI (annual average cpi all goods Poland) from Fred, downloaded on August 8 2022.

strong pro-growth effects found in, for instance, Rivera-Batiz and Romer (1991),[71] is explained by the weak dynamic 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, raising the dynamic knowledge spillover, or both. When keeping the adoption margin, production workers in rich countries become indifferent between autarky and integration when $\phi = .2$. Adoption still takes a hit but growing innovation 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.[72]

Efficiency: The adverse effects of integration on growth in my model suggest that the economy is inefficient and too little skill is devoted to technology adoption. To be clear, the paper has nothing to say about whether too much research is performed relative to producing final output, the classic trade-off studied in Romer (1990) or Jones (1995). My model suggests that given a fixed amount of skilled labor devoted to innovation and adoption, over-investment in research relative to adoption becomes a distinct possibility when ideas are getting harder to find.

To see this, I derive the planner solution in the autarky steady state.[73] The optimal allocation equals

$$\frac{h_D}{h_F} = \frac{\beta}{1-\theta}\left(1-\phi\right) , \tag{59}$$

which maximizes consumption per worker in the closed economy. It is unlikely that the market equilibrium coincides with the planner solution for three reasons. First, there are externalities in research. A negative value of ϕ implies a so-called "fishing out" of ideas where today's research makes finding ideas in the future harder. Second, there are markups which imply incomplete consumer surplus appropriation and ultimately tilt the balance toward too little research in equilibrium in the current setting.[74]

[71] See also 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).

[72] I maintain a long-run growth rate of 1%, so I have to adjust the population growth as follows $g_L = (1-\phi) * 0.01$ to make sure I compare economies with different autarky long run growth rates.

[73] Given log utility this is found by maximizing $\log(A_F z)$. 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 $(h_D)^{\frac{\beta}{1-\theta}}$ while the total number of ideas is proportional to $(h_F)^{\frac{1}{1-\phi}}$. Picking up some constant parameters in Λ_z (for instance the normalized measure of firms m among other variables), I obtain the following system

$$\begin{aligned} &\max \log(A_F z) \\ &s.t. \\ &z = \Lambda_z (h_D)^{\frac{\beta}{1-\theta}} \\ &A_F = \Lambda_F L^{\frac{1}{1-\phi}} (h_F)^{\frac{1}{1-\phi}} \\ &h_{tot} \geq h_D + h_F. \end{aligned} \tag{58}$$

[74] One could imagine that markups in the production sector create the opposite bias, i.e. too much research, but it turns out that this is not the case and in particular the value of σ does not show up at all in the decentral allocation.

Third, there is a spillover in adoption as entrants learn from incumbents after paying a fixed cost. This type of spillover is implicit in virtually any model of endogenous firm entry.[75] Since in my framework, the productivity of incumbents is the outcome of costly adoption decisions, it seems intuitive that this choice may not coincide with a planner solution as private firms fail to internalize the positive spillover on entrants. This is a force toward too little technology adoption in the de-central equilibrium.

To show these different channels in the most straightforward way, I compare the planner solution to the private allocation derived in (45) when effective discounting is low, i.e. $\tilde{\rho} \approx 0$. Under this scenario, the ratio of skilled labor in adoption relative to innovation reads

$$\frac{h_D}{h_F} = \frac{\beta}{1-\theta}\frac{1}{\alpha}\frac{1}{1+\frac{\delta_{ex}+g_M}{(1-\theta)(g_F+\delta_I)}}, \tag{60}$$

which is similar but not identical to (59).

The solution (60) highlights that the ratio of gross entry to effective technology depreciation $(\frac{\delta_{ex}+g_M}{(1-\theta)(g_F+\delta_I)})$ leads to insufficient adoption. Firms discount the future due to the random death shock δ_{ex}, and they do not take into account the positive spillover on new entrants that enter at a rate g_M. This leads to less adoption compared to the planner solution. Note that if the measure of firms is constant ($g_M = 0$) and there is no churn ($\delta_{ex} = 0$), the inefficiency disappears as infinitely lived firms fully internalize the benefits of their adoption activity.[76] Moreover, a markup $(\frac{1}{\alpha})$ in the research sector appears relative to the planner solution, while the research externality $(1-\phi)$ is missing in the private allocation. The markup leads to too little innovation, in contrast to the research externality, which leads to too much innovation. These last two forces are standard, see Jones (1995) for a discussion.

Given my calibration, the market equilibrium leads to under-investment in adoption relative to innovation from the point of view of a social planner. Market integration with emerging markets amplifies this inefficiency as the advanced economy has a comparative advantage in innovation. This feature of the model allows for the possibility of weak aggregate growth after market integration. In a world where ideas are harder to find, and new entrants experience a positive knowledge spillover from incumbents in the productions sector – two seemingly innocuous assumptions – an inefficient de-central equilibrium emerges with too little skilled labor devoted to technology adoption. Consequently, subsidizing innovation in this model would be counterproductive as it amplifies the initial inefficiency. The skill premium widens, and growth takes a hit as more labor is reallocated away form domestic technology adoption, which was under-supplied to begin with. The result of over-investment in research

[75] The reason that models of firm dynamics and growth require a spillover can be seen as follows. If, for example, entrants had to start at a fixed entry productivity, and economic growth leads to increasing productivity of incumbents, then no balanced growth path with a stationary size distribution would exist as the gap between entrants and incumbents keeps increasing over time. The standard assumption is that some type of learning spillover or "sampling from the distribution" takes place so that a stationary firm size distribution emerges. See Luttmer (2007), Lucas (2009b), Sampson (2016), Buera and Oberfield (2020), or Peters and Walsh (2019).

[76] This limiting case helps clarify the link to Parente and Prescott (1994), which is a model where adoption is efficient, and without endogenous innovation. I conjecture that introducing a positive gross entry rate into their model will lead to too little adoption as well.

contrasts Jones (1995), which finds that under-investment is the more likely outcome, even for negative values of ϕ.[77] First, the two results are not directly comparable as I don't consider a dynamic trade-off between consumption today vs. tomorrow. Second, and more importantly, the main difference to Jones (1995) is the technology adoption margin.

An important assumption underlying this welfare analysis hinges on the planner taking a national perspective. A planner that cares about world output instead faces a different trade-off. Suppose the planner puts equal weight on domestic and foreign income, and there are N such foreign economies of equal size. Again, assume that the planner only decides on the allocation of skilled labor within the advanced economy. The maximization problem becomes $\max\,(N+1)*\log\left(A_F\right)+\log z+\sum_{j=1,\dots,N}\log z_j$ and the optimal solution would change to $\frac{1}{N+1}\frac{\beta}{1-\theta}\left(1-\phi\right)=\frac{h_D}{h_F}$. Intuitively, pushing out the frontier helps both the domestic and foreign economies, so more labor is devoted to producing frontier technology. Any welfare statement thus crucially depends on whether the scope of the analysis is global or national.

Transition Dynamics: I next solve for transition dynamics. I focus on a simplified problem where I abstract away from capital accumulation dynamics, which are of second order importance in this setting.[78] Figure 6 plots the wage dynamics. The skill premium shoots up immediately and skilled labor gains instantaneously as market integration leads to rising rising returns to innovation "over night", which explains the jump in skilled workers' wages. The free entry condition in the research sector requires that the skill premium shoots up to make up for the rising value of innovation. As innovators enter the research sector they eventually push down the profits in innovation. This requires skilled labor to be reallocated toward innovation, and a new equilibrium with a higher overall level of frontier technology is reached. At the same time, a slow process of wage stagnation sets in for production workers. As the skill premium rises, firms in the production sector endogenously reduced their technology adoption effort. It takes more than 50 years for the technology adoption gap to reach its new, higher level. This slow process is not surprising as I picked θ to be consistent with the slow convergence dynamics found in the data.

Note that while z moves slowly, the equilibrium measure of ideas expands quickly. This coincides with an immediate and large reallocation of skilled labor towards innovative activity.

[77] Perla, Tonetti, and Waugh (2021) is another relevant benchmark. In their heterogeneous firm model larger firms fail to internalize the positive spillover they exert on smaller competitors or entrants, which copy the superior technology of larger incumbents. They show that a Melitz-type selection-into-exporting mechanism can alleviate this externality as market shares are reallocated toward larger incumbents, leading to faster growth in the open economy, similar to Sampson (2016). See the heterogeneous-firm open-economy model of Atkeson and Burstein (2010) for an efficient benchmark.

[78] Note that the long-run dynamics are pinned down by the smallest eigenvalue of the linearized system, which is given by the advantage of backwardness $1-\theta$ that is related to the process of technology adoption.

Figure 6. Wage Dynamics in Advanced Economy

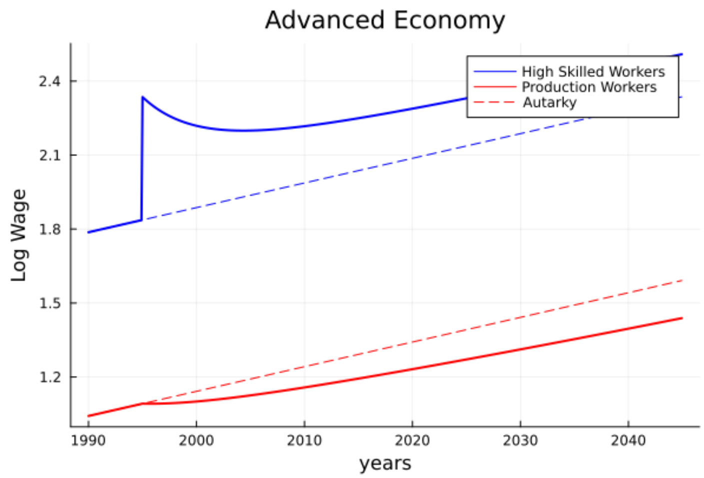
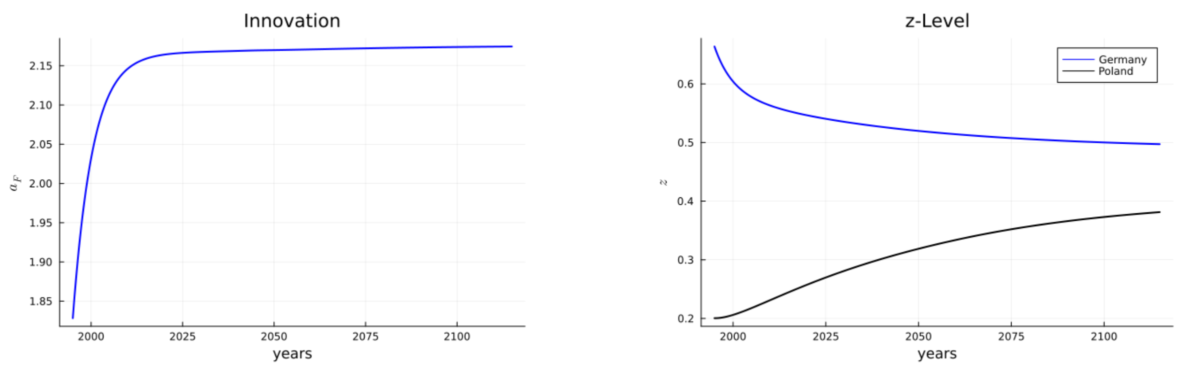


Figure 7. Innovation and Adoption in Open Economy



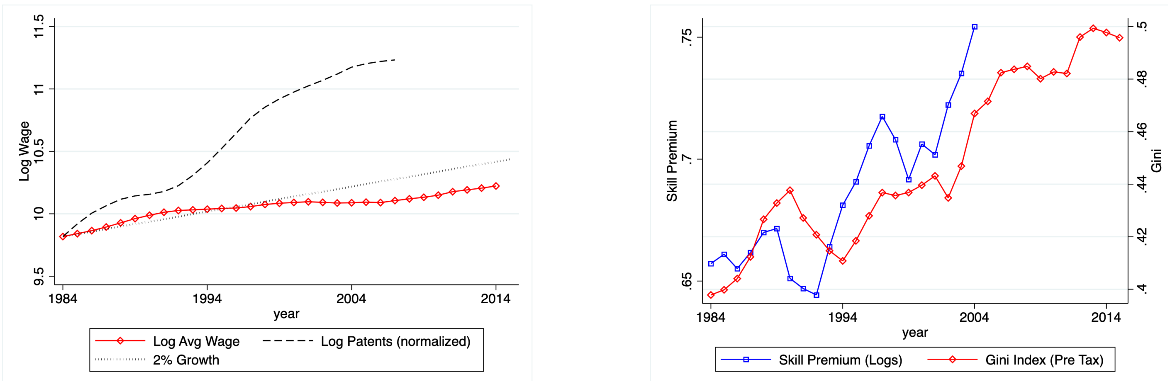
Time units are in $\Delta = 0.1$. Recall the definition of $a_F = \frac{A_F^{1-\phi}}{L}$.

