Essays on Growth, Globalization, and Inequality

by

Florian Trouvain

A dissertation submitted in partial fulfillment
of the requirements for the degree of
Doctor of Philosophy
(Economics)
in the University of Michigan
2023

Doctoral Committee:

Professor John V. Leahy, Co-Chair
Professor Dmitriy Stolyarov, Co-Chair
Professor Linda L. Tesar
Professor Xun (Brian) Wu

Florian Trouvain

ftrouv@umich.edu

ORCID iD: 0000-0001-5719-7367

ACKNOWLEDGEMENTS

I thank my committee members John Leahy (Co-Chair), Dmitriy Stolyarov (Co-Chair), Linda Tesar, and Brian Wu for patient advice, generous insights, and pivotal encouragement. I learned most of what I know about economics from these outstanding scholars, to whom I owe a huge debt.

I thank Dominick Bartelme, Andres Blanco, Hoyt Bleakly, John Bound, Paco Buera, Josh Hausman, John Laitner, Pablo Ottonello, Yongs Shin, and Sebastian Sotelo for generous advice.

I thank my wife for her unwavering support, trust, and love. I thank my brother for looking after me always.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

LIST OF APPENDICES

ABSTRACT

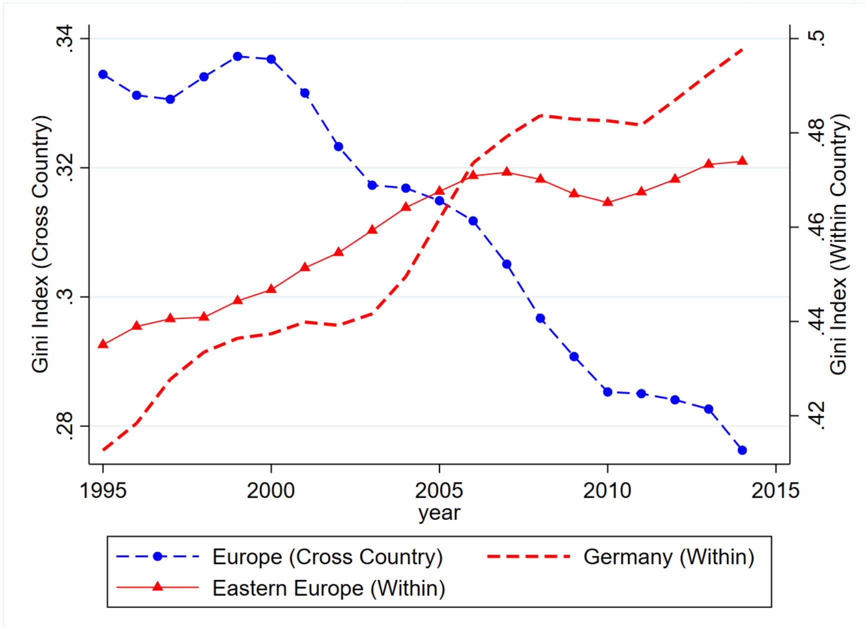
This set of research papers focuses on the interaction between growth, globalization, and inequality. The first chapter proposes a semi-endogenous growth model with skill-intensive technology adoption. The model addresses a number of important issues, and in particular, how globalization and market integration shape growth and inequality. In the second chapter, I take growth and inequality as given and study how growth & inequality shape aggregate patterns of structural change and household savings in a dual economy setting with a rural and an urban sector. The final chapter develops an incomplete market model with intergenerational human capital risk to study how the distribution of aggregate growth between high and low-income households during an episode of fast industrialization in emerging markets impacts aggregate savings, investment, and consumption.

Chapter I. Technology Adoption, Innovation, and Inequality

1.1 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]

[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 P. M. Romer (1990) or Jones (1995) and instead focus on the allocation of skilled labor devoted to innovation relative to

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 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.

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 P. M. 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 P. M. 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

[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.

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 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 P. M. 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, Jones, 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 P. M. 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 (2021a) 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, Akcigit, 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-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 S. L. 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 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 S. 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, Bergeaud, Boppart, 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 P. M. Romer (1991), Grossman and Helpman (1991a), Grossman and Helpman (2018),

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

[5]This is most easily seen in a version of the model of Jones (1995) without population growth. The assumption that ideas are getting harder to find leads to prohibitively high entry cost into research in the long-run so that share of resources devoted to innovation asymptotes to zero.

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 Marti 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, Ramondo, 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 A. B. 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 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

[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 S. 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 S. Redding (2010), R. Liu and Trefler (2008), Sampson (2014), or A. Burstein and Vogel (2017).

[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 D. Autor (2011) for a discussion.

literature has focused on the task content of work and automation (D. H. 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 1.2 presents a model of innovation and adoption. Section 1.3 introduces the open economy version. Section 1.4 offers a quantitative exercise after calibrating and estimating key parameters of the model. Section 1.5 provides empirical evidence to support the central mechanism. Section 1.6 concludes.

1.2 A Tractable Theory of Innovation and Adoption

1.2.1 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

$$\begin{aligned} &\max_{\{c,B\}} \int_0^\infty e^{-(\rho-g_L)t} \log c_t \, dt \\ &\text{s.t. } \dot{B} = rB - C. \end{aligned} \tag{1.1}$$

[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.

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 (1.2) where capitalists' per capita consumption grows at rate

$$\frac{\dot{c}}{c} \;=\; r - \rho. \tag{1.2}$$

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

$$Y \;=\; L^{-\delta_Y} \left(\int_{\Omega_M} (q_i)^{\frac{\sigma-1}{\sigma}} di \right)^{\frac{\sigma}{\sigma-1}} , \tag{1.3}$$

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

$$\dot{K} \;=\; Y - \tilde{C} - \delta_k K, \tag{1.4}$$

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

[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.

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

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

$$q_i \quad = \quad \left(\int_{j\in\Omega_{A_i}} \left(\frac{x_{ij}}{\alpha}\right)^{\alpha} dj\right)\left(\frac{l_i}{1-\alpha}\right)^{1-\alpha} \; . \tag{1.5}$$

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

[12] I will establish a link between available capital goods and innovation following P. M. 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.

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

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

$$mc_i = (p_x)^{\alpha} \left(\tfrac{w}{A_{iK}} \right)^{1-\alpha}. \tag{1.6}$$

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

$$p_i = \frac{\sigma}{\sigma - 1} mc_i, \tag{1.7}$$

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

$$
\begin{aligned}
w l_i &= \tilde{r}_i \tfrac{\sigma-1}{\sigma} (1-\alpha) \\
p_x \tilde{x}_i &= \tilde{r}_i \tfrac{\sigma-1}{\sigma} \alpha,
\end{aligned} \tag{1.8}
$$

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

$$\pi_i^o = \tfrac{\tilde{r}_i}{\sigma}. \tag{1.9}$$

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

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

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

$$\dot{A}_{iK} \;=\; \zeta A_F^{1-\theta} A_{iK}^{\theta} h_i^{\beta} - A_{iK}\delta_I, \tag{1.10}$$

where $\theta \in (0,1)$, $\zeta > 0$, and $\beta \in (0,1)$. The law of motion is similar to S. Parente and Prescott (1991), R. E. J. Lucas (1993), and more recently 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 discussed below. Note that (1.10) 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,

$$(r_t + \delta_{ex}) V(A_{iK}, t) \;\; -\dot{V} = \;\; \max_{h_i} \pi_t^o (A_{iK}) + \partial_{A_{iK}} V(A_{iK}, t) \left[\dot{A}_{iK}\right] \; - w_H h_i. \tag{1.11}$$

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

[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.

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

$$f_e w \geq V(A_K, t). \tag{1.12}$$

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

$$\dot{M} = \tfrac{L_E}{f_e} - M\delta_{ex} \tag{1.13}$$

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 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.1.6 I consider such a setting.[20]

[17]This argument has been made in Schumpeter (1942) and P. M. 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 1.3.1, together with several other extensions.

[19]See Luttmer (2007), R. E. J. Lucas (2009), 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.

Innovation: Innovators expend skilled labor to add novel technology to the stock of ideas, following P. M. 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

$$\dot{A}_F = \frac{1}{f_R}A_F^{\phi}H_F - \delta_I A_F, \tag{1.14}$$

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)

$$f_R = \frac{H_F^{1-\lambda}}{\gamma} \tag{1.15}$$

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

$$V_I A_F^{\phi} \leq f_R w_H. \tag{1.16}$$

where (1.16) 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 P. M. 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 profits

$$V_I = \int_{t+\tau_t}^{\infty} \exp\left(-\int_t^u (r_v + \delta_I)\, dv\right) \pi_{Iu} du \tag{1.17}$$

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

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 P. M. 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 $c\left(.\right)$ 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

$$p_{x_j} = \frac{1}{\alpha}\left(r+\delta_k\right) \quad \forall j, \tag{1.18}$$

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

[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 P. M. 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.

constant share of total revenue divided by the total measure of active ideas, which in turn is proportional to the wage bill in the economy due to Cobb-Douglas production[24]

$$\pi_I = \frac{\alpha L_P w}{A} \tag{1.19}$$

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 $W(t,t) = A_F - A$,

[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.

a trajectory of A_t that the innovators takes as given, and the differential equation

$$\dot{W} \;=\; -\delta_I W - A\left(\delta_I + g_A\right) \;. \tag{1.20}$$

1.2.2 Equilibrium Concept

I define an equilibrium on the balanced growth path of this semi-endogenous growth model as follows. A balanced growth path equilibrium, with constant population growth $g_L = g_H$ and $\phi < 1$, consist of a sequence of prices $\{w_t, w_{Ht}, r_t, p_{xt}, p_{it}, V_t, V_{It}\}$ and allocations $\{L_{Pt}, L_{Et}, H_{Dt}, H_{Ft}, X_t, K_t, M_t, A_t, A_{Ft}, C_t\}$ 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 (1.5) and (1.11) 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.