5 Empirical Evidence

5.1 Aggregate Evidence

A feature of the theory is its ability to reconcile rising innovative activity against the backdrop of stagnant real wages and weak TFP growth, as seen in figure 7. Wages grew at a rate above 2% up until 1995. From then onward, Germany experienced its worst two decades of economic growth since WW2, where per capita income growth fell to a meager 0.55% annually despite strong patent growth, a proxy for innovation, as can be seen in figure 8. Van Ark, O'Mahoney, and Timmer (2008) provide careful evidence showing that productivity growth slowed down dramatically. German TFP growth from 1995-2004 is estimated to be .3%, an all time low in post war history.[79]

Figure 8. 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 8 shows the evolution of the skill premium, and the Gini Index, both of which shoot up in the mid 1990s, consistent with the model and the impact of market integration on the returns to innovation.[80] This pattern of robust innovative activity, weak productivity growth, and a divergence in real wages across workers is not unique to the German economy but seems to hold across a number of advanced economies. This is a puzzle for benchmark models of endogenous growth, but the quantitative exercise shows that the model proposed in this paper can account for these puzzling patterns. The decoupling of innovation and wage growth visible in figure 8 is explained by weak technology adoption during an episode of globalization that drives apart the returns of *local* adoption vs *global* innovation.[81]

There are many concerns about using patent data to proxy for innovation, not least that patents likely only reflect a small share of innovation and productivity growth.[82] An alternative approach is

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

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

[81]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. 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).

[82]Recent work highlights how patents are used "defensively" to shut down competitors without producing novel content,

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 as consisting of establishments with sectoral codes such as research, consulting, patent law, headquarter services, etc. The production sector is the rest of the economy. Thus, I follow a broad notion of "innovative" employment, and I am missing out on research activity performed in production firms. A detailed discussion can be found in the appendix, but the idea is to map the simple two-sector structure of the model into the sectoral classification in the data. In equilibrium, rising returns to innovation should show up as elevated firm entry rates and increasing total employment in the innovation sector so that excess profits are arbitraged away.

The left panel in figure 9 shows a massive increase in the relative employment share in innovation, consistent with the substantial rise in innovative activity in the model. Quantitatively, 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.[83] Note that I abstract away from structural change toward services and away from agricultural and manufacturing production, which should explain some of the shifts in employment. On the other hand, that abrupt acceleration in the mid1990s suggests that this secular trend does not exclusively drive rising employment.[84] The right panel of figure 9 reports 3-year moving average establishment entry rates across both sectors, showing that net entry and overall business dynamics took off in the innovation sector relative to the rest of the economy in the mid 1990s, consistent with the predictions of the model. The reallocation is fast, and consistent with the transition dynamics computed in the previous section.

These reallocation patterns are consistent with trade-induced sectoral specialization patterns. There is one critical difference, however. Technology is fixed in models of international trade, which leads to the typical gains from trade. In contrast, the exodus of skilled labor from the production sector has adverse effects on the *level of technology adoption in the production sector* in this model. A sectoral "brain drain" sets in, allowing for more nuanced effects of openness on growth and inequality.

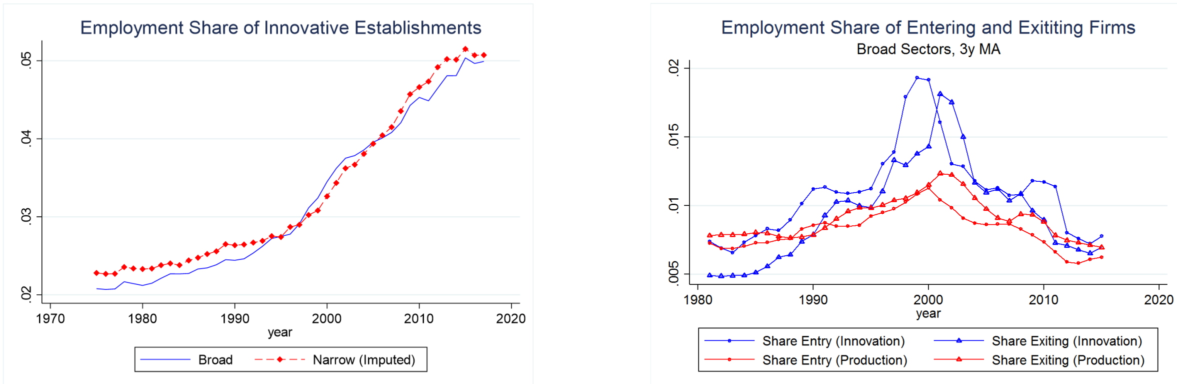
So far, I have shown that a model with an endogenous adoption gap can account for advanced

see Argente et al. (2020). Note that some of this literature is motivated by the fact that patenting activity has not translated into productivity growth. The model at hand provides an obvious explanation for this weak transmission since patents raise productivity only when the new technology is widely adopted.

[83] 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 20 in the appendix. However, while the skill share in the production sector is falling relative to the skill share in the research sector, both sectors exhibit a rising skill share. A more realistic model would thus include an additional source of skill-biased task-changing technological change 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.

[84] 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 are 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) could lead to more fragmentation between innovation and production. My establishment measure of innovative employment would understate the amount of research done in the early 1990s and thus overstate the growth rate.

Figure 9. 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.

economies' uneven and sluggish growth experience after market integration with emerging markets. Aggregate patterns are consistent with the theory and hard to reconcile with benchmark growth models that do not feature an adoption margin. In the German context, these patterns are particularly stark: A dramatic rise in exports from Germany from around 20% to 45% from 1995 to 2005 and German multinationals heavily invest in Eastern Europe, which leads to cross-border technology and profit flows between between Germany and the "East" (see Dauth, Findeisen, and Suedekum (2014))[85] and and rising innovative effort coincide with weak productivity growth and wage stagnation. 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.

5.2 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 should experience a positive shock due to the rising global demand for ideas. This idea demand shock should lead to elevated skilled employment and GDP growth in these regions, if comparative advantage in innovation is relatively fixed over time and unevenly distributed across space. Both assumptions are consistent with an extensive literature on the persistent clustering of innovation across space, see Feldman (1994). Moreover, to identify any effects, labor must be mobile across space. Weak labor mobility in German regions suggests that the exercise

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

is biased against finding any effects.

This cross-sectional approach relates to a recent literature in macroeconomics (Nakamura and Steinsson, 2014; Mian and Sufi, 2014; Chodorow-Reich, 2019) and international trade (Hummels et al., 2014; Autor, Dorn, and Hanson, 2013). I focus on West Germany to avoid dealing with the massive institutional change in East Germany after German unification around 1990.[86] 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 the overall trend toward integration with the West in Eastern Europe.[87]

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.[88] 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 prior to the beginning of each period.[89] Figure 10 plots the positive correlation between log patents and log population across counties (Kreis-level).

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,

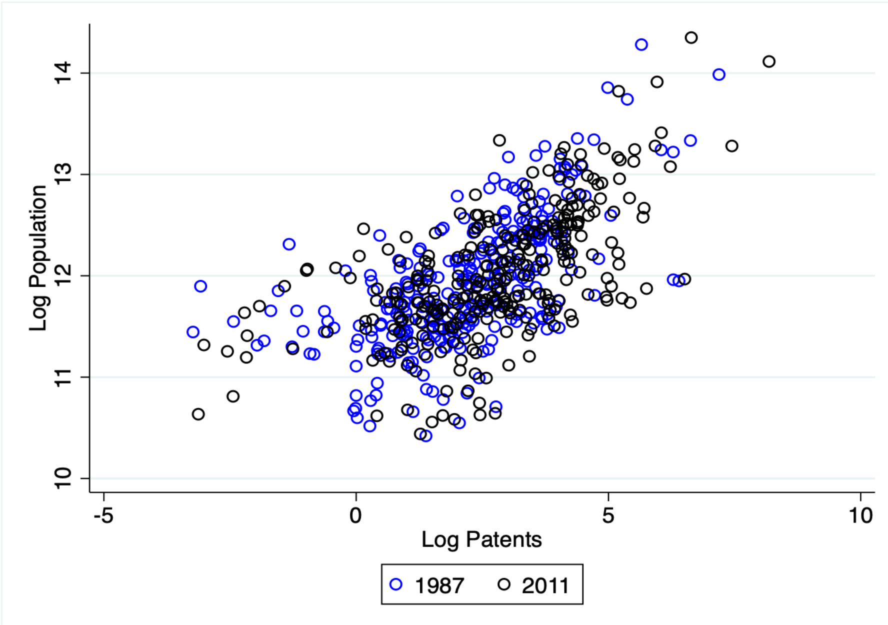
[86] See Findeisen et al. (2021) for work on employment reallocation in the East Germany. Note that most of the convergence within German between East Germany and West Germany occurs up until 1995, see Bachmann et al. (2022). Importantly, while some East German workers did migrate to the West, Findeisen 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.

[87] Another potential concern relates to the role of immigration in explaining weak wage growth in Germany in the 2000s. This hypothesis is largely rejected empirically. See Glitz (2012) and Dustmann and Glitz (2015) for the German context which report a null-finding when estimating the negative effect of immigration on native wages. Moreover, Dustmann, Frattini, and Preston (2013), Ottaviano and Peri (2012), and Card (2001) even find average wage gains in local labor markets more exposed to immigration for the UK and the US, respectively. See Dustmann, Schönberg, and Stuhler (2016) for a review of this large literature. I also show that the aggregate foreign employment share in Germany is at an all-time low in the early 2000s in the appendix in figure 15, casting further doubt on this argument.

[88] 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.

[89] 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.

Figure 10. Population & Patents Across Counties in West Germany



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

$$\Delta_k \log pop_{rt} = \alpha + \gamma_t + \left(\beta + \underbrace{\delta_{t>1995}}_{>0} \right) Patents_{rt} + \log\left(pop_{rt}/area_{rt}\right) + u_{rt}. \tag{61}$$

Controlling for density is essential since there is mean reversion in population growth. Moreover, density is a good proxy for average GDP per capita so that the specification effectively hods fixed the level of development. I report the results in table 3, which confirms that initial patent specialization strongly predicts population growth from 1996 onward but, crucially, 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 also run a version of (61) using log patents, which leads to a semi-elasticity that is easier to interpret: a 14% increase in initial patents leads to a 0.1 percentage point increase in population growth.[90]

A potential confounder is skill-biased technological change, see the recent work on urban-biased

[90] 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.

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.
* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

growth and technological change (Giannone, 2017; Rubinton, 2020; 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 explain the overall weak growth performance in the 90s and 2000s, one needs an additional mechanism since skill-biased technological change tends to raise aggregate productivity. The adoption margin is key to resolving this puzzle, and I offer some evidence on this channel next.

Missing Technology Adoption: While measurement of technology adoption is challenging,[91] there are clear tell-tale patterns in the data.[92] First, I document changing wage growth convergence patterns across labor markets where regional catch-up growth gave way to regional divergence since the fall of the Iron Curtain.[93] 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 1994 as the dividing line.[94]

[91] The classic reference on technology diffusion is Griliches (1957)'s study on hybrid corn. Comin and Hobijn (2010b) more recently measure the use of technologies on the country level, and Bloom et al. (2021) use data from company earnings call in combination with machine learning and text analysis tools to study the diffusion of technology. These papers suggest that the degree of technology diffusion differs substantially across countries and regions but are restricted to case studies of particular technology.

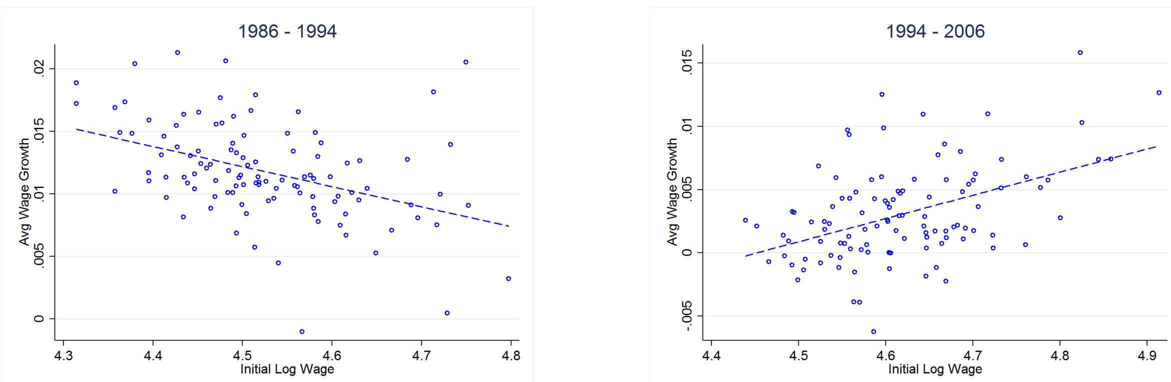
[92] While the sampling variation is too large to tease out regression-based effects, as reported in the previous section, the data is very useful to document broad trend breaks in wage growth and employment growth across local labor markets.

[93] 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.

[94] 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

Figure 11 plots average wage growth, defined as the total wage bill of full-time employees over total full-time employment, against the log of the initial average real wage for a local labor market, following Baumol (1986) and Sala-i-Martin and Barro (1997). While wage growth in the early period from 1986 – 1994 was, on average, higher for laggard regions. These growth patterns are turned upside down in the 2000s, where high-income places grew relatively fast while laggard regions stagnated.[95] 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, frontier growth could not compensate for weak growth in the hinterlands, consistent with the aggregate growth slowdown predicted by the model. These changing convergence patterns are more broadly true across the USA and advanced European Economies as I show in the appendix.

Figure 11. 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.

Through the lens of the model, stagnation in laggard regions represents an endogenous widening of the adoption gap due to a rising price of skill. The gap widens until the advantage of backwardness is sufficiently strong to compensate for the increased cost of technology adoption. A common concern is that international trade, and in particular import exposure following Autor, Dorn, and Hanson (2013), fully explains weak growth in laggard regions. To consider the effect of import exposure on

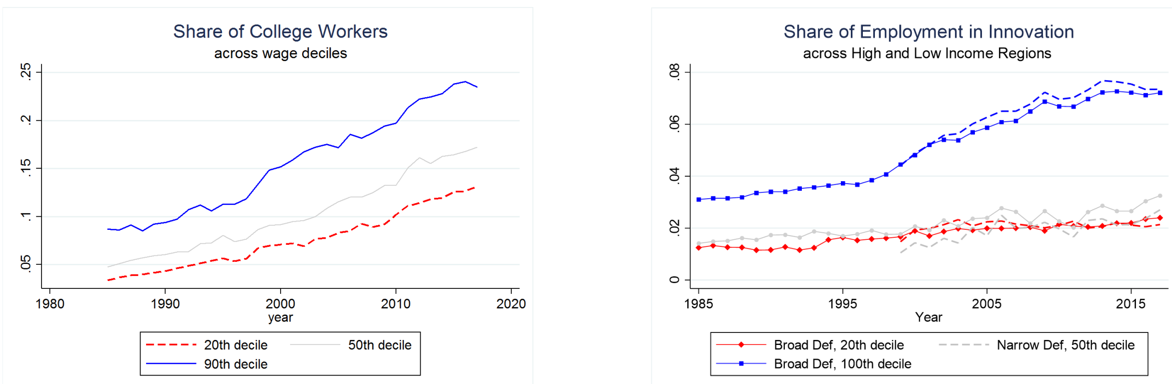
for this structural break from 1983 to 1984. My sample cut avoids this issue altogether. The period from 1975-1985 does not feature strong convergence dynamics in wages, but it does feature strong convergence dynamics in the skill ratio of each region. Overall, there is a clear trend of within-country regional convergence in Europe and the US from 1950 - 1990 as I show in the appendix.

[95] 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 considered against the thick right tail of the income distribution.

wage growth, I run a convergence regression with an additional import exposure variable as a control variable. Import competition accounts for virtually none of the stagnation in laggard regions. Results are reported in the appendix.

To corroborate the interpretation that a (relative) loss of skill hurts laggard regions, the left panel in figure 12 shows how the share of college workers in high-income regions has been diverging in the decade starting in 1994. An acceleration in the college share of high-income regions gave way to stagnation in the college share of 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 skill growth in innovative centers comes at the cost of production-focused regions.

Figure 12. 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.

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) and Moretti (2012). Table 4 reports that skilled labor growth is robustly correlated with wage growth in both periods, yet its moved from laggard to leading regions.

Note that technological change was also skill-biased in the early post-war period as documented in Goldin and Katz (2010). The crucial difference through the lens of this model is that adoption-driven growth gave way to frontier growth. When skilled labor helps adopt new technology, improving real wages for all worker types is a natural outcome and consistent with the positive correlation between low skilled wage growth and skilled employment growth. This association disappeared in the more recent period as seen in table 4. A model where adoption and innovation compete for skilled labor in

Table 4. Wage Growth & Total High Skill Employment Growth

		$g_H^{1986-1994}$		$g_H^{1994-2006}$		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.

a globalized world explains these changing growth patterns across space and workers all at once.

Lastly, related research lends credibility to the central weak-adoption-channel in the paper. Recall that a scarcity of human capital, and a rising skill premium, lead to a widening adoption gap in my model. Using micro data and a causal estimation design, Lewis (2011), Beaudry, Doms, and Lewis (2010), and Imbert et al. (2022), provide compelling evidence that a change in the local skill mix towards less skilled workers reduces a local labor market's ability to adopt frontier technology. This is precisely what my theory would predict. In my closed economy version, an increase in the low-skilled labor force shows up as a rise in the skill premium, leading to a larger equilibrium adoption gap. My model thus provides a tractable micro-foundation in a dynamic general equilibrium setting that highlights the central role of the skill premium and its negative impact on technology adoption.

6 Conclusion

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, while domestic technology adoption stalls. I make this argument precise by generalizing the model of Jones (1995) to include two types of labor and an endogenous technology adoption gap. The theory highlights how innovation and technology adoption are complementary on the market for ideas, but at the same time compete for skilled labor on factor markets. This leads to a novel role for the skill premium, which directly impacts productivity through its effect on equilibrium adoption effort.

In my calibration, weak domestic technology adoption entirely erases gains from additional innovation in the aftermath of market integration between advanced economies and emerging markets. The mechanism can generate sizable real wage losses for production workers in rich countries, and a rising skill premium. The theory matches weak aggregate growth in advanced economies despite rising innovative efforts and increasing globalization, which eludes benchmark growth models. Empirical evidence from Germany is consistent with the key mechanism. Notably, the broad patterns in the data – uneven growth across space and workers where the innovative sector runs away from the rest of the economy – have been documented elsewhere, in countries like the UK, France, or the USA.

Much work remains to be done to discipline the innovation-adoption tradeoff that is the focus of the paper. Yet, I hope that the framework's simplicity and its ability to explain several patterns in the data all at once will contribute to the reader's understanding of the nexus of technological change, inequality, and globalization. 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 ideas. For this to be the case, human capital accumulation and rising research efforts in emerging markets are crucial. Recent concerns about the adverse effects of the ability of emerging markets to compete with advanced economies in high-tech may be misplaced. Innovation in emerging markets would push down the skill premium by reducing advanced economies' global market share in idea production. This, in turn, would induce a reallocation of skilled labor toward adoption activity and broad-based wage growth.

References

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

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

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

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

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.

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

Adao, Rodrigo et al. (2020). *International trade and earnings inequality: A new factor content approach.* Tech. rep. National Bureau of Economic Research.

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.

— (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.

Aristizabal-Ramirez, Maria, John Leahy, and Linda L Tesar (2022). "A North-South Model of Structural Change and Growth". In: *Journal of Monetary Economics*.

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

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

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

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.

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.

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

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. (2020). *The Stable Transformation Path*. Tech. rep. National Bureau of Economic Research.

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 (2001). “Immigrant inflows, native outflows, and the local labor market impacts of higher immigration”. In: *Journal of Labor Economics* 19.1, pp. 22–64.

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.

— (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, Tommaso Frattini, and Ian P Preston (2013). “The effect of immigration along the distribution of wages”. In: *Review of Economic Studies* 80.1, pp. 145–173.

Dustmann, Christian and Albrecht Glitz (2015). “How do industries and firms respond to changes in local labor supply?” In: *Journal of Labor Economics* 33.3, pp. 711–750.

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

Dustmann, Christian, Uta Schönberg, and Jan Stuhler (2016). "The impact of immigration: Why do studies reach such different results?" In: *Journal of Economic Perspectives* 30.4, pp. 31–56.

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.

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.

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

Feenstra, Robert C and Gordon H Hanson (1996). "Globalization, outsourcing, and wage inequality". In: *The American Economic Review* 86.2, p. 240.

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.

Galle, Simon, Andrés Rodríguez-Clare, and Moises Yi (2017). *Slicing the pie: Quantifying the aggregate and distributional effects of trade.* Tech. rep. National Bureau of Economic Research.

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.*

Glitz, Albrecht (2012). "The labor market impact of immigration: A quasi-experiment exploiting immigrant location rules in Germany". In: *Journal of Labor Economics* 30.1, pp. 175–213.

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.

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

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, 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.

Hummels, David et al. (2014). "The wage effects of offshoring: Evidence from Danish matched worker-firm data". In: *American Economic Review* 104.6, pp. 1597–1629.

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

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.

Jones, Charles I (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.

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

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.*

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

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

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.

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

Moll, Benjamin (2014). "Productivity Losses from Financial Frictions: Can Self-Financing Undo Capital Misallocation?" In: *American Economic Review* 104.10, pp. 3186–3221.

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

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.

Obstfeld, Maurice and Kenneth Rogoff (1995). "The intertemporal approach to the current account". In: *Handbook of international economics* 3, pp. 1731–1799.

Ottaviano, Gianmarco IP and Giovanni Peri (2012). "Rethinking the effect of immigration on wages". In: *Journal of the European economic association* 10.1, pp. 152–197.

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

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

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

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

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

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

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

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.

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

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

Sampson, Thomas (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.

Smith, Matthew et al. (2019). "Capitalists in the Twenty-first Century". In: *The Quarterly Journal of Economics* 134.4, pp. 1675–1745.

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

$$\begin{array}{ll} \max & \int_0^A \left(\frac{x_j}{\alpha}\right)^\alpha dj \\ s.t. & \int p_j x_j dj \leq I. \end{array} \tag{62}$$

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

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

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

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

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

The problem further simplifies to

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

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

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

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

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

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,

$$\begin{aligned} v\left(\rho+\delta_{ex}\right) &= \max_h \frac{\pi_t(z)}{w}-s_th+v_z\dot{z}+\dot{v} \\ s.t. & \\ \dot{z} &= \zeta z^{\theta}h^{\beta}-\left(g_F+\delta_I\right)z. \end{aligned} \tag{63}$$

A solution to the program (63) needs to satisfy the following first order condition

$$\left\{\frac{v_z\beta\zeta z^{\theta}}{s}\right\}^{\frac{1}{1-\beta}} \quad = \quad h \ . \tag{64}$$

Equation (64) 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

$$\frac{v_{zz}}{v_z} + \frac{\theta}{z} = (1-\beta)\frac{h_z}{h}$$

$$\frac{v_{tz}}{v_z} - \frac{\dot{s}}{s} = (1-\beta)\frac{h_t}{h}.$$

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

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

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

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

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

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

In the steady state, we have that

$$h^{1-\beta} = \frac{1}{s}\frac{\beta(1-\alpha)(\sigma-1)(g_F+\delta_I)}{\rho+\delta_{ex}+(1-\theta)(\delta_I+g_F)}\left[\frac{\pi}{w}\right]\frac{\zeta z^{\theta-1}}{(g_F+\delta_I)} \tag{65}$$

$$z^{1-\theta} = \frac{\zeta h^\beta}{g_F+\delta_I} \tag{66}$$

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 effective costs in steady state, which is related to effective discounting as well as the advantage of backwardness. The firm needs to take into account that as it climbs up the technological ladder, the pull force introduced through the advantage of backwardness diminishes. This gives rise to an endogenous adoption gap as a function of the relative price of skill. Moreover, climbing up the ladder is costly when discounting is high since the benefits only accrue in the future, which is heavily discounted.

A.1.3 Firm value function off and on the balanced growth path

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

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

This dynamic HJB equation is tied to the free entry condition in a useful way, as shown in Peters and Walsh (2019). Note that the free entry condition implies $v_z = -\dot{v}$, i.e. totally differentiate $f = v(z,t)$. I can use this relationship to simplify the HJB equation where it must be understood that h solves the dynamic adoption problem. Rearranging yields

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

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

To check that this is a solution, it should be the case that it is consistent with the free entry condition and the present discounted value of entry,

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

Indeed, we have found a solution. This simplicity follows from the fact that the free entry condition at any point disciplines the profits that an incumbent firm can earn, see Peters and Walsh (2019) for a lucid application. 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.

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,

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

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

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

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

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

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

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

In standard q-theory applications, θ equals one and π_z is falling in z.[96] 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.

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

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. If $\beta \to 1$, adjustment occurs extremely fast as there is no curvature in h. In that case, the model jumps to the new steady state instantaneously.

Partial Equilibrium Investment vs. General Equilibrium Dynamics

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

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

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.

Nonetheless, when using global solution methods below every individual firm needs to be allowed to move to any z_i they desire. The inequality offered in proposition 1 is important for this solution to exist and if it doesn't hold, firms want to adopt too much technology in the sense that their individual incentive to learn is so high that they end up making flow profits below of what the fixed cost of entry requires for them to break even. Ex ante, no firm should enter and equilibrium is not well defined. Being in the market and refusing to make these large learning efforts does not help either. Note that profits are proportional to $\frac{z_i}{z_{agg}}$ and so if the aggregate moves but z_i is fixed, again, profits will be too low to make up for the entry cost. The inequality is thus essential and bounds the benefit of adoption to ensure a free entry equilibrium concept is well defined.