1.2.3 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 (1.11), 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

$$v\left(r+\delta_{ex}-g_w\right)=\max_h \frac{\pi_t\left(z\right)}{w}-sh+\left(\partial_z v\right)\dot{z}+\dot{v} \tag{1.21}$$

$$\text{s.t.} \tag{1.22}$$

$$\dot{z}=\zeta z^{\theta}h^{\beta}-\left(g_F+\delta_I\right)z, \tag{1.23}$$

where $\frac{\dot{A_F}}{A_F}=g_F$, and $\dot{z}=\dot{v}=0$ in the steady state. A solution to (1.21) needs to satisfy the first order condition

$$\left(\partial_z v\right)\beta\zeta z^{\theta}h^{\beta-1} \quad = \quad s. \tag{1.24}$$

Equation (1.24) 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

$$\frac{\dot{h}}{h} \;=\; \frac{1}{1-\beta}\left\{\underbrace{\rho+\delta_{ex}}_{\text{effective discounting}}+\underbrace{\left(1-\theta\right)\left(g_F+\delta_I\right)}_{\text{effective depreciation}}-\underbrace{\left\{\frac{\beta z^{\theta}\zeta h^{\beta-1}}{s}\left[\frac{\pi_t}{w}\frac{\left(1-\alpha\right)\left(\sigma-1\right)}{z}\right]+\frac{\dot{s}}{s}\right\}}_{\text{marginal benefit of extra unit of skilled labor}}\right\}. \tag{1.25}$$

Equation (1.25) is similar in spirit to the well-known q-theory of investment, and I show the mathematical equivalence in the appendix A.1. 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

[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.

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. 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 1.2.3 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

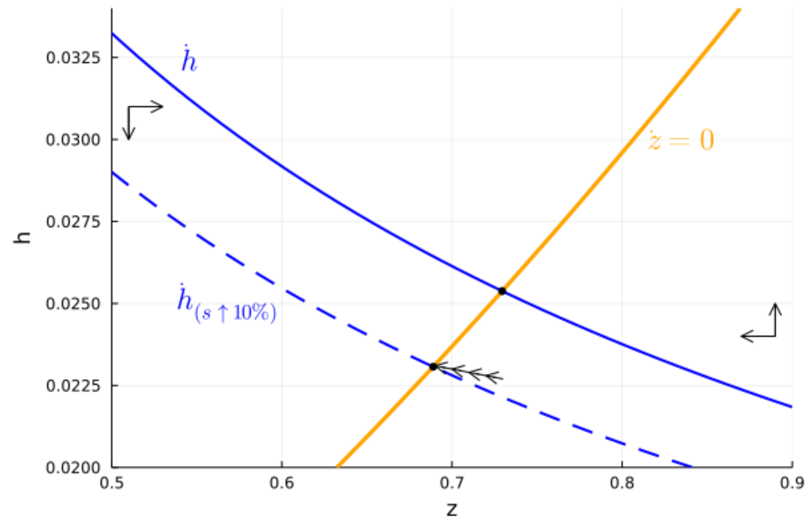
$$h \;=\; \tfrac{1}{s}\tfrac{\beta(1-\alpha)(\sigma-1)(g_F+\delta_I)}{\rho+\delta_{ex}+(1-\theta)(g_F+\delta_I)}\left[\tfrac{\pi}{w}\right]. \tag{1.26}$$

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]

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 (1.25) 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.

[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.

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

$$v \;=\; \frac{\frac{\pi}{w} - sh}{r + \delta_{ex} - g_w} \tag{1.27}$$

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 (1.27) and (1.26) the normalized operating profits on the balanced growth path are pinned down

$$f_e\,(\rho + \delta_{ex})\,\kappa_2 = \frac{\pi}{w}, \tag{1.28}$$

where I used $r = \rho + g_w$. Equation (1.28) 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

[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 1.2.3 holds, $\kappa_2 > 1$ will hold as well.

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

$$f_e\left(\rho+\delta_{ex}\right)\kappa_2 \quad = \quad \frac{l_P}{m}\frac{1}{(1-\alpha)(\sigma-1)}. \tag{1.29}$$

Together with the normalized firm entry resource constraint

$$\dot{m} \quad = \quad \frac{l_E}{f_e} - \left(\delta_{ex}+g_L\right)m, \tag{1.30}$$

the steady state normalized measure of firms reads

$$m \quad = \quad \frac{1}{f_e\left[(\rho+\delta_{ex})(1-\alpha)(\sigma-1)\kappa_2+g_L+\delta_{ex}\right]}. \tag{1.31}$$

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 (1.21) with the firm's demand for skill (1.26). Taking logs leads to

$$\log z \quad = \quad -\frac{\beta}{1-\theta}\log s + \frac{1}{1-\theta}\log\left(\frac{\zeta}{(g_F+\delta_I)}\left(\frac{\pi}{w}\kappa_1\right)^{\beta}\right). \tag{1.32}$$

Expression (1.32) 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 $(1-\theta)$ 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

[33] This is an implication of (1.27) 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.

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

$$V_I \quad = \quad \int_{t+\tau}^{\infty} \exp\left(-\int_t^u (r_x + \delta_I)\, dx\right) \pi_{Iu} du \qquad (1.33)$$

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

$$V_I = \exp\left(-\int_t^{t+\tau} [r_u + \delta_I]\, du\right)(1+\tau')\, \frac{\pi_{I,t+\tau}}{r + \delta_I - g_{w_H} - (1-\lambda)\, g_{H_F} + \phi g_F} \qquad (1.34)$$

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

$$\tau \quad = \quad -\frac{\log z}{g_A + \delta_I}. \qquad (1.35)$$

Proof in A.1.2. Intuitively, equation (1.35) takes physical units of productivity $\left(\log \frac{A}{A_F} = \log z\right)$ 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]

$$V_I \;=\; \frac{1}{\tilde{\rho}+g_F+\delta_I}\left(\frac{\alpha L_P w}{A_F}\right) z^{\frac{\tilde{\rho}}{g_A+\delta_I}} \tag{1.36}$$

where $\widetilde{\rho} := \rho - g_L > 0$ is the effective discount factor of the dynastic household and I substituted out τ using (1.35). 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 (1.36) 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,

$$\frac{1}{\gamma} \;=\; \frac{1}{s}\frac{\alpha l_P}{\tilde{\rho}+g_F+\delta_I}\left(\frac{h_F^{\lambda-1}}{a_F}\right) z^{\frac{\tilde{\rho}}{g_A+\delta_I}}. \tag{1.37}$$

Two important economic mechanisms are captured in (1.37). 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

$$\begin{aligned} h_D &= m h_i \\ &= \tfrac{1}{s}\Lambda^D \end{aligned} \tag{1.38}$$

where Λ^D collects elements that are constants in the steady state.[36] Using a normalized version of the law of motion of ideas (1.14), I get

$$\frac{\gamma h_F^{\lambda}}{(g_F+\delta_I)} \;=\; a_F. \tag{1.39}$$

[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.

Combining (1.37) with (1.39) leads to the research sector's normalized demand for skilled labor

$$\begin{aligned} h_F &= \frac{1}{s}\left(\frac{g_F+\delta_I}{\tilde{\rho}+g_F+\delta_I}\right)\alpha l_P \left(z\right)^{\frac{\tilde{\rho}}{\delta_I+g_F}} \\ &= \frac{1}{s}\left(z\right)^{\frac{\tilde{\rho}}{\delta_I+g_F}}\Lambda^F. \end{aligned} \tag{1.40}$$

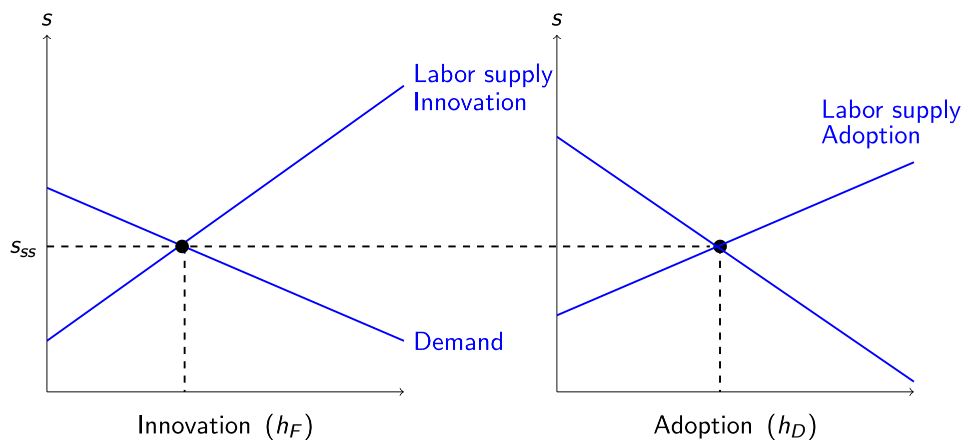
Adding up (1.38) and (1.40) and imposing market clearing, I obtain the following equation that implicitly defines the relative price of skill

$$\left\{\tfrac{1}{s}\left(z\right)^{\frac{\tilde{\rho}}{\delta_I+g_F}}\Lambda^F+\tfrac{1}{s}\Lambda^D\right\} \quad = h_{tot} \tag{1.41}$$

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. 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,

$$Y = \left(\frac{K}{\alpha}\right)^{\alpha} \left(\frac{z A_F L^P}{1-\alpha}\right)^{1-\alpha} . \tag{1.42}$$

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 $k := \frac{K}{zA_F L_P}$ 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

$$\begin{aligned} w &= (1-\alpha) \frac{\sigma-1}{\sigma} \Lambda_\alpha z A_F k_{ss}^\alpha \\ w_H &= sw. \end{aligned} \tag{1.43}$$

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.1.3. 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.

1.2.4 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 (1.37), the partial equilibrium elasticity of the total measure of ideas A_F with respect to the adoption gap in the steady state equals[37]

$$\frac{\partial \log A_F}{\partial \log z} = \frac{\lambda}{1-\phi} \frac{\tilde{\rho}}{g_A+\delta_I}. \tag{1.44}$$

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

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

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 (1.39), 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 (1.38) and (1.40)

$$\frac{H_F}{H_D} = \left(\frac{g_F+\delta_I}{\tilde{\rho}+g_F+\delta_I}\right)\frac{\alpha(\sigma-1)(1-\alpha)}{\kappa_1}(z)^{\frac{\tilde{\rho}}{\delta_I+g_A}}. \tag{1.45}$$

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 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 1.3.1. 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 (1.45). 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.

1.3 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

[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 S. 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, Ramondo, 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.

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 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

$$\frac{w^*}{w} = \frac{z^*}{z}. \tag{1.46}$$

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

$$z_c \propto h_c^{\frac{\beta}{1-\theta}}. \tag{1.47}$$

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 (R. E. J. Lucas (1988), R. E. J. Lucas (2009)) 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 pro-

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.

$$\dot{A}_F = \frac{\left(A_F^W\right)^{\phi} H_F}{f_R} - \delta_I A_F. \tag{1.48}$$

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

$$\frac{h_F}{h_F^*} = \left(\frac{\frac{\gamma}{w_H}}{\frac{\gamma^*}{w_H^*}}\right)^{\frac{1}{1-\lambda}} \tag{1.49}$$

duction, 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.

[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.