Phase Diagram:

Next, I show that a unique saddle-path stable equilibrium obtains where I keep s fixed. To show that first note that (65) establishes a negative link between z and h, while (66) 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 Innovation problem and market clearing conditions for high skilled labor

A.2.1 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

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

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

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

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

A.2.2 Innovation on and off the balanced growth path

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

$$(r+\delta_I)\,V_I-\dot{V}_I=\exp\left(-\int_t^{t+\tau}(r+\delta_I)\,dx\right)\frac{\alpha L_{t+\tau}^P w_{t+\tau}}{A_{F,t+\tau}z_{t+\tau}}\left[1+\tau'\right]. \tag{67}$$

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 (67) leads to

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

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

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

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

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

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

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

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

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

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

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

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

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

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.2.3 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, $g_A = g_F$. Then,

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

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

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

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:

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

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

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

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

$$\begin{aligned}\log A_{Ft} - \log A_t &= \delta_I \tau + \log A_{t+\tau} - \log A_t \\ \tfrac{\log A_{Ft} - \log A_{t+\tau}}{\delta_I} &= \tau.\end{aligned}$$

I totally differentiate this expression to obtain

$$\begin{aligned}g_F dt &= \delta_I d\tau + g_A (t+\tau) d\tau + g_A (t+\tau) dt \\ &\Leftrightarrow \\ \tfrac{d\tau}{dt} &= \tfrac{g_F - g_A(t+\tau)}{\delta_I + g_A(t+\tau)}\end{aligned}$$

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

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

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 if there was no learning whatsoever, it would be the case that $g_A = -\delta_I$ emerges and the derivative τ' would 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.2.4 Stochastic Adoption

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

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

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

A.3 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.[97]

A.4 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,

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

[97]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

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

which intuitively follows from total output minus total consumption of the final good.[98]

A.5 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

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

It follows that

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

[98] 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

$$\frac{\chi}{1-\chi} = \left(\frac{z}{z^*}\right)\frac{sh^{tot} - \Lambda^D}{s^*h^{tot,*} - \Lambda^D}$$

$$\left(\frac{\gamma}{\gamma^*}\right)^{\frac{1}{1-\lambda}} \left(\frac{s}{s^*}\frac{z}{z^*}\right)^{-\frac{\lambda}{1-\lambda}} = \left(\frac{z}{z^*}\right)\frac{sh^{tot} - \Lambda^D}{s^*h^{tot,*} - \Lambda^D}$$

$$\Leftrightarrow$$

$$\left(\frac{\gamma}{\gamma^*}\right)^{\frac{1}{1-\lambda}} \left(\frac{s}{s^*}\frac{z}{z^*}\right)^{-\frac{\lambda}{1-\lambda}} = \left(\frac{z}{z^*}\right)\frac{sh^{tot} - \Lambda^D}{s^*h^{tot,*} - \Lambda^D}$$

$$\left(\frac{\gamma}{\gamma^*}\right)^{\frac{1}{1-\lambda}} \left(\frac{s}{s^*}\left(\frac{s}{s^*}\right)^{-\frac{\beta}{1-\theta}}\right)^{-\frac{\lambda}{1-\lambda}} = \left(\frac{s}{s^*}\right)^{-\frac{\beta}{1-\theta}}\frac{sh^{tot} - \Lambda^D}{s^*h^{tot,*} - \Lambda^D}$$

$$\left(\frac{\gamma}{\gamma^*}\right)\left(\left(\frac{s}{s^*}\right)^{\frac{1-\theta-\beta}{1-\theta}}\right)^{-\lambda} = \left(\frac{s}{s^*}\right)^{-\frac{(1-\lambda)\beta}{1-\theta}}\left(\frac{sh^{tot} - \Lambda^D}{s^*h^{tot,*} - \Lambda^D}\right)^{1-\lambda}$$

$$\left(\frac{\gamma}{\gamma^*}\right) = \left(\frac{s}{s^*}\right)^{\frac{(1-\theta-\beta)\lambda-\beta(1-\lambda)}{1-\theta}}\left(\frac{sh^{tot} - \Lambda^D}{s^*h^{tot,*} - \Lambda^D}\right)^{1-\lambda}$$

$$\left(\frac{\gamma}{\gamma^*}\right) = \left(\frac{s}{s^*}\right)^{\frac{(1-\theta-\beta)\lambda-\beta(1-\lambda)}{1-\theta}}\left(\frac{sh^{tot} - \Lambda^D}{s^*h^{tot,*} - \Lambda^D}\right)^{1-\lambda}$$

$$\left(\frac{\gamma}{\gamma^*}\right) = \left(\frac{s}{s^*}\right)^{\frac{(1-\theta-\beta)\lambda-\beta(1-\lambda)-(1-\lambda)(1-\theta)}{1-\theta}}\left(\frac{s}{s^*}\right)^{1-\lambda}\left(\frac{sh^{tot} - \Lambda^D}{s^*h^{tot,*} - \Lambda^D}\right)^{1-\lambda}$$

$$\left(\frac{\gamma}{\gamma^*}\right) = \left(\frac{s}{s^*}\right)^{\frac{(1-\theta-\beta)\lambda+(1-\lambda)(1-\theta-\beta)}{1-\theta}}\left(\frac{h^{tot} - \frac{\Lambda^D}{s}}{h^{tot,*} - \frac{\Lambda^D}{s^*}}\right)^{1-\lambda}$$

$$\left(\frac{\gamma}{\gamma^*}\right) = \left(\frac{s}{s^*}\right)^{\frac{1-\theta-\beta}{1-\theta}}\left(\frac{h^{tot} - \frac{\Lambda^D}{s}}{h^{tot,*} - \frac{\Lambda^D}{s^*}}\right)^{1-\lambda} \tag{68}$$

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.

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 68 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 68.

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 and pushes down the skill premium in the foreign economy.

A.5.1 Open Economy Real Wage Effects

Recall the expression for real wages in the open relative to the closed economy ("Welfare"),

$$\frac{w^{open}}{w^{closed}}=\underbrace{\left(\frac{h_F^{open}}{h_F^{closed}}\right)^{\frac{\lambda}{1-\phi}}\left(\frac{1}{\chi}\right)^{\frac{1}{1-\phi}}}_{\text{gains from frontier innovation}}\underbrace{\left(\frac{s^{open}}{s^{closed}}\right)^{-\frac{\beta}{1-\theta}}}_{\text{loss from missing adoption}}$$

and

$$\frac{w_H^{open}}{w_H^{closed}}=\underbrace{\left(\frac{h_F^{open}}{h_F^{closed}}\right)^{\frac{\lambda}{1-\phi}}\left(\frac{1}{\chi}\right)^{\frac{1}{1-\phi}}}_{\text{gains from frontier innovation}}\underbrace{\left(\frac{s^{open}}{s^{closed}}\right)^{\frac{1-(\beta+\theta)}{1-\theta}}}_{\text{gains from rising skill premium}}.$$

One can derive this expression as follows, starting with the law of motion of ideas in the open economy along the balanced growth path,

$$
\begin{aligned}
\dot{A}_F &= \left(A_F^W\right)^{\phi} H_F^{\lambda} - \delta_I A_F \\
&\Rightarrow \\
\left(g_F + \delta_I\right)\chi &= \left(A_F^W\right)^{\phi-1} H_F^{\lambda} \\
\left(g_F + \delta_I\right)\chi &= L^{\lambda}\left(A_F^W\right)^{\phi-1}\left(\frac{H_F}{L}\right)^{\lambda} \\
\left(g_F + \delta_I\right)\chi \frac{\left(A_F^W\right)^{1-\phi}}{L^{\lambda}} &= \left(\frac{H_F}{L}\right)^{\lambda} \\
\left(g_F + \delta_I\right)\chi a_F^W &= \left(h_F\right)^{\lambda} \\
&\Leftrightarrow \\
a_F^W &= \left(h_F\right)^{\lambda}\frac{1}{\chi}\frac{1}{g_F+\delta_I}
\end{aligned}
$$

Now the ratio of frontier technology in the open and closed economy is simply given by $\frac{A_F^{W,open}}{A_F^{W,closed}} = \left(\frac{a_F^{W,open}}{a_F^{W,closed}}\right)^{\frac{1}{1-\phi}} = \left(\frac{h_F^{open}}{h_F^{closed}}\right)^{\frac{\lambda}{1-\phi}}\left(\frac{1}{\chi}\right)^{\frac{1}{1-\phi}}$ where I used the fact that $\chi^{closed} = 1$. In order to study the real wage effects I also have to account for the adoption margin since $\frac{w^{open}}{w^{closed}} = \frac{A_F^{W,open}}{A_F^{W,closed}}\frac{z^{open}}{z^{closed}}$. Note that $\frac{z^{open}}{z^{closed}} = \left(\frac{s^{open}}{s^{closed}}\right)^{-\frac{\beta}{1-\theta}}$ which delivers the result. The effects for skilled wages are almost identical, but the ratio of the skill premium needs to be added, i.e. $\frac{w_H^{open}}{w_H^{closed}} = \frac{A_F^{W,open}}{A_F^{W,closed}}\frac{z^{open}}{z^{closed}}\frac{s^{open}}{s^{closed}}$. Since this is a growth model, wages are growing at some constant long-run growth rate. The wage ratio reflects the long-run difference in wages after all temporary adjustments have taken place. In particular, since the long-run supply of capital is perfectly elastic, the capital-effective labor ratio is the same in the open and closed economy and is thus netted out in the ratio. This concludes the derivation.

A.5.2 Innovation profits in the open economy in special case with $\gamma^* = 0, \lambda = 1$

$$
\begin{aligned}
V_{It} &= \int_{t+\tau}^{\infty} \exp\left(-\left(r+\delta_I\right)\left(u-t\right)\right) L_{Pu} w_u \alpha \left(\frac{1}{A_u}\right) du \\
&+ \int_{t+\tau^*}^{\infty} \exp\left(-\left(r+\delta_I\right)\left(u-t\right)\right) L_{Pu}^* w_u^* \alpha \left(\frac{1}{A_u}\right) du \\
&= \left(\frac{\alpha}{r-g_w-g_L+g_F+\delta_I}\right)\left\{\left(\frac{L_{Pt}w_t}{A_t}\right)\exp\left(-\left(r-g_w-g_L+g_A+\delta_I\right)\tau\right) + \left(\frac{L_{Pt}^*w_t^*}{A_t^*}\right)\exp\left(-\left(r-g_w-g_L+g_A+\delta_I\right)\tau^*\right) \\
&= \left(\frac{\alpha}{r-g_w-g_L+g_F+\delta_I}\right)\frac{L_t^P w_{L,t}}{A^F}\left\{\left(\frac{1}{z}\right)\exp\left(-\left(r-g_w-g_L+g_A+\delta_I\right)\tau\right) + \frac{L_{Pt}^*w_t^*}{L_{Pt}w_t}\left(\frac{1}{z^*}\right)\exp\left(-\left(r-g_w-g_L+g_A+\delta_I\right)\tau \\
&= \left(\frac{\alpha}{\rho-g_L+g_F+\delta_I}\right)\frac{L_{Pt}w_t}{A_F} z^{\frac{\rho-g_L}{g_A+\delta_I}}\left\{1 + \frac{L_{Pt}^*w_t^*}{L_{Pt}w_t}\left(\frac{z^*}{z}\right)^{\frac{\rho-g_L}{g_A+\delta_I}}\right\}
\end{aligned}
$$

A.6 Entry with Partial Knowledge Spillovers

A simplifying assumption in the paper is the complete knowledge spillover from incumbents to entrants. 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 $A_K = \sup\{A_{Ki} : i \in \Omega_M\}$. This tweak turns the setting into a heterogeneous firm model where entrant learns from the most sophisticated incumbent, but imperfectly so, hence $\iota < 1$. I make one technical assumption, similar to the work in Benhabib, Perla, and Tonetti (2021) where an entrant arrives with a small probability p right at the frontier $A_{K,\max}$.[99]

A well-defined equilibrium is characterized by a distribution $f(z)$ with support $z \in [\iota z_{\max}, z_{\max}]$. This leads to a normalized free entry condition

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

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

Now, consider the problem of some firm i using the HJB approach in the steady state (so that $\dot{v} = 0$)

$$(\rho + \delta_{ex}) v(z_i) = \max_h \pi(z_i) - s h_i + (\partial_{z_i} v) \cdot \left[\zeta z_i^\theta h_i^\beta - (\delta_I + g_F) z_i\right]$$

with the first order condition

$$h_i = \left\{\frac{(\partial_{z_i} v) \beta \zeta z_i^\theta}{s}\right\}^{\frac{1}{1-\beta}}.$$

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

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

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

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

which helps to pin down the equilibrium dynamics. By construction, the most productive firm hires a constant amount of skilled labor with the only difference to the homogenous firm model being that

[99] This is a technical assumption to ensure a stationary distribution emerges, similar to Benhabib, Perla, and Tonetti (2021). In a stationary distribution, the share of maximum productivity level firms needs to be constant. Note that due to the death shock δ_{ex} a fraction of top firms dies each instant. And while other firms converge to the frontier, they may never fully reach it. So in the long run the share of top firms becomes arbitrarily small, even though all firms converge to this maximum productivity level. This leads to troubling limiting properties of the distribution, and a simple fix is to allow for a few firms to get luck and arrive at the top productivity level instantaneously. Given the probability $p > 0$, this leads to a consistent and smooth distribution as $T \to \infty$ with a mass point at $z_{\max}$.

the steady state profits are larger. This is a direct consequence of starting out with an initially lower productivity. Higher long-run profits have to make up for low profits after the firm just entered, since the entry cost are the same in both cases, i.e.

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

Define $(\rho + \delta_{ex} + (1-\theta)(\delta_I + g_F)) = \hat{\kappa}$ and note

$$\frac{\dot{h_i}}{h_i}(1-\beta) = \hat{\kappa}\left(1 - \left(\frac{z_i}{z_{\max}}\right)^{\theta-1+(1-\alpha)(\sigma-1)}\left(\frac{h_{\max}}{h_i}\right)^{1-\beta}\right).$$

This structure gives rise to a meaningful stationary distribution whereby firms start out small and improve their productivity over time, given regularity conditions. In particular, a well-defined unique solution emerges when the determinant of the linearized system is negative so that a negative and a positive eigenvalue leads to unique saddle-path stable convergence dynamics. For this to be the case, note that the matrix $\boldsymbol{A}$ defined as

$$\begin{pmatrix} \frac{\partial \log z}{dt} \\ \frac{\partial \log h}{dt} \end{pmatrix} \approx \underbrace{\begin{pmatrix} -(1-\theta)(\delta_I + g_F) & \beta(\delta_I + g_F) \\ -\hat{\kappa}\left[\theta - 1 + (\sigma-1)(1-\alpha)\right] & \hat{\kappa}(1-\beta) \end{pmatrix}}_{\boldsymbol{A}} \cdot \begin{pmatrix} \log(z/z_{ss}) \\ \log(h/h_{ss}) \end{pmatrix}$$

guarantees a unique saddle-path stable solution when its determinant is negative. For this to be the case, the following inequality needs to hold

$$\left[\theta - 1 + (\sigma-1)(1-\alpha)\right]\frac{\beta}{1-\theta} < 1 - \beta.$$

Given that a stationary equilibrium is well defined, a number of features are noteworthy.