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 (1.48), it follows that[44]

$$\left(\frac{\chi}{\chi^*}\right) = \frac{\gamma}{\gamma^*}\left(\frac{h_F}{h_F^*}\right)^{\lambda}. \tag{1.50}$$

Combining this expression with (1.49) and noting that $\frac{w_H}{w_H^*} = \frac{s}{s^*}\frac{z}{z^*}$ leads to

$$\left(\frac{\chi}{1-\chi}\right) \quad = \quad \left(\frac{\gamma}{\gamma^*}\right)^{\frac{1}{1-\lambda}}\left(\frac{s}{s^*}\frac{z}{z^*}\right)^{-\frac{\lambda}{1-\lambda}}. \tag{1.51}$$

Equation (1.51) 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

$$\begin{array}{rcl} \left\{\frac{\chi}{z}\Lambda_{FO}\left((z)^{\frac{\tilde{\rho}}{g_A+\delta_I}+1} + (z^*)^{\frac{\tilde{\rho}}{g_A+\delta_I}+1}\right) + \Lambda_D\right\} & = & sh_{tot} \\ \left\{\frac{\chi^*}{z^*}\Lambda_{FO}\left((z)^{\frac{\tilde{\rho}}{g_A+\delta_I}+1} + (z^*)^{\frac{\tilde{\rho}}{g_A+\delta_I}+1}\right) + \Lambda_D^*\right\} & = & s^*h_{tot}^* \end{array} \tag{1.52}$$

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 (1.52).

---

[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.

[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.

Balanced trade implies

$$\underbrace{\{Y+Y^*\}\frac{\sigma-1}{\sigma}\alpha\left(1-\alpha\right)\left(\chi-\chi^*\right)}_{\text{Innovator Profits/Royalty}}=\underbrace{Y^*-\tilde{C}^*-I^*}_{\text{Final good exports}} \tag{1.53}$$

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 (1.53).[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 1.3 I consider what happens if the research productivity γ of the home economy increases. An increase in the home economy's absolute advantage in research γ, given that $\beta+\theta<1$, leads to an increase in the skill premium in the home economy while the skill premium in the foreign economy falls. Proof see appendix A.1.5. Proposition 1.3 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,

[46]Note that total aggregate profits that accrue to innovators are proportional to total spending on capital goods, i.e. $\sum_c \int \pi_{j,c} dj = \frac{P_x X}{(1-\alpha)^{-1}} + \frac{P_x X^*}{(1-\alpha)^{-1}}$. These are the royalties that are paid each instant, and using the Cobb-Douglas assumption spending on capital goods reads $P_x X = \frac{\sigma-1}{\sigma}\alpha Y$. 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.

[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.

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. 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, without negative distributional effects or changes in the adoption gap. 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 P. Krugman (1980) where

[48]If on the other hand N new countries join the world economy and hav the same skill endowment and

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 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

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 P. R. 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 P. Krugman (1980).

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

$$V_I = \underbrace{\left(\frac{\alpha}{\tilde{\rho} + g_F + \delta_I}\right) \frac{L_P w}{A^F} z^{\frac{\tilde{\rho}}{g_A + \delta_I}}}_{\text{same as closed economy}} \left\{ 1 + \underbrace{\frac{L_P^* w^*}{L_P w} \left(\frac{z^*}{z}\right)^{\frac{\tilde{\rho}}{g_A + \delta_I}}}_{\text{additional market size effect}} \right\}. \tag{1.54}$$

Equation (1.54) 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 (1.52) when using $\frac{w^*}{w} = \frac{z^*}{z}$. The model can easily accommodate countries of different size as equation (1.54) shows, and a larger foreign labor force exerts more pull on innovation in the advanced economy.

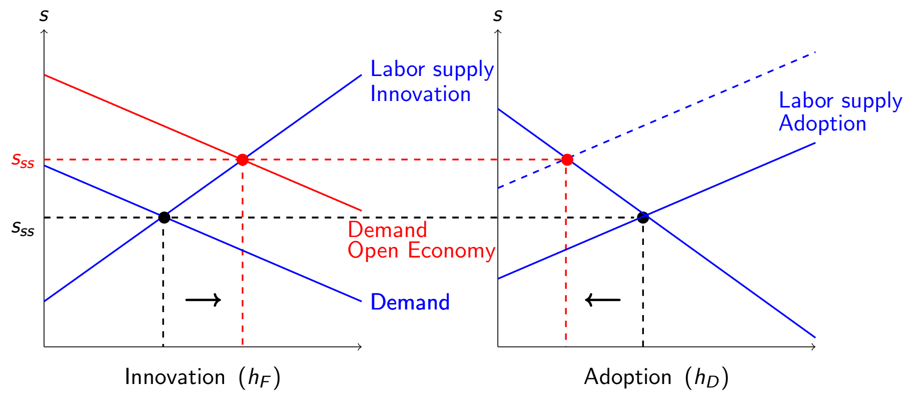
Note that market integration directly increases the market size of innovators, which raises profits that are arbitraged away by increasing entry into innovation. Importantly, convergence in the emerging market further raises the returns to innovation. Note that a rising wage rate ($w^* \uparrow$) and 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 that is capturing the benefits from market integration in advanced economies is skilled labor. Figure 4 summarizes the main argument of this paper

[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.

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 S. L. 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.

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

Second, a related literature has focused on the task content of work (D. H. 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,

$$\frac{1}{s}\left(\tilde{\Lambda}_F z^{\frac{\tilde{\rho}}{g_F+\delta_I}} + \tilde{\Lambda}_D + \tilde{\Lambda}_{\tilde{\beta}}\right) = h_{tot}. \tag{1.55}$$

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

[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 A8 in the appendix.

[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.

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 P. M. Romer (1990)). Slower population growth induces slower productivity growth which requires a smaller share of labor devoted to the production of new ideas.[55] Yet, an increasing share of employment is devoted to research activity, see Bloom, Jones, et al. (2020) for the US, and evidence that I compile for Germany in section 1.4. 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

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

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

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 Ryan A Decker et al. (2017), Ryan A. 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 1.3.1, 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.

1.4 Quantification

1.4.1 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, Jones, 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

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

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

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, Akcigit, 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 $\hat{\beta}_B = -.02$. 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

$$z_c = \left(\frac{\zeta h_c^{\beta}}{g_F + \delta_I}\right)^{\frac{1}{1-\theta}} \tag{1.56}$$

so the real wage in any country is proportional to $h^{\frac{\beta}{1-\theta}}$. Conditional on a distribution of the

[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.

[59]Mankiw, D. 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.

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 (1.56) and adding a measurement error u allows me to back out β by running the following regression

$$\log z_{ct} \quad = \quad \alpha + \delta_t + \tfrac{\beta}{1-\theta} \log h_{ct} + u_{ct}. \tag{1.57}$$

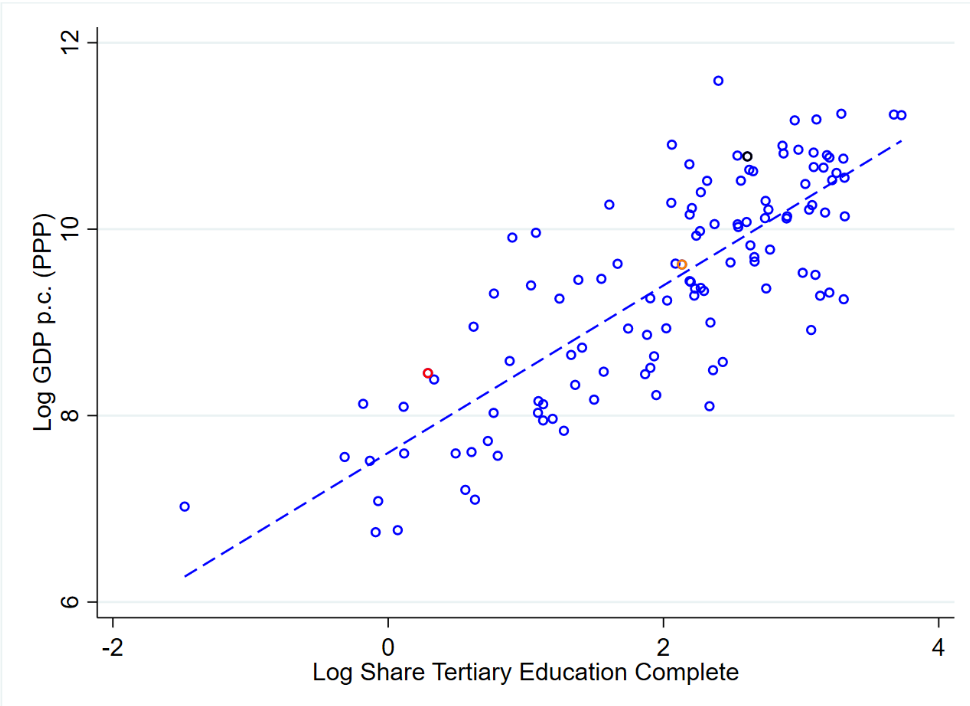
The slope coefficient through the lens of the model equals $\frac{\beta}{1-\theta}$ where I proxy for production worker wages using GDP per capita and I proxy for h using the share of college-educated workers in each country, i.e. h_{tot}. Since most countries don't perform frontier innovation this simplification 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.

[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.

Moreover, I set α to be equal to .5, which is the capital share in production. Once one takes into account that there are overhead labor costs both in terms of production labor for firm entry one arrives at the usual share of capital in total income of 33%.[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, Kaboski, Martí Mestieri, 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 J. G. 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 1.2.3. Precisely

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]

this inequality suggests a smaller σ is appropriate so as to avoid "too much" adoption.

[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.

1.4.2 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 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]

[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. (2019a) 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.

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 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]

[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.

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 strong pro-growth effects found in, for instance, Rivera-Batiz and P. M. 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]

[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 R. E. J. Lucas (2009) and S. S. 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 S. 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.

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 P. M. 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

$$\frac{h_D}{h_F} = \frac{\beta}{1-\theta}\left(1-\phi\right) , \tag{1.59}$$

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] 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

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

$$\begin{array}{ccc} & \max\log\left(A_F z\right) & \\ s.t. & & \\ z & = & \Lambda_z\left(h_D\right)^{\frac{\beta}{1-\theta}} \\ A_F & = & \Lambda_F L^{\frac{1}{1-\phi}}\left(h_F\right)^{\frac{1}{1-\phi}} \\ h_{tot} & \geq & h_D + h_F. \end{array} \tag{1.58}$$

[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.

[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

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 (1.45) 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

$$\frac{h_D}{h_F} = \frac{\beta}{1-\theta}\frac{1}{\alpha}\frac{1}{1+\frac{\delta_{ex}+g_M}{(1-\theta)(g_F+\delta_I)}}, \tag{1.60}$$

which is similar but not identical to (1.59).

The solution (1.60) 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

distribution emerges. See Luttmer (2007), R. E. J. Lucas (2009), Sampson (2016), Buera and Oberfield (2020), or Peters and Walsh (2019).

[76] This limiting case helps clarify the link to S. L. 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.

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 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 $\max{(N+1)} * \log{(A_F)} + \log z + \sum_{j=1,\ldots,N} \log z_j$ and the optimal solution would change to $\frac{1}{N+1}\frac{\beta}{1-\theta}(1-\phi) = \frac{h_D}{h_F}$. 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

[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 A. T. 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.

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.

Figure 6: Wage Dynamics in Advanced Economy

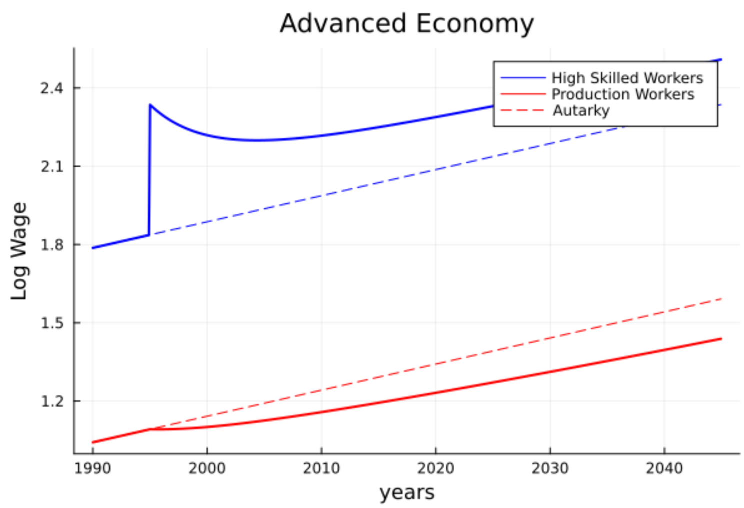
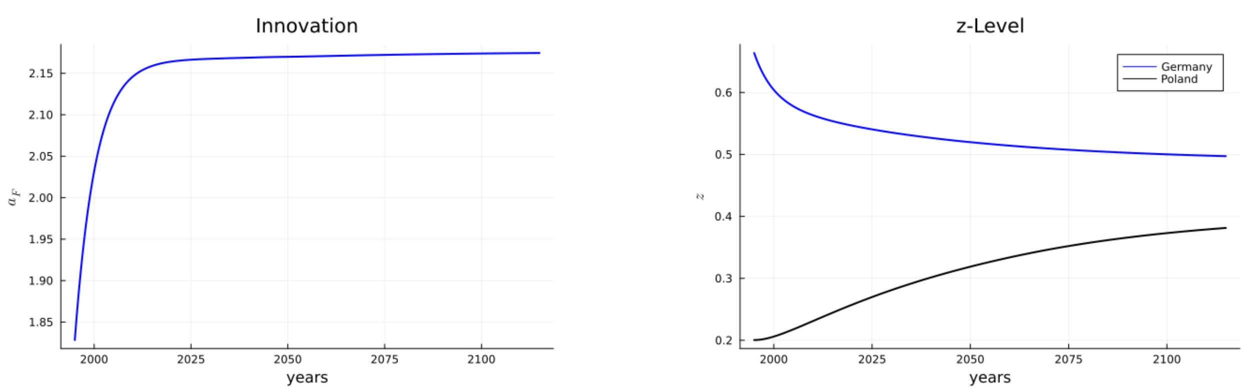


Figure 7: Innovation and Adoption in Open Economy



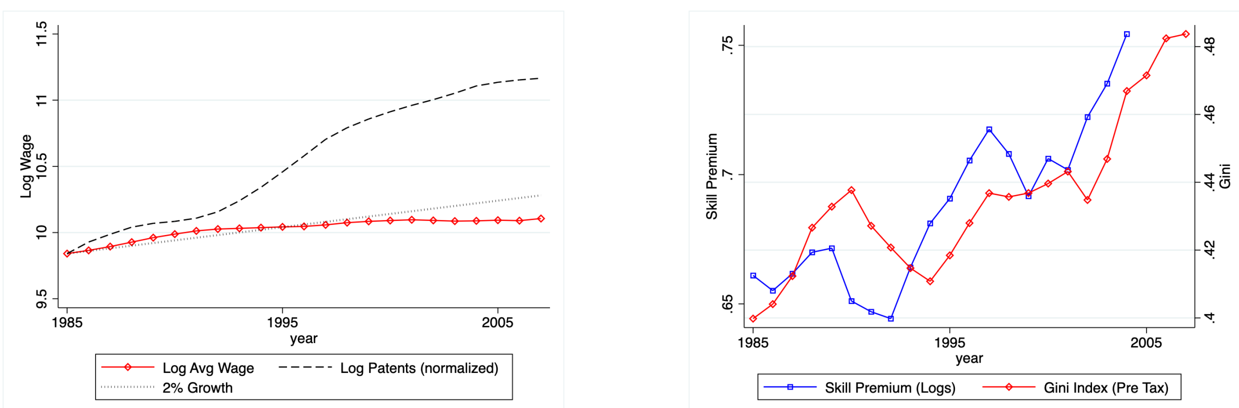
Time units are in $\Delta = 0.1$. Recall the definition of $a_F = \frac{A_F^{1-\phi}}{L}$.

1.5 Empirical Evidence

1.5.1 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, Kaboski, Rogerson, 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

[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. J. Song et al. (2019) find

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 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

similar results in the US, while Haltiwanger, Hyatt, and Spletzer (2022) argue that the industry plays the dominant role in the rise in inequality.

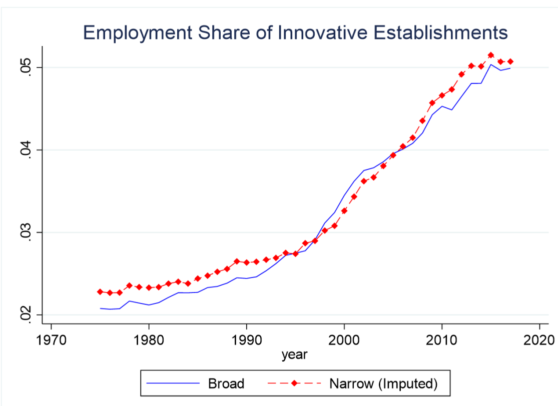
[81] Note that a declining labor share as argued in Karabarbounis and Neiman (2014) is not able to quantitatively account for weak wage growth. Using the KLEMS data, the labor share from 1995 to 2004 fell only from 67.7% to 65.6%. Assuming constant GDP per capita growth of 2%, this would have led to 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).

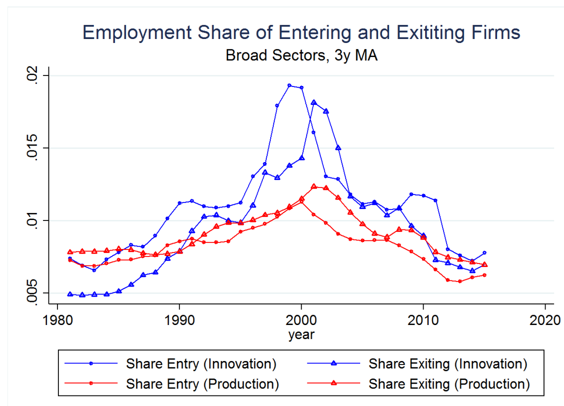
[82] Recent work highlights how patents are used "defensively" to shut down competitors without producing novel content, see Argente et al. (2020). Note that some of this literature is motivated by the fact that patenting activity has not translated into productivity growth. The model at hand provides 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 A8 in the appendix. However, while the

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.

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.

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*

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 D. 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, Kaboski, Rogerson, 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.

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 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.

1.5.2 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 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; D. Autor, Dorn, and Hanson 2013). I focus on West Germany to avoid dealing with the massive institutional change in East Germany after German unification

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

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).

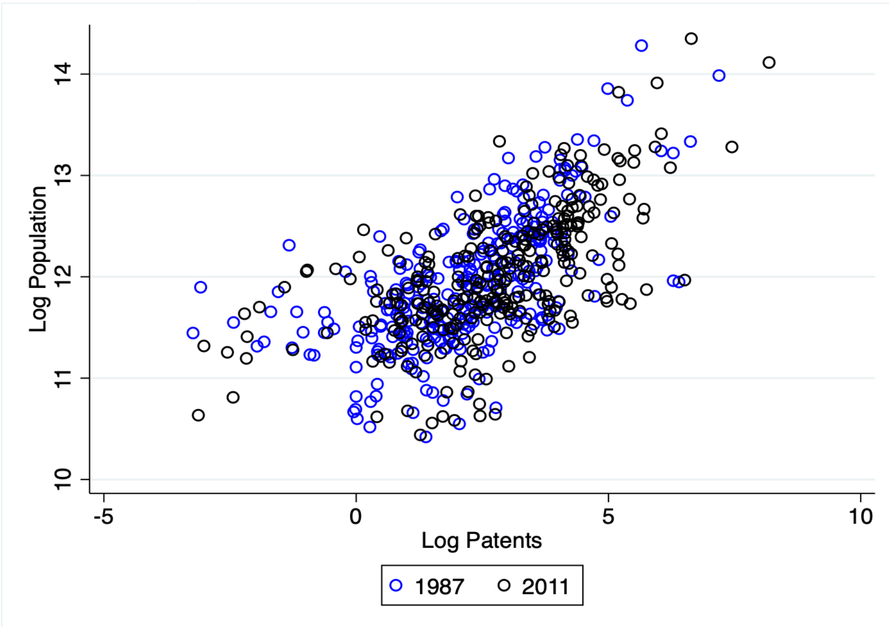
[86] See Findeisen et al. (2021) for work on employment reallocation in the East Germany. Note that most of the convergence within German between East Germany and West Germany occurs up until 1995, see Bachmann et al. (2022). Importantly, while some East German workers did migrate to the West, Findeisen 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 I. P. 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 A3, 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.

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,

$$\Delta_k \log pop_{rt} = \alpha + \gamma_t + \left(\beta + \underbrace{\delta_{t>1995}}_{>0}\right) Patents_{rt} + \log\left(pop_{rt}/area_{rt}\right) + u_{rt}. \tag{1.61}$$

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 (1.61) 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]

Table 3: Innovation and Population Growth

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

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

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

[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, Donaldson, et al. (2019) or Aghion, Bergeaud, Lequien, 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.

[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, Hassan, 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

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]

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-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.