First, the firm size distribution is independent of the relative price of skill s. What this suggests is that a new stationary equilibrium with a higher price of skill produces an identical wave but shifted to the left, i.e. a permanently lower level of adoption across all firms. This traveling wave property is not surprising in light of recent work on heterogeneous firms, see König, Lorenz, and Zilibotti (2016), Luttmer (2007), Sampson (2016),Benhabib, Perla, and Tonetti (2021) and Perla and Tonetti (2014). This is to say, the partial equilibrium elasticity $\frac{\partial \log \mathbb{E}[z]}{\partial \log s} = -\frac{\beta}{1-\theta}$ computed in the main text still applies except with an expectation operator.

Second, demand for skilled labor in the production sector can be derived by integrating over all productivity levels

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

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

$$V = \mathbb{E}_z \left[V(z) \right] \tag{69}$$

where $V(z) = \int_{t+\tau(z)}^{\infty} \exp\left(-\int_t^s [r_u + \delta_x]\, du\right) \pi_I(z,s)\, ds$ is a function of the firm-specific z-level which matters both in terms of firm size and how long it takes for an idea to be adopted by a firm of type z. The problem is conceptually the same as before except now one needs to keep track of the distribution of firm-specific adoption gaps. Of course, equation (69) is also conceptually very close to the value function of an innovator in the open economy in the main text, where many heterogenous countries (with different z-levels) would give rise to a similar integral.

B Transition Dynamics

B.1 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, an implication of log utility which leads to a constant saving rate when capital income is the only income. 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

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

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

$$\begin{aligned}
\frac{\dot{k}}{k} &= \frac{\dot{K}}{K} - g_{L^P} - g_F - g_z \\
&= \hat{\alpha}\frac{Y}{K} - (\rho - g_L + \delta_k) - g_{L^P} - g_F - g_z \\
&= \underbrace{\hat{\alpha}\frac{y}{k} - (\rho + g_F + \delta_k)}_{=0 \text{ in steady state}} - \underbrace{(g_{L^P} - g_L) - g_z}_{=0 \text{ in steady state}} \\
&= \hat{\alpha}\left(\frac{1}{\alpha}\right)^{\alpha}\left(\frac{1}{1-\alpha}\right)^{1-\alpha}(k)^{\alpha-1} - (\rho + \delta_k + g_A + g_{l^P})
\end{aligned}$$

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.2 Transition Dynamics Computation

I next derive transition dynamics and develop a solution algorithm that allows me to study the time path of the economy based on forward-looking optimal adoption and entry decisions in each sector and country. I normalize the system as follows $a = \frac{A^{1-\phi}}{L}, l^P = \frac{L^P}{L}, l^E = \frac{L^E}{L}, m = \frac{M}{L}, h^F = \frac{H^F}{L}$ so as to obtain a system of equations that admits a steady state.

I use global solution methods in continuous time following the algorithms developed in Achdou et al. (2022). While I don't have idiosyncratic risk, solving for the transition dynamics in the mode is hard and simple shooting algorithms (boundary value problem with iteration) don't work well. The problem is this: There is an endogenous price of skill that needs to be picked so that markets clear. In a neoclassical economy or Aiyagari type incomplete market model this is not very difficult because the demand for capital is static and simply downward sloping in the interest rate. The problem here is much harder because entry into innovation is endogenous, forward-looking, a function of the price of skill, and responds to all future adoption choices.

I solve this problem as follows. I use global solution methods in the production sector that gives a demand for skilled labor which moves smoothly in changing the skill-premium sequence $\{s_t\}_{t\in[0,T]}$. This gives much needed stability in the production sector, where a shooting algorithm explodes for a slightly wrong sequence of prices.[100]

B.3 Production Firm Problem

B.3.1 Preliminaries

Note that the profit term π is itself endogenous as the production side is closed by a free entry. In normalized form, this free entry condition reads $f_e = v(z,t)$ .Totally differentiating this expression and plugging it into the HJB equation (21) yields the following free entry condition that holds off and on the balanced growth path

$$f_e \quad = \quad \frac{\frac{\pi}{w} - sh}{r + \delta_{ex} - g_w} \quad . \tag{70}$$

Rearranging (70) leads to $f_e = \frac{\frac{l^P}{m(1-\alpha)(\sigma-1)} - sh}{r+\delta_{ex}-g_w}$ where l^P adjusts to ensure free entry holds. Inverting this relationship and accounting for the fact that the free entry condition may not be binding, I obtain the share of labor devoted to production $l^P = \min\{1, (1-\alpha)(\sigma-1)\, m\, (f_e[r+\delta_{ex}-g_w]+sh)\}$. The free entry condition may not be binding in particular when spending on technology adoption is very large. At that point, net profits are very small and no potential firm may be willing to incur the fixed cost of entry. In the same way, I can rewrite firm profits $\frac{\pi}{w} = \min\left\{f_e(r+\delta_{ex}-g_w)+sh, \frac{1_{\{l^P=1\}}}{m(1-\alpha)(\sigma-1)}\right\}$. Plugging this relationship into the law of motion of adoption leads to a dynamic equation that automatically respects potential corner solution along the transition path, i.e. I have

$$\frac{\dot{h}}{h}(1-\beta) = \rho + \delta_{ex} + (1-\theta)(g_F + \delta_I) \tag{71}$$

$$- \frac{\beta z^{\theta-1}\zeta h^{\beta-1}(1-\alpha)(\sigma-1)}{s}\left[\min\left\{f_e(r+\delta_{ex}-g_w)+sh, \frac{1}{m(1-\alpha)(\sigma-1)}\right\}\right] + \frac{\dot{s}}{s} \tag{72}$$

Lastly, I need to keep track of normalized firm entry m

$$\dot{m} = \frac{1-l^P}{f_e} - (\delta_{ex}+g_L)\, m.$$

B.4 Solving the model

I use global solution methods based on the finite differenced method in Achdou et al. (2022).

[100]In standard shooting algorithms, this is no problem as one can iterate on the initial value (or shoot backwards) up until the sequence converges. The problem here is that this will in general not be consistent with optimality in the forward-looking research sector. Again, a problem that does not arise in the benchmark neoclassical or incomplete market model because capital demand is static. Iterating on one sector, keeping variables in the other sector fixed, should in principal work but I have found this procedure to be extremely unstable. The global solution method suffers less from this issue.

The normalized firm problem reads

$$(r_t + \delta_{ex} - g_w)\, v\left(z_i, \overrightarrow{X}\right) = \frac{\pi_i}{w} - sh_i + (\partial_{z_i} v_i) * \left[\zeta_i z_i^{\theta} h_i^{\beta} - (\delta_I + g_F)\, z_i\right] + \mathcal{A} v_i$$

subject to the law of motion for z_i and the evolution of aggregate state variables captured in $\overrightarrow{X} = \{m, z_{agg}\}$. The differential operator $\mathcal{A}$ captures the effect of the drift in $\overrightarrow{X}$.[101] Before I derive optimality conditions, it is useful to consider some special aspects of this problem. First, note that the free entry condition (as long as it binds) implies $v = f_e$. I can thus directly infer the amount of labor l^P devoted to entry as a function of m and z as $l^P = \min\left\{m\,(1-\alpha)\,(\sigma-1)\left[sh + f_e\,(r + \delta_{ex} - g_w)\right], 1\right\}$. The fact that the value function is equal to the properly discounted flow profits minus the learning cost also offers me a way to ensure that my solution algorithm makes sense and coincides with the free entry condition $v = f_e$ whenever the free entry condition binds, and $v < f_e$ otherwise. This is useful because the global solution method is a somewhat non-standard application of the method of Achdou et al. (2022) with endogenous entry both in innovation and the production sector, which tend to be harder to solve than models of perfect competition.

One non-standard aspect is that I have to keep some aggregate z_{agg} which proxies for the choice of all other firms, from the point of view of an individual firm i that makes an optimal investment decision. Even though all firms are the same, but nonetheless choices are made individually so a clear distinction needs to be drawn. One can see in equation (71) that imperfect competition necessitates this more complicated approach. The reason is that the elasticity of substitution from the final goods aggregator shows up, precisely because each individual firm considers how much they can boost their profits relative to anyone else if they invested more.

Thus, for the individual choice to be optimal, it needs to i) take into account the choice of everyone else as an aggregate state variable that is moving (z_{agg}), and ii) the choice must be the optimal solution among all kinds of $z's$ that it could pick, even if they are "off" the equilibrium. In equilibrium, these must be the same, which means one has to find a fixed point by iteration where this is indeed the case. Convergence thus requires for the value function change below a set tolerance and for the endogenous drift on the diagonal to agree.

With this in mind, I can derive the first order condition of the firm which reads

$$h_i = \left\{\frac{\beta \zeta_i z_i^{\theta} \left(\partial_{z_i} v_i\right)}{s}\right\}^{\frac{1}{1-\beta}}.$$

[101] Both the aggregate movement in z_{agg} and the endogenous change in the measure of firms m are captured by the infinitesimal generator $\mathcal{A}$. To be clear, I have no risk and no additional Brownian motion so there are no second-order terms. Given that I already have blown up the state space to include the individual z_i, one can add shocks to firm productivity without much extra work. Note that the net present discounted value of an innovator might change, as it depends on the speed of adoption which might interact with the exogenous productivity shock and the entire distribution. In follow up work I focus on this margin but abstract away from it here.

Now construct

- grid with states $\{z_{i,}z_{agg}, m\}$
- choice $\{h_i\}$
- and when you make that grid you need to satisfy certain inequality that I have derived some place else...
- state variables that are constant in the steady state but change along a transition path are $\{s, g_F, r, g_w\}$
- Note that s is an endogenous outcome that I need to solve for in the steady state
- Note also that I have a solution for s independent of this quantitative routine because the steady state can be solved very easily as I show in the paper, yet the transitions are quite hard
- This also allows me to assess the quality of the approximation inherent to any quantitative solution routine

The finite difference solution routine requires boundary conditions at the lower and upper end of the state space of z such that $\dot{\underline{z}} \geq 0$ and $\dot{\overline{z}} \leq 0$. The appropriate boundary conditions can be enforced by setting the derivative for the backward drift and forward drift at these boundary points to

$$v_{\underline{z_i}} = \frac{s}{\beta (\zeta)^{\frac{1}{\beta}}} \left(\underline{z_i}\right)^{\frac{1-\theta-\beta}{\beta}} (\delta_I + g_F)^{\frac{1-\beta}{\beta}}$$
$$v_{\overline{z_i}} = \frac{s}{\beta (\zeta)^{\frac{1}{\beta}}} (\overline{z_i})^{\frac{1-\theta-\beta}{\beta}} (\delta_I + g_F)^{\frac{1-\beta}{\beta}}.$$

.

The derivation follows from setting

$$\dot{\underline{z}} = 0$$

$$\Leftrightarrow$$

$$\underline{h} = \left(\frac{\underline{z}^{1-\theta}}{\zeta}\left(\delta_I + g_{A_F}\right)\right)^{\frac{1}{\beta}}$$

$$v_{\underline{z_i}} \geq \frac{s}{\underline{z_i}^{\theta}\beta\zeta}\left(\underline{h}\right)^{1-\beta}$$

$$v_{\underline{z_i}} \geq \frac{s}{\underline{z_i}^{\theta}\beta\zeta}\left(\left(\frac{\underline{z}^{1-\theta}}{\zeta}\left(\delta_I + g_{A_F}\right)\right)^{\frac{1}{\beta}}\right)^{1-\beta}$$

$$v_{\underline{z_i}} \geq \frac{s}{\beta\zeta}\underline{z_i}^{-\theta+(1-\theta)\left(\frac{1-\beta}{\beta}\right)}\left(\frac{\delta_I + g_{A_F}}{\zeta}\right)^{\frac{1-\beta}{\beta}}$$

$$v_{\underline{z_i}} \geq \frac{s}{\beta\left(\zeta\right)^{\frac{1}{\beta}}}\left(\underline{z_i}\right)^{\frac{1-\theta-\beta}{\beta}}\left(\delta_I + g_{A_F}\right)^{\frac{1-\beta}{\beta}}$$

The case for $\overline{z}$ is analogous.

Note that as long as $(\sigma-1)(1-\alpha) < 1$ the payoff function is concave in the individual z which makes using the right drift easy. The concavity implies $v_{z,F} < v_{z,B}$ (forward and backward difference) and in turn $\dot{z}_F > \dot{z}_B$. Setting the backward difference to zero thus implies that a non-negative drift will be picked at the lower boundary. For my preferred calibration this inequality does not hold and I follow the strategy proposed in Achdou et al. (2022) and use the Hamiltonian instead to pick the right drift, a strategy the authors discuss in their supplementary material.

B.4.1 Transition dynamics in production sector given a sequence of $\{s\}$

Next, consider computing transition dynamics where a sequence $\{s_t\}_{t\in[0,T]}$ where T is a terminal time at which the system has converged to its new steady state.[102] Given and initial steady state z_0 and m_0 and a terminal steady state with z_T and m_T, the routine runs backward and computes the demand for skill in adoption in each instant. I have to discretize time, and use small .1 steps and solve the problem backwards.