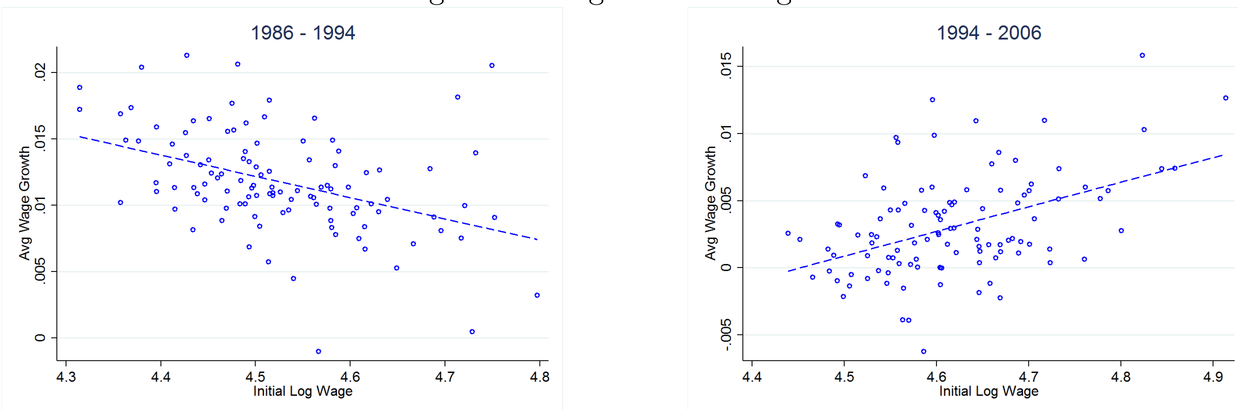
section, the data is very useful to document broad trend breaks in wage growth and employment growth across local labor markets.

[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 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.

Figure 11: Regional Convergence

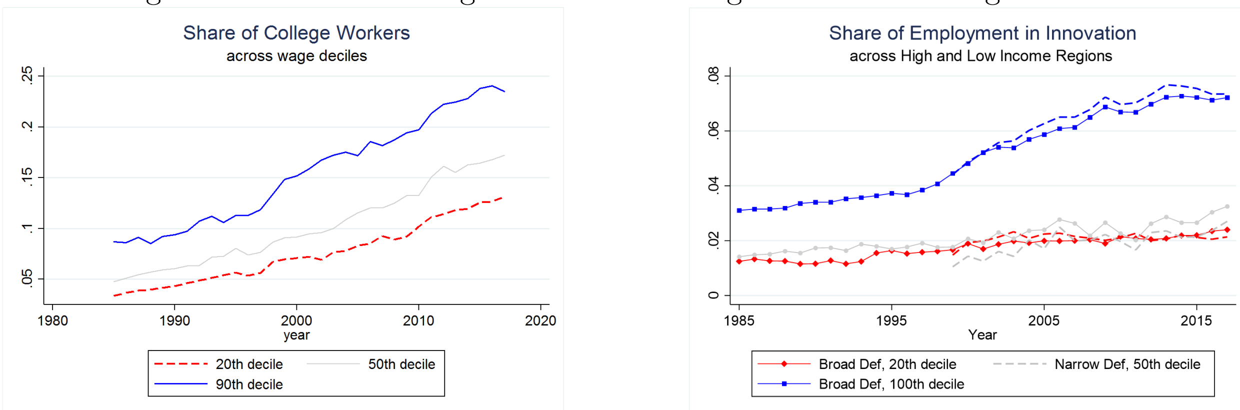


Using data from the BHP establishment sample, the figure plots average wage growth against initial the initial average wage in real terms. The plot shows how growth pre 1994 was biased towards lagging regions, while from 1994 onwards growth was biased towards high income regions. I stop short of the 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 D. Autor, Dorn, and Hanson (2013), fully explains weak growth in laggard regions. To consider the effect of import exposure on wage growth, I run a convergence regression with an additional import exposure variable as 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 a globalized world explains these changing growth patterns across space and workers all at once.

Table 4: Wage Growth & Total High Skill Employment Growth

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

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.

1.6 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.

Chapter II. Structural Change, Inequality, and Capital Flows

2.1 Introduction

In a seminal article R. E. J. Lucas (1990) asks: "Why doesn't capital flow from rich to poor countries?" Worse even, from the standpoint of neoclassical theory, is that capital tends to flow out of fast-growing emerging markets into slow-growing advanced economies. Prominent examples for the combination of strong growth and capital outflows include Taiwan, Japan, Germany, Korea, Singapore, Hongkong, and most recently China. The fact that capital flows out of fast-growing emerging markets constitutes a major puzzle since it seemingly contradicts the cornerstone on which modern macroeconomics is built: the permanent income hypothesis (PIH). The PIH implies that households that are relatively poor today, say compared to the US, but grow relatively fast (and hence will be relatively rich in the future) should smooth consumption by running current account deficits during their catch-up phase. Alas, several studies have documented how this prediction has failed (Hausmann, Pritchett, and Rodrik 2005; Gourinchas and Jeanne 2013; Prasad, Rajan, and Subramanaian 2007).

This paper proposes a novel theory where structural change from traditional rural production into modern human-capital intensive sectors generates household saving pressure during a growth miracle. The key insight of the model is that unevenly distributed income in modern productive activity together with ex-ante uncertainty of a household's position on this evolving distribution can lead to very powerful precautionary savings as the economy is transitioning. In fact, depending on the degree of urban inequality, this precautionary motive can dominate the consumption smoothing force despite miraculously fast aggregate income growth.

In a first step, I document salient differences in savings behavior across rural (agricultural) and urban (non-agricultural) households. In particular, I focus on differences in asset-to-income ratios and asset growth rates for Chinese households, which are systematically higher for urban residents in the Chinese Family Panel Study (CFPS) (Xie and Hu 2014). This is true for both total assets, as well as for a more narrow subset of "safe" assets where I exclude housing and other productive assets. In combination with fast-paced structural change out of urban or agricultural production, a first order feature of growth miracles, these differences can help rationalize the surprising built-up of aggregate savings and safe assets during an episode of fast catch-up growth. China is a case in point: while roughly 80% of workers were employed in agricultural activity in the early 80s, this number dropped to less than 30% in 2020. To my knowledge, this is the first paper that highlights the potential of urban-rural differences in the demand for assets and its interplay with fast-paced structural change to explain the positive association between savings, capital outflows, and catch-up growth.

In a second step, I propose a tractable theory that rationalizes urban-rural differences in savings behavior during the catch-up phase of the economy. Importantly, the model features a growth miracle in the urban sector, which, on its own would leads to low saving rates for urban households along the transition path driven by the standard consumption smoothing motive.[96] Building on the neoclassical tradition, households in my economy are infinitely lived, have perfect foresight regarding the aggregate trajectory of the economy, and feature standard preferences of the CRRA type. I depart from the benchmark neoclassical model in two crucial ways.

First, I use a two-sector setting where human capital risk is larger in the modern sector compared to traditional rural production. Even though rural production may be risky, for instance because of its dependence on weather conditions, workers face little uncertainty about the value of their human capital. Their productivity is tied to their physique as well as access to land. In contrast, modern productive activity with highly specialized human capital tends to yield very uneven outcomes for ex ante similar workers. I introduce this urban human capital risk in the form of an "inequality shock" that works like a draw from a lottery pushing effective human capital up or down. To the extent that households have imperfect knowledge of their productivity in non-agricultural production, which seems intuitive during an episode of fast structural change, ex post inequality in the urban sector is going to represent ex ante risk leading to a strong precautionary savings motive.

Second, a key difference to the benchmark neoclassical framework as well as the canonical incomplete market models of Carroll (1997) or Kaboski and Townsend (2011), is that catch-up growth itself is unevenly distributed across households. I introduce catch-up growth in the urban sector, where entering households experience fast income growth for a random time interval. I assume that households are pulled out of this fast-growth regime according to a Poisson process, which has several desirable features. The household problem becomes

[96] See the recent paper by Coeurdacier, Rey, and Winant (2019) for a quantitative model that features consumption smoothing along the transition path.

extremely tractable since it delivers a structure similar to the perpetual-youth model of Yaari (1965) and Blanchard (1985). Moreover, it leads to a thick-tailed income distribution which is very useful to quantitatively account for the capital flow puzzle, while also being consistent with empirically observed income distributions. This "uneven" growth helps a great deal because it substantially reduces the expected lifetime income growth along the transition path, which is the force in the benchmark model that induces borrowing. To see this, consider standard CES preferences with a very large coefficient of relative risk aversion. In that case, expected utility is mostly informed by the worst growth path which may be substantially below the aggregate (average) growth path. Loosely speaking, the rising tail inequality provides much aggregate catch-up with relatively little consumption-smoothing motive since risk-averse households heavily discount the possibility of landing somewhere on the very right tail. All this idiosyncratic risk averages out conveniently in the aggregate and leads to smooth and strong catch-up growth.

The main result of the theoretical section is a simple sufficient statistic where the trade-off between consumption smoothing on the one hand, and precautionary savings on the other, is pinned down by primitives of the model. While the model is stylized, it allows for sharp predictions and clear insights into the relationship between growth, human capital risk, and savings. It highlights the potential of urban-rural differences, structural change, and human capital risk to account for one of the most persistent puzzles in international macroeconomics.

Another strength of the model is that it can give rise to hump shaped saving rates along the transition path. Representative agent models fail to match this feature of the data, and the literature has resorted to explanations based on habit in consumption (Carroll, Overland, and Weil 2000).[97] Even if one were to consider a closed economy setting, it is difficult to generate hump-shaped saving rates in the benchmark neoclassical model. It is tempting to argue that fast productivity growth could lead to high saving rates since the marginal product of capital is rising (a substitution effect). In a closed-economy version of the neoclassical model with CRRA preferences, however, aggregate saving rates are inversely related to productivity growth, given estimated elasticities of intertemporal substitution well below unity (Hall 1988), leading to a falling aggregate saving rate. The income effect simply dominates the substitution effect.[98] In the model at hand the importance of the

[97] A standard explanation for hump-shaped saving rates is habit in consumption. Yet, empirical studies based on micro level data find that habit in consumption is at odds with actual household consumption choices (M. Chamon, K. Liu, and Eswar Prasad 2013).

[98] See the unpublished manuscript by Antras (2001) for a solution to this puzzle based on non-standard preferences and production technology.

precautionary motive in the aggregate is mediated by the fraction of agents that build up precautionary savings, and their income share in the economy. Initially, structural change adds to the savings pressure by reallocating households into the urban sector where they try to build up an asset position. Overtime, as growth slows down for most households, and their human capital type is revealed, precautionary savings peter out. These compositional effects can aggregate up in way to yield hump-shaped saving dynamics where the aggregate saving rate picks up initially but reverts back in the long run.