This provides a sequence of transition matrices $\{A_t\}_{t\in[0,T]}$. In a standard incomplete market model this matrix can be used directly to compute the evolving distribution g_t that then delivers the aggregate demand for capital. Here, I am looking for the demand for skilled labor in production but it is conceptually quite similar, which in turn shapes the stock of knowledge z which is the key state variable.

[102] That is, T needs to be large enough so that this approximation is reasonable. To check this, I increase T and note that the convergence dynamics are not changed relative to convergence when T is smaller.

To compute the aggregate evolution of m and h I essentially rely on heterogeneous agent distributional economics to compute aggregate that mimic the behavior of the economy. Note that the finite difference scheme turns the whole problem into a discrete Markov chain in continuous time which means that transitioning from one state to another is stochastic. Clearly, since this is a homogenous firm model the probability of the measure of firms to be in one particular point on the grid space is one, and zero elsewhere. Yet, the finite difference approximation turns the whole system into one of random Poisson transitions. If the grid space is finite enough, this random transition, by a law of large numbers, is the same as a deterministic transition where the randomness in the drift gives way to deterministic laws of motion.

Keeping this approximation in mind, the aggregate transition is best approximated by starting out at a point on the grid and then letting the system run forward with the endogenous adoption decisions that shift mass from one grid point to another. For example, if the skill price falls, firms move a permanently higher level of z over time. To obtain the appropriate aggregate demand for human capital I need to know the measure of firms m_t and the choice $h_t(m, z_i, z_{agg})$. I reduce the state space by dropping z_{agg} which I know is equal to z_i in equilibrium. Denote this smaller system using a *small*-superscript.

Consider I have the matrix A_t^{small} and a small initial distribution g_0^{small} where all mass is on the grid point $g(z_0, m_0) = 1$ which also has a unique choice $h_0(z_0, m_0)$. Now, I apply the transition matrix A_t^{small} to obtain an evolving distribution g_t. Once I have the sequence g_t, together with the policy function h_t, normalized aggregate demand for skilled labor is given by

$$m_t h_t = \mathbb{E}_{z,m}[m_t h_t] = \sum_{m,z} g_t(m,z)\, h_t(m,z)\, m$$

This leads to normalized demand for skilled labor as a function of the sequence of skill premia. Importantly, to solve this problem an initial guess is required about what the innovation sector is doing since g_F is a key input into the production firm problem. I turn to the innovation sector next, and it is clear that a solution will require to iterate back and forth between the two sector, which are connected through a market clearing condition.

B.4.2 Recursive Innovation Firm Problem Closed Economy

Even though I have derived the present discounted value in closed form, note that this expression features growth rates in the denominator. It is thus inherently unstable when a guess is slightly off the equilibrium growth rate form a computational point of view, as one finds oneself dividing through by something close to zero. I obtain a more stable system by computing a normalized version of the innovator value function recursively as follow.

First, recall the free entry condition, and after some algebra

$$
\begin{aligned}
A_F^{-\phi} w_H f_R &= \int_{t+\tau}^{\infty} \exp\left(-\int_t^u \left[r_x+\delta_I\right]dx\right)\pi_u du \\
A_F^{-\phi} s f_R &= \int_{t+\tau}^{\infty} \exp\left(-\int_t^u \left[r_x+\delta_I-g_w\right]dx\right)\frac{\alpha L_u^P}{A_{F,u}z_u}du \\
s f_R &= \int_{t+\tau}^{\infty} \exp\left(-\int_t^u \left[r_x+\delta_I-g_w\right]dx\right)A_F^{\phi}\frac{\alpha L_u^P}{A_{F,u}z_u}du \\
s f_R &= \int_{t+\tau}^{\infty} \exp\left(-\int_t^u \left[r_x+\delta_I-g_w\right]dx\right)\left(\frac{A_{F,t}}{A_{F,u}}\right)^{\phi}\frac{\alpha L_u}{A_{F,u}^{1-\phi}z_u}\frac{L_u^P}{L_u}du \\
s f_R &= \int_{t+\tau}^{\infty} \exp\left(-\int_t^u \left[r_x+\delta_I-g_w+\phi g_F\right]dx\right)\frac{\alpha l_u^p}{a_u z_u}du \\
s f_R &= \exp\left(-\int_t^{t+\tau}\left[r_x+\delta_I-g_w+\phi g_F\right]dx\right)\int_{t+\tau}^{\infty}\exp\left(-\int_{t+\tau}^u\left[r_x+\delta_I-g_w+\phi g_F\right]dx\right)\frac{\alpha l_u^p}{a_u z_u}du
\end{aligned}
$$

Introduce notation $v_{IN,t+\tau} = \exp\left(-\int_t^{t+\tau}\left[r_x+\delta_I-g_w+\phi g_F\right]dx\right)\int_{t+\tau}^{\infty}\exp\left(-\int_{t+\tau}^u\left[r_x+\delta_I-g_w+\phi g_F\right]dx\right)\frac{\alpha l_u^p}{a_u z_u}du$ is the normalized present discounted value of a firm that is selling goods in the market at time $t+\tau$. The normalizing factor is $wA_F^{-\phi}$. Now exploit the recursion and compute $v_{t+\tau}^{I,N}$ backwards. First, in the steady state it must be that

$$
v_{IN,T} = s_T f_R
$$

Out of steady state, I compute the evolution of v_{IN} by iterating backwards, using the following notation $\psi = [r_x+\delta_I-g_w+\phi g_F]$ and an approximation $dt \approx \Delta = 0.1$ to make this operational off the steady state

$$
\begin{aligned}
v_{IN,t} &= \exp\left(-\int_t^{t+\tau}\psi dx\right)\int_{t+\tau}^{\infty}\exp\left(-\int_{t+\tau}^{u}\psi dx\right)\frac{\pi_u}{w}du \\
v_{IN,t} &= \Lambda_t\left(\{\psi\},t,t+\tau\right)\int_{t+\tau}^{\infty}\exp\left(-\int_{t+\tau}^{u}\psi dx\right)\frac{\pi_u}{w}du \\
&\approx \Lambda_t\left(\{\psi\},t,t+\tau\right)\left\{\Delta\left(1+\tau'\right)\frac{\pi_{t+\tau}}{w_{t+\tau}}+\exp\left(-\int_{t+\tau}^{t+\tau+\Delta}\psi dx\right)\int_{t+\tau+\Delta}^{\infty}\exp\left(-\int_{t+\tau+\Delta}^{u}\psi dx\right)\frac{\pi_u}{w}du\right\} \\
&= \Lambda_t\left(\{\psi\},t,t+\tau\right)\left\{\Delta\left(1+\tau'\right)\frac{\pi_{t+\tau}}{w_{t+\tau}}+\exp\left(-\Delta\psi_{t+\tau}\right)\int_{t+\tau+\Delta}^{\infty}\exp\left(-\int_{t+\tau+\Delta}^{u}\psi dx\right)\frac{\pi_u}{w}du\right\} \\
&= \Lambda_t\left(\{\psi\},t,t+\tau\right)\left\{\Delta\left(1+\tau'\right)\frac{\pi_{t+\tau}}{w_{t+\tau}}+\exp\left(-\Delta\psi_{t+\tau}\right)\int_{t+\tau+\Delta}^{\infty}\exp\left(-\int_{t+\tau+\Delta}^{u}\psi dx\right)\frac{\pi_u}{w}du\right\} \\
&= \Lambda_t\left(\{\psi\},t,t+\tau\right)\Delta\left(1+\tau'\right)\frac{\pi_{t+\tau}}{w_{t+\tau}}+\exp\left(-\Delta\psi_t\right)v_{t+\Delta}^{IN}.
\end{aligned}
$$

Note that the recursion looks very similar in the open economy where the value function is split into a domestic and a foreign piece.[103]This approximation leads to a slightly different steady state value function, i.e.

$$
v_{IN,t} = \frac{\Lambda_t\left(\{\psi\},\tau\right)\Delta\pi}{1-\exp\left(-\Delta\psi\right)}
$$

where t is not an argument anymore. It is still true that

[103]To be precise, $v^{open} = \underbrace{\Lambda_t\left(\{\psi\},t,t+\tau\right)\Delta\left(1+\tau'\right)\pi_{t+\tau}+\exp\left(-\Delta\psi_t\right)v_{t+\Delta}^{IN}}_{domestic}+\underbrace{\Lambda_t\left(\{\psi^*\},t,t+\tau^*\right)\Delta\left(1+\tau'^*\right)\pi_{t+\tau^*}^*+\exp\left(-\Delta\psi_t^*\right)v_{t+\Delta}^{IN,*}}_{foreign}$

$$v_{IN} = f_R s$$

so

$$a_F = \frac{l^P}{z} \frac{\alpha}{f_R s} \frac{\Delta \exp(-\tau\psi)}{1 - \exp(-\Delta\psi)}$$

is the free entry condition where $\psi_{ss} = \tilde{\rho} + \delta_I + \frac{g_L}{1-\phi}$. Together with market clearing, this requires a slightly different measure of innovators in equilibrium relative to the analytical solution. Recall the law of motion of normalized ideas which reads

$$\dot{a}_F = (1 - \phi) \frac{h_F}{f_R} - a_F \left[(1 - \phi) \delta + g_L\right]$$

so in steady state demand must be equal to $h_F = f_R a_F \left[\delta + \frac{g_L}{1-\phi}\right]$ and the equilibrium consistent demand thus needs to satisfy

$$\begin{aligned}
h_F &= \frac{l^P}{z} \frac{\alpha}{s} \frac{\Delta \exp(-\tau\psi)}{1 - \exp(-\Delta\psi)} \left[\delta + \frac{g_L}{1 - \phi}\right] \\
h_F &= \frac{l^P}{z} \exp(-\tau\psi) \frac{\alpha}{s} \frac{\Delta}{1 - \exp(-\Delta\psi)} \left[\delta + \frac{g_L}{1 - \phi}\right] \\
h_F &= \frac{l^P}{z} \exp\left(\log z \frac{\tilde{\rho} + \delta_I + g_F}{g_F + \delta_I}\right) \frac{\alpha}{s} \frac{\Delta}{1 - \exp(-\Delta\psi)} \left[\delta + \frac{g_L}{1 - \phi}\right] \\
h_F &= z^{\frac{\tilde{\rho}}{g_F + \delta_I}} * \frac{\alpha l^P}{s} \frac{\Delta \left[\delta_I + \frac{g_L}{1-\phi}\right]}{1 - \exp\left(-\Delta\left[\tilde{\rho} + \delta_I + g_F\right]\right)}
\end{aligned}$$

Moreover, note that the equilibrium solution to the computational solution reads $\mathbb{E}[hm]$ which is also not exactly equal to the closed form solution. Thus, a new s^{comp} needs to be found that is consistent with the computational approximations. Specifically, I need to make sure the market clearing condition is satisfied so

$$h_F(s^{comp}) + h(s^{comp})\, m = h^{tot}.$$

Now the following algorithm can be used to compute a transition path:

1. Guess the sequences $\{g_F, z, \tau, s, l^P\}$
2. Solve production firm problem backwards
3. Use new $\{z\}$ to solve innovator problem backwards
4. Use $\{hm\}$ and $\{h^{tot}\}$ to obtain entry into innovation and the evolution of a_F as a residual, i.e. $h^F = h^{tot} - hm$
5. Compute $v_{IN,t}$ backwards, using the guessed sequences

6. Check if $v_{IN,t} = s_t f_R$ is consistent with free entry, this will usually not be the case

7. Update the skill premium gently upward if $v_{IN,t} > s_t f_R$ and vice versa if the inequality is reversed

8. Go back to step 1, update all guesses and keep iterating since you found a fixed point where

   (a) $\{s'\} = \alpha\{s\} + (1-\alpha)\{s_{old}\}$(whole vector being updated)

   (b) $a_{F,T} = \alpha a_{F,T} + (1-\alpha) a_{F,ss}$(only one scalar being updated)

The idea is that an increase in s directly reduces the wedge in 7) but it also has an indirect effect by inducing less demand for skill in the production sector so that an indirect effect raises a_F since more skilled labor is available for innovation, which in turn further pushes down $v^{I,N}$. Note that step b) in 8) is non standard. I update the steady state normalized idea stock because my analytical solution is not perfectly consistent with the approximation, which leads to slightly wrong dynamics at the very end of the transition. This is due to i) approximations in the stationary grid space (z, m) as well as approximations in time units (Δ).

Next I discuss how to find the sequence τ which depends on both g_A and g_F. First, note that the waiting time is implicitly defined by

$$\frac{-\log z}{\delta_I + \frac{\int_t^{t+\tau} g_A dx}{\tau}} = \tau \tag{73}$$

as I show in the theory section. In the steady state this is a simple closed form expression. Of the steady state, one has to find the initial τ_0 iteratively, i.e. try out different $\tau' s$, start with $\tau = 0$ and stop when (73) holds. Given a guess on $\{g_F\}$ and a sequence $\{g_A\}$ this can be computed. After that, I can use the fact that changes in the waiting time are equal to

$$\frac{d\tau}{dt} = \frac{g_F - g_A(t+\tau)}{\delta_I + g_A\,(t+\tau)}$$

which means, using the discrete approximation, that $\tau_{t+\Delta} = \tau_t + \Delta \frac{d\tau}{dt}$ and so the sequence can be solved forward using $\{g_F\}$ and $\{g_A\}$. This sequence, together with the evolution of $\{a_{F,} l^P, s, z\}$ is the information needed to compute the value of an innovation backwards. In the open economy, the same variables also need to be known about the foreign economy except for a_F since innovation only happens in the advanced economy.

B.5 Convergence with Both Countries Innovating

In contrast to the previous section I know consider a world where both countries are innovating. I assume that in the open economy the research technology is the same everywhere, and complete the degree of convergence in this world is then only dependent on the evolution of the skill ratio. I use the case of the Korean Growth miracle as an optimistic case to feed in an exogenous process of skill accumulation and show the evolution of the real wage and the skill premium in both advanced economies and emerging markets.

s

Here is the math:

First, compute the world value function of an innovation.

dition, and after some algebra

$$
\begin{aligned}
\frac{A_F^{-\phi} w_H H_F^{1-\lambda}}{\gamma} &= \int_{t+\tau}^{\infty} \exp\left(-\int_t^u \left[r_x + \delta_I\right] dx\right) \pi_u du + \int_{t+\tau^*}^{\infty} \exp\left(-\int_t^u \left[r_x^* + \delta_I\right] dx\right) \pi_u^* du \\
\frac{s H_F^{1-\lambda}}{\gamma} &= \int_{t+\tau}^{\infty} \exp\left(-\int_t^u \left[r_x + \delta_I - g_w\right] dx\right) \frac{\alpha L_u^P}{A_{F,u}^{1-\phi} z_u} du + \int_{t+\tau^*}^{\infty} \exp\left(-\int_t^u \left[r_x^* + \delta_I - g_w^*\right] dx\right) \frac{\alpha L_u^{P,*}}{A_{F,u}^{1-\phi} z_u^*} \left[\frac{w^*}{w}\right] du \\
\frac{s h_F^{1-\lambda}}{\gamma} &= \int_{t+\tau}^{\infty} \exp\left(-\int_t^u \left[r_x + \delta_I - g_w\right] dx\right) \frac{\alpha l_{Pu}}{a_{Fu} z_u} du + \int_{t+\tau^*}^{\infty} \exp\left(-\int_t^u \left[r_x^* + \delta_I - g_w^*\right] dx\right) \frac{\alpha l_{Pu}^*}{a_{Fu} z_u^*} \frac{z^*}{z} du \\
\frac{s h_F^{1-\lambda}}{\gamma} &= v_I + v_I^*
\end{aligned}
$$

Like before, the normalized value functions can be computed recursively for computational stability so that

$$
\begin{aligned}
v_I^* &= \Lambda_t\left(\{\psi^*\}, t, t+\tau^*\right) \Delta \left(1+\tau'^*\right) \frac{\pi_{t+\tau}^*}{w_{t+\tau}^*} \frac{w_{t+\tau^*}}{w_{t+\tau}} + \exp\left(-\Delta \psi_t^*\right) v_{I,t+\Delta}^* \\
&= \Lambda_t\left(\{\psi^*\}, t, t+\tau^*\right) \Delta \left(1+\tau'^*\right) \frac{\alpha l_{P,t+\tau}^*}{a_{Ft+\tau} z_{t+\tau}^*} \frac{w_{t+\tau^*}}{w_{t+\tau}} + \exp\left(-\Delta \psi_t^*\right) v_{I,t+\Delta}^* \\
&= \Lambda_t\left(\{\psi^*\}, t, t+\tau^*\right) \Delta \left(1+\tau'^*\right) \frac{\alpha l_{P,t+\tau}^*}{a_{Ft+\tau} z_{t+\tau}^*} \frac{z_{t+\tau^*}}{z_{t+\tau}} + \exp\left(-\Delta \psi_t^*\right) v_{I,t+\Delta}^* \\
&= \Lambda_t\left(\{\psi^*\}, t, t+\tau^*\right) \Delta \left(1+\tau'^*\right) \frac{\alpha l_{P,t+\tau}^*}{a_{Ft+\tau} z_{t+\tau}} + \exp\left(-\Delta \psi_t^*\right) v_{I,t+\Delta}^*
\end{aligned}
$$

Note that when computing the present discounted value from the point of view of the emerging market, the expression changes slightly in the sense that the foreign skill premium is the key price and the foreign relative technology level shows up in the denominator. Note that the frontier level of technology is global

$$
\begin{aligned}
\frac{s^* h_F^{*1-\lambda}}{\gamma} = \; & \Lambda_t\left(\{\psi^*\}, t, t+\tau^*\right) \Delta \left(1+\tau'^*\right) \frac{\alpha l_{P,t+\tau}^*}{a_{Ft+\tau} z_{t+\tau}^*} + \exp\left(-\Delta \psi_t^*\right) v_{I,t+\Delta}^* \\
& + \Lambda_t\left(\{\psi\}, t, t+\tau\right) \Delta \left(1+\tau'\right) \frac{\alpha l_{P,t+\tau}}{a_{Ft+\tau} z_{t+\tau}^*} + \exp\left(-\Delta \psi_t\right) v_{I,t+\Delta}
\end{aligned}
$$

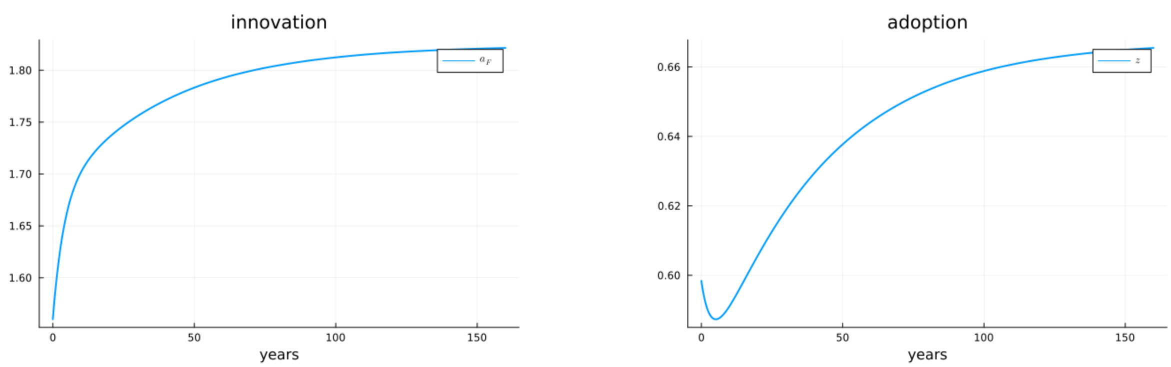
Focusing on the advanced economy, for a sequence of skill prices etc, one can invert the right hand side to obtain equilibrium demand for normalized skilled labor. Then update the price of skill gently to get at something reasonable. Note that this system depends on what happens in the foreign economy, everything is related!

B.6 Convergence Dynamics when Skill Ratio Expands in Closed Economy

The results here are preliminary and might be revised for a later version of the paper. First, I am going to consider a positive shock to the skill-ratio ($h^{tot} \uparrow$). I consider an increase from .13 to .15, which means that the relative supply of skilled labor increases by roughly 15%. In this model, this leads to powerful medium-term growth effects that fuel both innovation and technology adoption. As argued in the main part of the paper, adoption and innovation are complementary, and both activities interact with each other in a way to lead to long-run growth dynamics. I keep all other parameters the same as in the baseline calibration in the paper. The virtues cycle between innovation and adoption is consistent with the account of rising skilled labor shares in Goldin and Katz (2010) and its impact on economic growth in the US after WW2.

Figure 13 shows how this shock to the supply of skill leads to an immediate response in the forward-looking innovation sector. The normalized measure ($a_F = \frac{A_F^{1-\phi}}{L}$) expands. This means that the adoption gap initially widens, but it closes over time and in the long run z increases to a higher steady state. Most adjustments happen in the first 50 years, but I plot out the response 150 years in.

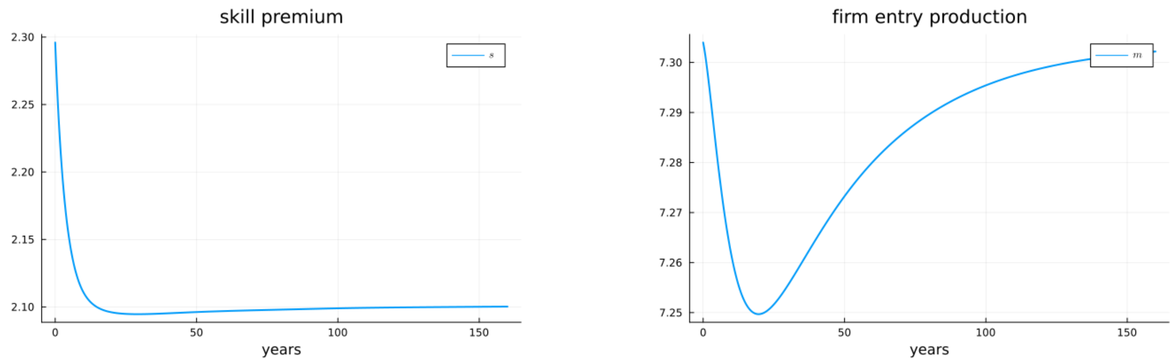
Figure 13. Innovation and Adoption in Closed Economy



Time units are in $\Delta = 0.1$.

Consistent with this trajectory is that the price of skill is still relatively high initially, but it converges quickly to a lower level. Note how the measure of firms endogenously shrinks when adoption effort is high, while it expands back to its long run level that is independent of the price of skill or the level of frontier technology.

Figure 14. Innovation and Adoption in Closed Economy



Time units are in $\Delta = 0.1$.

C Extensions

C.1 Immigration

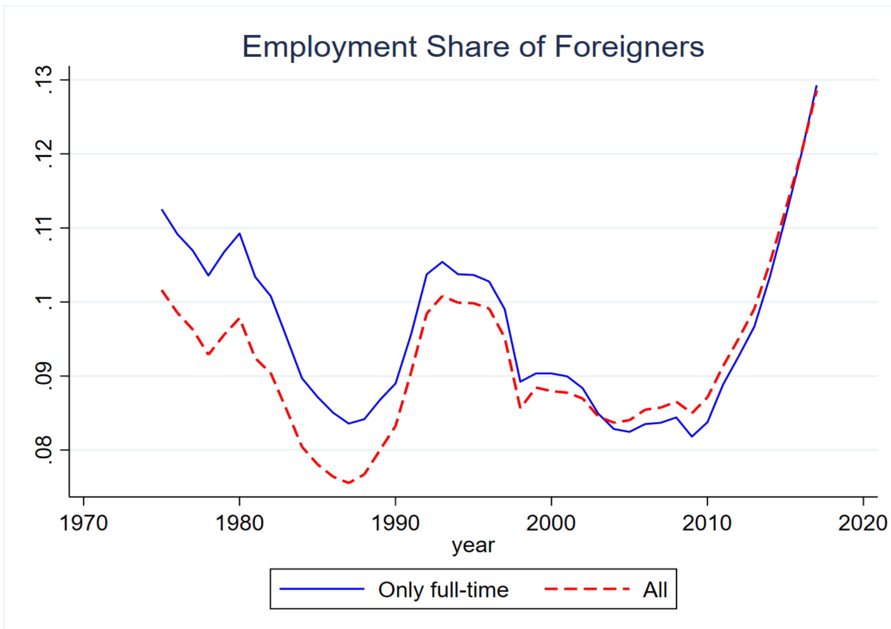
A fully integrated equilibrium behaves differently from the baseline model. Note that the factor price equalization theorem does not hold precisely because countries have different research productivities so goods market trade is no substitute for immigration. World output would be maximized by moving 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 (41). 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. 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.

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.

Figure 15 helps us assess the relevance of immigration into Germany as a potential reason for weak wage growth. It turns out that the foreign employment share is falling since the mid 1990s, leading to an all-time low in the 2000s. This figure suggests that immigration is of second-order with regard to wage stagnation in Germany. This is consistent with the numerous micro studies on the labor market effect of immigration cited in the main text.

Figure 15. Immigration



The figure plots the share of foreign workers in West Germany, using the BHP of the IAB.

C.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 $2^{\frac{1}{1-\phi}} - 1 \approx 40\%$ 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.

C.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 the 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 (41) 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 $h(1) = 0$.[104] Again, such a model offers more scope for production labor to gain from market integration.

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

Part I

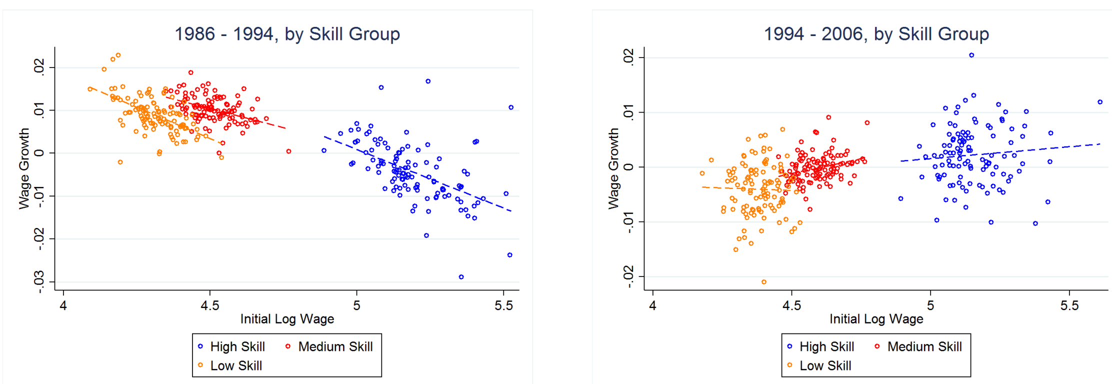
Empirical Appendix

D 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 109 regions.