Finally, I simulate a version of the model where I feed in a growth miracle that increases the per capita income of the miracle economy relative to the United States by a factor of around 6 (by a factor of 8 in absolute terms) to study saving rates and capital flows along the transition path. The literature on global imbalances and north-south capital flows mostly has abstracted away from transitional growth dynamics (Caballero, Farhi, and Gourinchas 2008; Mendoza, Quadrini, and Rios-Rull 2009), precisely because infinitely-lived forward-looking households would borrow against future income. Both the fact that growth is unevenly distributed, and reinterpreting ex-post inequality in urban production as ex-ante risk are key to resolve the tension between empirically observed outflows and predictions from the benchmark neoclassical model. When simulating the full dynamics of the model, the baseline parameterization delivers capital outflows along the transition path with a current-account-to-GDP ratio of around 5%, consistent with outflows observed during the Taiwanese or Chinese growth miracle. The simulation also delivers a realistic decline of the agricultural share, and a rise in inequality along the transition path.

The rest of the paper is structured as follows: Section 2.2 reviews relevant literature. Section 2.3 provides a set of stylized facts relating to miracle growth, structural change, and urban-rural differences. Section 2.4 develops a simple model that connects those facts and studies the tradeoff between catch-up growth and risk. Section 2.5 provides a simple quantification of the model. Section 2.8 concludes.

2.2 Literature Review

The model draws heavily on insights developed in the literature on precautionary savings (Deaton 1991; Carroll 1997; Gourinchas and Parker 2002; Carroll and Kimball 1996) and incomplete markets(Bewley 1977; Huggett 1993; Aiyagari 1994). In particular, the model builds on Huggett (1993) with a risk-free asset and human capital risk. My model shares the main predictions as the benchmark framework of Carroll (1997) but is a simplified version that adds convergence growth and structural change. In line with the precautionary savings literature, the model suggests that asset-to-income ratios are positively related to a household's risk exposure, providing theoretical context to the empirically different asset-to-income ratios across urban and rural households. Most incomplete market models are hard to handle and require heavy computational methods and approximations. In contrast, I derive the evolution of the income distribution along the transition path in closed form, and offer a particularly tractable precautionary savings framework where the tradeoff between consumption smoothing and savings boils down to a simple and intuitive sufficient statistic.[99] It is important to note, however, that the residual component of household income fluctuations a la Blundell, Pistaferri, and I. Preston (2008) is usually not sufficient to generate capital outflows during a growth miracle. Coeurdacier, Rey, and Winant (2019) provide a quantification of this claim by introducing a volatility of income that is twice as high in the emerging market without changing the prediction of the neoclassical model substantially. In contrast, if ex post inequality that is building up during a transition to a market-based economy is ex ante unknown, then a very powerful precautionary savings motive emerges.

Related to the focus on precautionary savings is the assumption that human capital risk is higher in urban relative to rural communities. The seminal paper by Townsend (1994) shows that rural village economies are close to a complete market benchmark in the sense that idiosyncratic income shocks are fully insured,[100] an implication of the potentially strong informal institutions in rural village communities as argued by Rosenzweig and Stark (1989). In contrast, idiosyncratic income risk is large in modern market economies, see for instance Heathcote, Storesletten, and Violante (2014a).

The focus on urban rural differences relates the paper at hand to a vast literature both in macroeconomics and development. The work by Harris and Todaro (1970) is the seminal paper that studies urban-rural wage gaps in a two sector economy. R. E. Lucas J. (2004) models the connection between development, urban-rural migration, and human capital accumulation. I add a central aspect to these "dual" economies by modeling modern human capital as fundamentally more risky relative to the "raw" labor input in traditional agriculture. There is also a recent literature on migration and risk (Bryan, Chowdhury, and

[99]The importance of precautionary savings for the Chinese growth miracle have been highlighted in several papers (M. Chamon, K. Liu, and Eswar Prasad 2013; Ding and He 2018; He et al. 2018) but the aforementioned papers abstract away from capital flows and urban-rural differences.

[100]Santaeulalia-Llopis and Zheng (2018) use the method of Blundell, Pistaferri, and I. Preston (2008) to show how the transmission of shocks to consumption has changed in China as growth took off. The results for their early period for urban households, that show a very low transmission of shocks to consumption, are the ones that I am basing this claim on.

Mobarak 2014; Morten 2019; Lagakos, Mobarak, and Waugh 2018). In contrast to this literature that tends to focus on temporary migration, this paper is concerned with long-run changes away from rural production in the broadest sense. I also build on the literature on the agricultural productivity gap (Restuccia, D. T. Yang, and X. Zhu 2008; Caselli 2005) and the related concept of the urban-rural wage gap. This urban-rural wage gap, which I take as given in the model, delivers an additional boost to catch-up growth as more and more households earn the higher urban wage.[101] Differences in total wealth accumulation between urban and rural households in Subsaharan Africa have also been documented by De Magalhães and Santaeulàlia-Llopis (2018a).

Financial frictions feature prominently in many theories of south-to-north capital flows. One strand of the literature argues that the flows occur due to developing economies' inability to produce safe assets (Caballero, Farhi, and Gourinchas 2008; Mendoza, Quadrini, and Rios-Rull 2009). My paper is consistent with and builds on these models as I assume that the risk-free asset is produced in the developed economy. Alfaro, Kalemli-Ozcan, and Volosovych (2007) and Alfaro, Kalemli-Ozcan, and Volosovych (2008) also provide empirical evidence that the capital flow puzzle originates from safe assets, while FDI for example tends to flow from rich to fast growing economies. Relative to this work, I incorporate urban-rural differences and consider transitional growth dynamics. I consider a growth miracle that pushes up GDP per capita by roughly a multiple of eight times, consistent with the Taiwanese experience, and orders of magnitude larger than what previous papers have considered.[102] Another central paper in this field is Z. Song, Storesletten, and Zilibotti (2011) which combines financing frictions and a heterogeneous firm model to study the Chinese growth miracle, and Buera and Shin (2017) which employ a similar model but focus more broadly on miracle economies. I view this paper as complementary to the literature centering around financial frictions. While financial frictions reflect an important aspect of emerging markets, they don't account for the high savings pressure of households that are

[101] There is a current debate to what extent the urban rural wage gap reflects selection (Lagakos and Waugh 2013; Young 2013; Gollin, Lagakos, and Waugh 2014; Hicks et al. 2017; Lagakos, Marshall, et al. 2020), casting doubt on the idea that urban-rural structural change can boost growth. Note, however, that much of the work focuses on stagnant economies. Urban-rural migration seems much more important during a growth miracle, and it is hard to imagine China would have been able to grow at 10% if 80% of its population had stayed in agricultural production.

[102] Caballero, Farhi, and Gourinchas (2008) and Mendoza, Quadrini, and Rios-Rull (2009) mostly abstract away from growth dynamics, which are a first order feature of the economies that display large capital outflows. Buera and Shin (2017) and Sandri (2014) consider transitional growth dynamics that are an order of a magnitude smaller than the ones considered here.

not involved in entrepreneurial activity.[103] A model without frictions on the household side delivers borrowing due to consumption smoothing of workers. Micro level household data is inconsistent with this strong consumption smoothing motive for the emerging middle class in emerging markets.

Several authors focus on demographic factors to explain household saving rates. İmrohoroğlu and Zhao (2017) and İmrohoroğlu and Zhao (2018) argue how the one-child policy in China can lead to savings pressure and capital outflows. Wei and X. Zhang (2011) posit that the high Chinese saving rates are driven by the gap in the sex ratio, and Curtis, Lugauer, and Mark (2015) highlight the relationship between demographics, age, and saving rates in China. Importantly, when documenting empirical differences in savings behavior across urban and rural households I show that the differences are robust to demographic controls. The main argument against demography-based explanations is, however, that other miracle economies have displayed similar dynamics with very different demographic fundamentals. For instance, Taiwan did not impose any restrictions on the number of children per household, and the marriage market in post-war Germany very much favored men due to the death of a disproportionate amount of male soldiers. One thing that all these miracle economies had in common, however, is fast-paced structural change out of agriculture as I will show in the next section. This structural change in combination with urban-rural differences is able to reconcile the puzzling relationship between catch-up growth and capital outflows.

2.3 Empirics of Miracle Growth and Structural Change

In this section I provide a set of stylized facts relating to the macro as well as the micro dynamics of miracle growth. While the macro facts of growth, savings, and capital flows are well known, I relate them to the fast-paced structural change in the miracle economy. I offer novel facts from Chinese and Thai household data that highlight urban-rural differences in saving behavior and asset accumulation, and how they might relate to uneven and uncertain labor market outcomes in the emerging urban economy. In what follows I use the terms city versus countryside, urban versus rural, and agricultural versus non-agricultural interchangeably.[104]

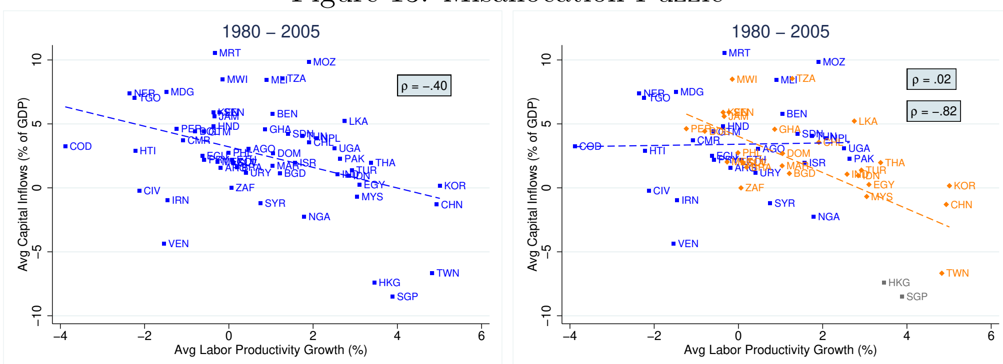
[103] Fan and Kalemli-Özcan (2016) cast doubt on the positive relationship between financial frictions and corporate savings in Asia.

[104] While this is not ideal, the categories are strongly correlated. It would be challenging to study change in the one, without change in the other. The actual measure used depends mostly on the available data, see for instance the work by Young (2013), Gollin, Lagakos, and Waugh (2013), Hicks et al. (2017), or Hnatkovska and Lahiri (2018). Hence, I lump them together, as is often done in the literature. I will make sure to point

2.3.1 Macro Facts

Figure 13 is a version of the main figure in the influential paper of Gourinchas and Jeanne (2013). While the left panel plots the familiar puzzling negative relationship between productivity growth and capital inflows, the right panel separates countries into economies that exhibit relatively fast or relatively slow structural change. In particular, the orange diamonds represent economies that display above average declines of the agricultural employment share.

Figure 13: Misallocation Puzzle



Relationship between average capital inflows and average labor productivity growth for a cross section of emerging markets following Gourinchas and Jeanne (2013). Orange diamonds represent emerging markets with above average decline of the agricultural employment share. Current-account-to-GDP is taken from WDI. Agricultural employment shares are taken from the GGDC ten sector database, and WDI. Labor productivity growth rates are computed from real national series from the PWT 9.1.

The countries that drive this negative correlation are also known as "miracle economies", a term coined by R. E. J. Lucas (1993), and usually referring to the East Asian tiger economies who have experienced unprecedented per capita growth. Unquestionably, this fast reallocation of labor out of agricultural production is itself a by product of massive increases in labor productivity in the manufacturing sector.[105] As mentioned in the introduction, fast catch-up growth in the benchmark neoclassical model implies consumption smoothing and current account deficits. In this paper, however, we have an additional lever to approach the puzzle, because fast productivity growth leads to fast structural change

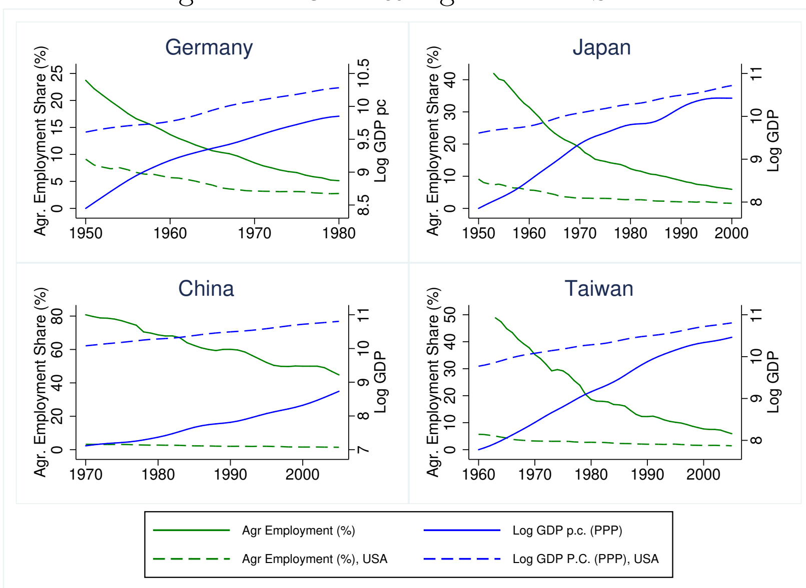
It is useful to go through the aggregate dynamics of the growth miracle, which are well known, and juxtapose them with structural change out of agriculture. Figure 14 highlights

out what concepts are used where in the empirical work.

[105] There is a debate about the importance of factor accumulation (Young 1995) relative to TFP growth (Hsieh 1999).

the relationship between catch-up growth and structural change in the form of a declining agricultural employment share for four miracle economies (Japan, Germany, Taiwan, China), loosely following Buera and Shin (2013).[106]

Figure 14: GDP & Agricultural Share



Relationship between agricultural employment share and convergence in GDP for Germany, Japan, China, and Taiwan. GDP series in purchasing power parity taken from the Penn World Tables 9.1. GDP is smoothed using an hp-filter with smoothing parameter of 8.5.

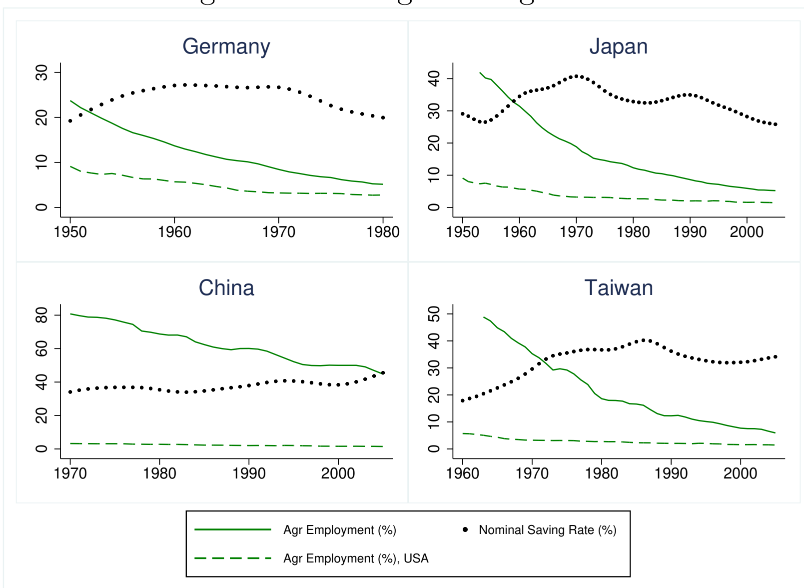
Figure 15 displays national saving rates over time, and shows a hump shaped pattern of the saving rate over the convergence process, except for China which is still in the catch-up phase. The saving rate picks up, with a lag, as the agricultural share declines relatively fast (compared to the US agricultural share) and growth takes off. The growth in the saving rate peters out as the country's convergence process comes to an end, and so does the spectacular decline in the agricultural share.

The rising aggregate saving rate becomes even more problematic in the open economy when the identity of savings and investment breaks down. Figure B2, reported in the appendix to save space, depicts the positive current account balance that has been identified as a robust feature of growth accelerations (Hausmann, Pritchett, and Rodrik 2005).[107] Based

[106] The recent handbook chapter by Herrendorf, Rogerson, and Valentinyi (2014) discusses this shift from the agricultural to the manufacturing and service sector as a general pattern in the process of economic development and industrialization. While this pattern holds across virtually any country, the speed of structural change in miracle economies is exceptional.

[107] Consistent with the findings of Buera and Shin (2017) the positive current account dynamics are more pronounced in the 1980s when most countries liberalized their capital accounts.

Figure 15: Savings and Agr. Share



Relationship between agricultural employment share and savings, data from the WDI. Same smoothing procedure applies.

on an accounting identity in the national accounts, the positive current account balance implies that aggregate national savings must exceed domestic investment leading to capital outflows.[108] Household saving rates were increasing during the growth acceleration in all economies, and more so for high-income households (Attanasio and Székely 2000), ultimately driving outflows.

To summarize, the main macro facts are i) fast-paced structural out of rural or agricultural production, ii) a hump-shaped saving rate that seems to be inversely related to the speed of structural change, and iii) capital outflows. The theoretical model will be able to replicate all three macro facts.[109]

2.3.2 Evidence from Household-Level Data

In this section I complement the macro facts by studying urban-rural differences in savings behavior and asset accumulation on the household level employing a median-regression approach. The main analysis is centered around the Chinese data which is uniquely suitable to measure urban-rural differences. Aggregate income in China has been growing at a rate

[108] The measurement of global capital flows is challenging (Coppola et al. 2020), but the qualitative finding that growth miracles are associated with capital outflows is robust (Gourinchas and Jeanne 2013).

[109] To corroborate the relationship between structural change and savings pressure, I offer additional evidence from cross country regressions in the appendix **??**.

of around 10 % for more than two decades, accompanied by fast structural change and urbanization. On the other hand, the Chinese economy is characterized by large differences in the level of development across its provinces, which allows me to compare urban households to rural ones. The main dataset is the Chinese Family Panel Study (CFPS). The CFPS is a household panel dataset that comprises detailed information on family structure, income, expenditure, assets, and other demographics. The survey was launched in 2010 by the Institute of Social Science Survey (ISSS) of Peking University, China.[110] The dataset is similar to the PSID, but many survey questions are designed to capture relevant variables for Chinese families. The CFPS data for 2010 contains roughly 15,000 households, in 25 provinces excluding Hong Kong, Macao, Tibet, Qinghai, Inner Mongolia, Nigxia, Hainan. An eligible household refers to an independent economic unit with at least one Chinese national. I use the CFPS to study differences in household savings and asset-accumulation behavior across urban and rural areas.

2.3.3 Urban-Rural Differences in Asset-to-Income Ratios

In order to learn about urban-rural differences in savings behavior I focus on differences in asset-to-income ratios. While it may seem more straightforward to measure saving rates directly, it turns out that households saving rates are often poorly measured. In fact, in the CFPS, which is of high quality and employs similar techniques as its US equivalent, the PSID, it is not uncommon to find households saving rates of minus 400 %. Of course, this very negative saving rate might reflect measurement error, but perhaps equally likely an inability to account for shifting positions of asset classes. Imagine a household that took out a mortgage and bought a house with a 30 % down-payment. This may look like large negative savings, while the household may actually be building up savings but turned liquid assets into a fixed asset. An alternate strategy is to focus instead on asset-to-income ratios, which has a long tradition in the precautionary savings literature (Carroll and Samwick 1997). Asset-to-income ratio use a stock-concept that reflects past saving and consumption choices, but are inherently more stable than saving rates.[111] Consider a standard budget constraint with a risk-free asset a, labor income y, and consumption c of the form $\dot{a}_t = ra_t + y_t - c_t$. Suppose that the household consumes a fraction $(1 - s)$ of labor income and starts out with

[110] After applying for access online, the data is in principal accessible to any researcher. For more information, see `https://opendata.pku.edu.cn/dataverse/CFPS?language=en`.

[111] As Carroll and Samwick (1997) point out, in a buffer-stock savings model saving rates are only higher for households that are below their optimal buffer-stock asset level. Once the household has accumulated a sufficient amount of wealth, income and consumption grow at the same rate.

zero assets, then it immediately follows that the asset-to-income ratio reads

$$\frac{a_T}{y_T} = s \int_0^T \exp(-rt) \frac{y_t}{y_T} dt$$

If T gets large and household income grows at a constant rate g_h the expression simplifies to $\frac{a}{y} = \frac{s}{r-g}$ and is directly proportional to the saving rate. Moreover, in canonical precautionary savings models (Carroll 1997), asset-to-income ratios are sufficient statistics for the precautionary motive since greater risk induces households to accumulate larger buffer-stock savings relative to their income. The simple model I sketch out in the next section features the same positive relationship between human capital risk and asset-to-income ratio. Through the lens of these models, significant differences in asset-to-income ratios, after controlling for a number of other factors, is suggestive of greater savings in the urban sector due to a more risky environment. Importantly, these ratios are naturally normalized by income, which is in the denominator, and therefore are not simply a by product of higher income in one area relative to another. Of course, other factors such as household age should affect this ratio as well. The benefit of the median regression approach is that I am able to control for demographics and other confounders.[112]

Since this ratio-based measure is inherently unstable and explodes for large levels of income, it is common to employ a median (quantile) regression as in Fagereng et al. (2019).