E Changing Convergence Dynamics

Figure 16. Convergence by Skill Group



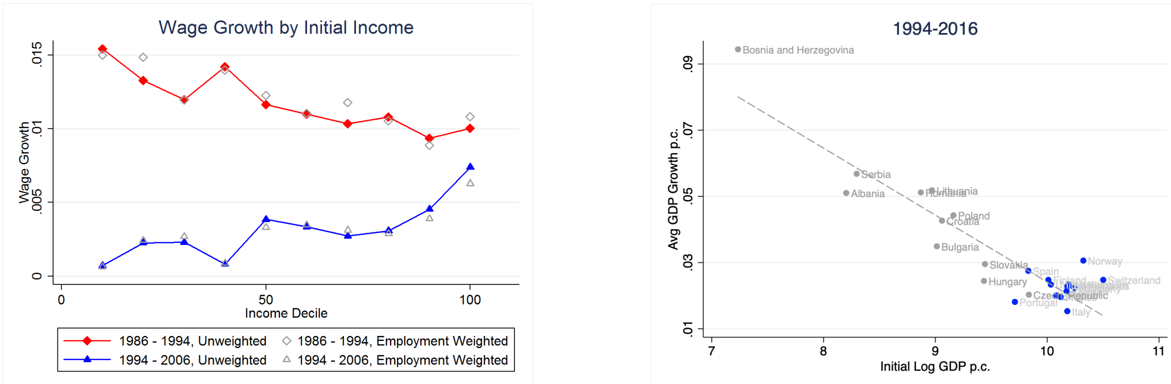
IAB BHP data. My plots.

F 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

Figure 17. Regional Divergence and Global Convergence



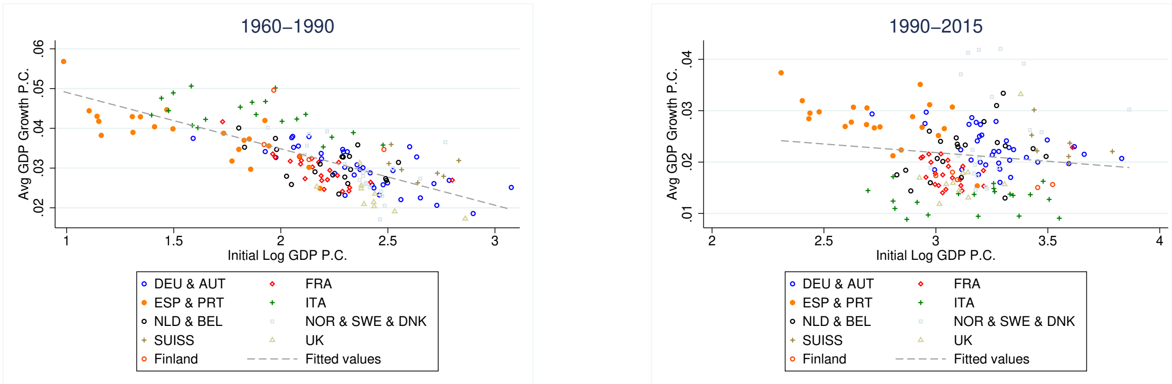
The left panel is based on the BHP dataset of the IAB. Regions are defined as local labor markets following Kosfeld and Werner (2012) which implies that there are 109 local labor markets in West Germany, each of which is assigned to a wage decile based on the average wage in the base period. The right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

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, $\Delta\,(Import\,exp)^{East}_{j,t} = \sum_j \frac{E_{j,i,t}}{E_{i,t}} \frac{\Delta Im^{D \leftarrow East}_{i,t}}{E_{j,t}}$, where $\Delta Im^{D \leftarrow East}_{i,t}$ 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 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

Figure 18. 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.

German industry classification, see their paper for details.

G Additional Information on Patent Data

TO BE COMPLETED.

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.

H Additional Information on Aggregate Wages and Employment

TO BE COMPLETED.

H.1 Sectoral Classification and Rising Skill Share in Research Sector

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. In the following plot 20 I show how the skill share in the research sector diverges from the skill share in the production sector.

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 22, 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 5. 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.[105]

Table 5. 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	**-0.0211**	.00558	**0.0109**	.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	**0.0178**	.00750	109
7.	import exposure (shift share)	**NA**	NA	**0.0154**	.00312	109
8.	export exposure (shift share)	**NA**	NA	**0.0120**	.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).

I Patents and Growth

TO BE COMPLETED.

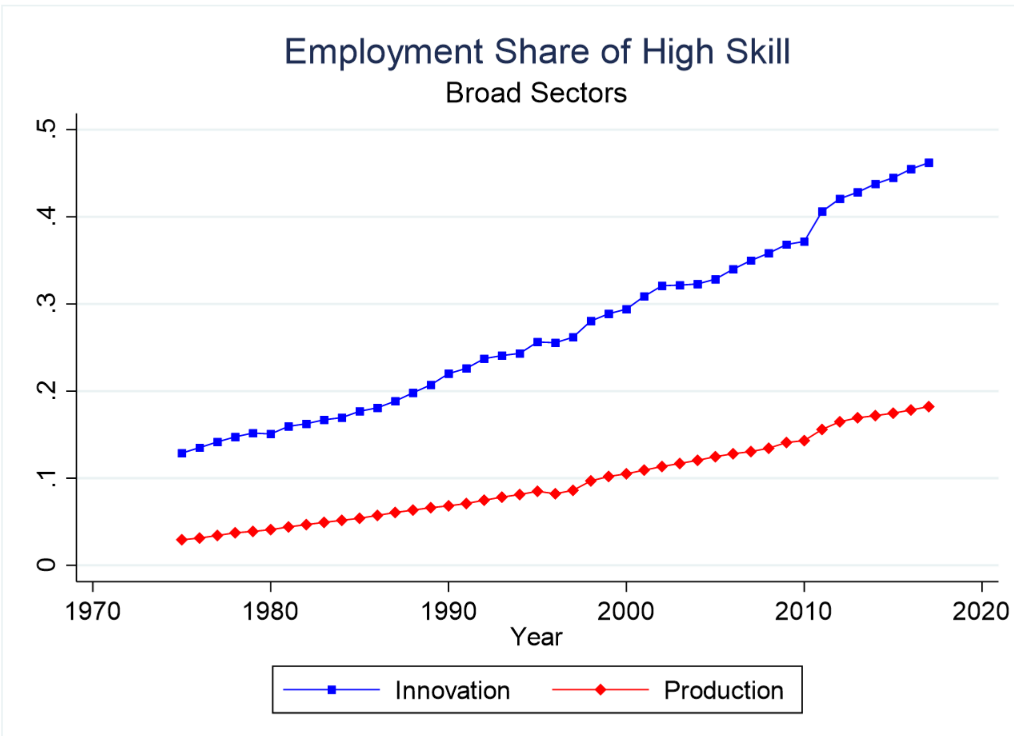
[105] Autor, Dorn, and Hanson (2013) focus on the employment margin of the China shock, and do not find strong wage effects. In the simple cross-sectional setting I use here, and without using their instrument, wage growth is positively related to both import and export exposure.

Figure 19. 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 v	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 und	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 Finance and IT
3-Steller	(w93_3)	651	Zentralbanken u. Kreditinst.	0	1
3-Steller	(w93_3)	652	So. Finanzierungsinstitute	0	1
3-Steller	(w93_3)	722	Softwarehaeuser	0	1
3-Steller	(w93_3)	723	Datenverarbeitungsdienste	0	0
3-Steller	(w93_3)	724	Datenbanken	0	0
3-Steller	(w93_3)	726	Verb Ttg. der Datenverarb.	0	0
3-Steller	(w93_3)	731	F&E Naturwissenschaft	1	1
3-Steller	(w93_3)	732	F&E Recht, Wirtschaft usw.	1	1
3-Steller	(w93_3)	741	Beratungsunternehmen	1	1
3-Steller	(w93_3)	742	Architektur- u. Ingenieurbuero	0	0
3-Steller	(w93_3)	743	Tech., physik. u. chem.	0	0
3-Steller	(w93_3)	744	Werbung	0	0
5-Steller	(w03_5)	74131	Marktforschung	1	1
5-Steller	(w03_5)	74132	Meinungsforschung	0	0
5-Steller	(w03_5)	74141	Unternehmensberatung	1	1
5-Steller	(w03_5)	74142	Public-Relations-Beratung	1	1
5-Steller	(w03_5)	74151	Managementtaetigkeiten von 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 c	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 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. Wage Stagnation in Germany over the long run

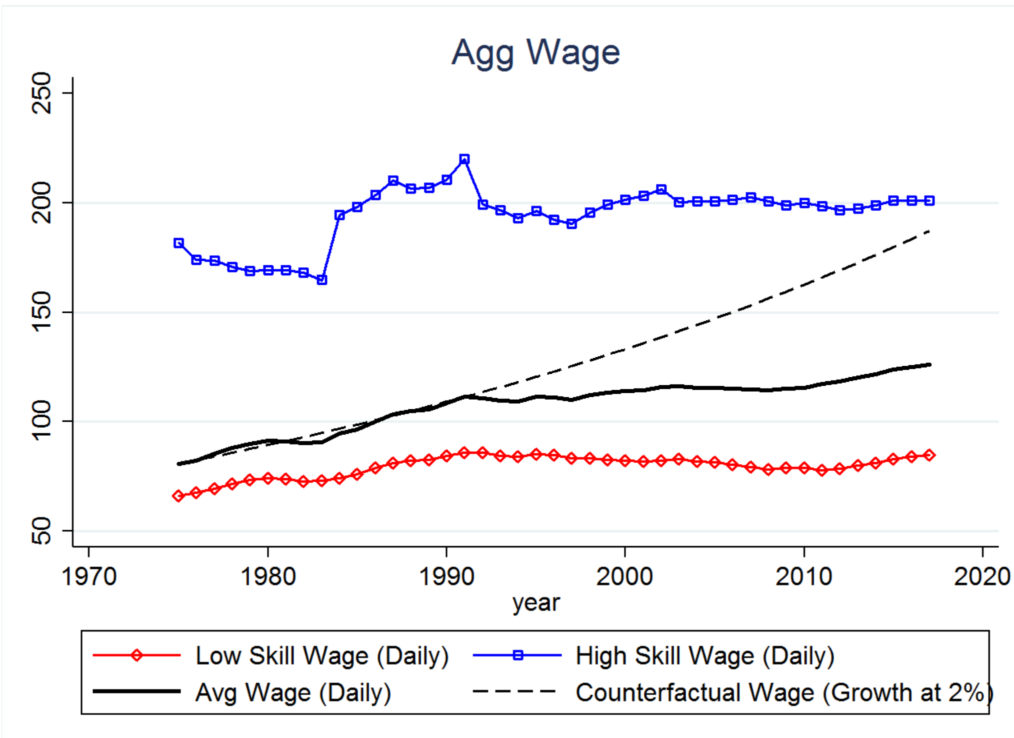
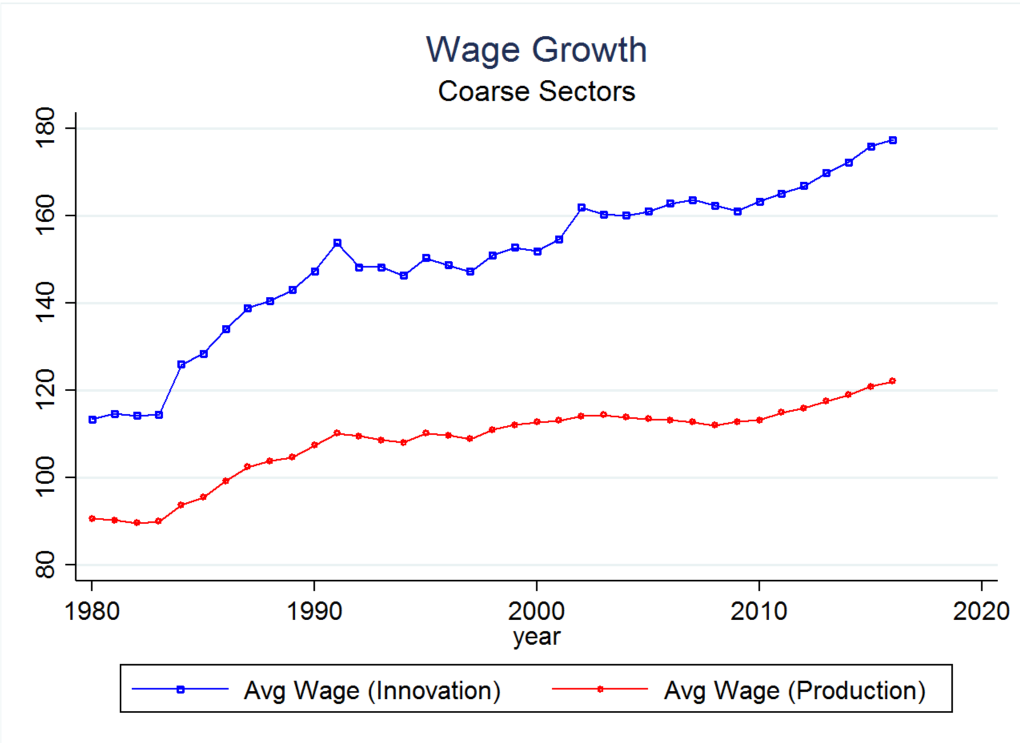


Figure 22. Wages in Innovation and Production



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