I estimate the following linear specification for the 2012 cross section of the CFPS

$$\frac{a^i}{y^i} = \alpha + \beta D_i + \Gamma' X_i + \epsilon_i \tag{2.1}$$

where $\frac{a^i}{y^i}$ is the asset-to-income ratio and D_i is a dummy variable that takes on a value of one if the household is non-agr. (urban), and zero otherwise. X is a vector of controls that contains income, education, demographics, and other covariates. ϵ is assumed to be a random error term.

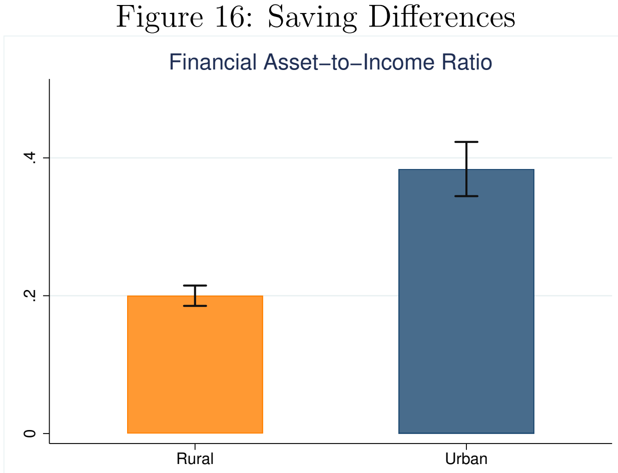
To run this regression, I restrict the sample to employed household heads that are be-

[112] There is an important subtlety here: the theoretically consistent measure in the literature on precautionary savings would be the asset-to-permanent-income ratio. I do not attempt to estimate permanent income which seems particularly challenging in the fast-changing environment of the Chinese Growth miracle. For example, the rising returns to education might have been hard to foresee in 1995 for individual households where the internationalization of the Chinese economy had yet to happen. Instead, I offer robustness checks based on asset-to-consumption ratios, which is a good proxy of permanent income for forward-looking households that are not borrowing constrained. Papers that aim to estimate the permanent component of income are Carroll, Dynan, and Krane (2003), Carroll and Samwick (1997), Fuchs-Schündeln and Schündeln (2005) or He et al. (2018).

tween 23 and 60 years old, in line with previous work (He et al. 2018; Storesletten, Telmer, and Yaron 2004). Additional details on sample selection is provided in the appendix. I focus on urban-rural differences, which I think best captures the distinction between a stable agriculture-based society relative to a fast-paced and uneven growth experience in the urban-based Chinese economy. Additional results for different years and for agr vs. non-agr households are offered in the appendix. My preferred variable to understand urban rural differences is $urban\_cfps$ which is a community based measure that groups villages into urban and rural areas provided by the CFPS. This measure is different from the Census Bureau's definition. In the appendix in subsection I discuss the official definition, and highlight some problems with it. Here, I also focus on financial "safe" assets which connects more closely with the previous literature (Caballero, Farhi, and Gourinchas 2008; Mendoza, Quadrini, and Rios-Rull 2009) and the theoretical model in the next section. Tables in the appendix comprise summary statistics for the raw sample of the CFPS. Urban (rural) household in 2012 have a mean income per capita of 20,434 (9,976) Yuan, and the household has, on average, close to 10 (6.5) years of schooling. Rural household heads are younger (42.5 vs 47 years) and rural families are larger (4 vs 3.3 people).

Figure 2.3.3 plots the regression coefficient for rural ($\hat{\alpha}$) and urban ($\hat{\alpha} + \hat{\beta}$) households, based on equation 2.1, without any controls. In this case, β reflects the difference between the median financial asset-to-income ratio of urban and rural households. Urban households hold substantially larger financial asset-to-income ratios with a median value of .4, relative to rural households with a median value of .2.

Figure 16: Saving Differences



Cross sectional Urban-rural median differences in financial asset-to-income ratios for CFPS 2012 with 95% confidence intervals. See table B16 for additional information.

Table 5: Median regression with urban-rural dummy for CFPS 2012

	(1) g_fin_asset	(2) g_fin_asset	(3) g_fin_asset	(4) g_fin_asset	(5) g_fin_asset
hukou_switcher	0.0504*	0.0449	0.0465	0.0478	0.0393
	(0.0296)	(0.0297)	(0.0307)	(0.0307)	(0.0324)
_cons	0.271***	0.242***	0.565**	0.544**	0.515**
	(0.0189)	(0.0212)	(0.235)	(0.237)	(0.250)
income growth	No	Yes	Yes	Yes	Yes
age & demographics	No	No	Yes	Yes	Yes
education	No	No	No	Yes	Yes
province fe	No	No	No	No	Yes
N	1115	1110	1110	1110	1110

Note: The dependent variable is the household financial asset-to-income-ratio. This contains bank deposits, stocks, derivatives, bonds, cash, and other financial assets. Robust Standard errors in parentheses. *, **,*** denote statistical significance at 1, 5, and 10 percent level.

Of course, one major concern is selection and omitted variable bias which I address next. It is well known that there are many other reasons that drive household saving behavior, for instance life cycle motives (Modigliani 1986), or, in the Chinese context, a competitive sex motive (Wei and X. Zhang 2011). In table B16 I report the results for the median regression where I control for a second order polynomial in income, a second order polynomial in age as well as additional demographics,[113] The differences found are robust and remain highly significant well below the 1 % level. From a theoretical point of view, however, it is unclear that controlling for education or income is appropriate. In the model in the next paragraph, human capital and income increase in urban economic activity and are tightly connected to savings. Through the lens of the model, controlling for education or income in a world with risky human capital takes out the essence of "urban" production.

The reader should consider the higher asset-to-income ratio together with the fact that median income is more than twice as high in urban ares compared to the country side. That is, not only do urban households accumulate a larger asset position relative to their income, their income is also a multiple of rural income. This highlights how urbanization is a driver of the demand for safe assets. I also report results for the total-asset-to-income ratio in the appendix[114]

Through the lens of an incomplete market model, it looks like urban households have a

[113] This includes the sex of the household head, the share of household members over 60, as well as whether there is a male heir in the household. In Chinese culture it is common for the male heir to look after the parents in old age, which might interact with life cycle saving motives.

[114] The estimated magnitudes are much larger in absolute terms which is intuitive since total assets are much larger than financial assets. On the other hand, the relative differences are somewhat comparable, i.e. the ratio is 20% higher for the median urban (non-agr.) household. The results for consumption are less clear-cut, as well as the results for agricultural occupations which lines up with the importance of productive assets in agricultural activity.

stronger precautionary motive, and the main point in this paper is to highlight the massive human capital risk in modern productive activity that is mostly absent in traditional agricultural production. The simple model in the next section makes the link between precautionary motive and a relatively high demand for safe assets precise. The fact that the demand for safe assets seems to be so much higher for urban households opens up the possibility that urban-rural differences combined with fast-paced structural change is quantitatively important to account for the capital flow puzzle.

2.3.4 Human Capital Risk and Uneven Growth

The reduced-form results suggest that urban-rural differences in savings behavior are large. Together with the rise of urban economic activity and fast structural change, this is suggestive that periods of fast structural change are followed by strong savings pressure.[115] The harder problem, however, is to write down a model with forward-looking agents that allow for the coexistence of catch-up growth and savings pressure. The Lucas puzzle really is a puzzle for the theorist who insists on models with forward-looking expectations that respect a version of the permanent income hypothesis.

Foreshadowing the theoretical framework in section 2.4, there are two ingredients that are necessary for capital outflows during a growth miracle to occur. First, there needs to be a source of risk that can leave households worse off for some time. I identify this as human capital risk, and I am arguing that this risk is acute in modern production and mostly absent in traditional farming. Second, I need convergence growth itself to be unevenly distributed, with households not knowing ex-ante how much they will participate in the aggregate growth miracle. This is a key departure from the previous literature and quantitatively important in accounting for the capital flow puzzle. The theory in section 2.4 will lay out why these two features are so central. Intuitively, you need the possibility to be worse off to have an incentive to save. If the future is always brighter than the past, agents want to borrow. Uneven growth, on the other hand, is equally important from a quantitative point of view but by itself will not generate capital outflows.

As mentioned already, while income processes are volatility in rural areas, due to the importance of weather shocks to production, consumption profiles are surprisingly smooth as

[115] The reader might wonder whether the argument implies a similar savings pressure during the process of industrialization in the US or UK. The answer is no. The key difference here is the speed at which people move out of agriculture to generate aggregate savings pressure. If this process happens slowly then the share of households that accumulate is relatively small given the size of the economy.

has been documented by a number of papers in the literature (Townsend 1994; Santaeulalia-Llopis and Zheng 2018; De Magalhães and Santaeulàlia-Llopis 2018a).[116] This suggests that own savings are potentially more important as an insurance tool in modern productive activity in urban areas. To tackle the Lucas puzzle, however, a large source of risk is needed in the urban areas since income growth is phenomenal. It is clear that simple measured volatility of the income process in the spirit of Blundell, Pistaferri, and I. Preston (2008) is not going to be powerful enough.[117] The argument proposed in this paper is that the inequality that emerges in modern market economies provides such a source of risk, if ex-ante households do not know where they will land on the income distribution. It is hard to discipline the question of what households know ex-ante, but I will offer a set of facts that are consistent with the interpretation that ex-post inequality represents ex-ante risk. Ultimately, if one dismisses this approach it is hard to see how high saving rates of ordinary households can be reconciled with miracle growth.[118]

Figure 17 plots the rise in wage-inequality for the case of urban China.[119] The key question is whether rising inequality represents ex-ante risk during a growth miracle, which leads to powerful precautionary savings pressure. In contrast, if households knew where they will end up on the distribution we would expect a lot of borrowing and little saving since, on average, households clearly are much richer in the future. While I am not able to measure the information set of households directly, I can assess the extent to which observables explain variation in log wages, especially human capital and experience. The answer is extremely little. This is a point worked out more carefully by Ding and He (2018) who show that rising inequality in China is mostly driven by residual income inequality. This supports the possibility that income growth might be hard to forecast.

Another important takeaway from figure 17 is that despite the right shift of every wave, there is an overlap of the distributions – that is to say, at least in the cross section, there

[116] Santaeulalia-Llopis and Zheng (2018) build on Blundell, Pistaferri, and I. Preston (2008) and provide evidence from China that (informal) insurance of rural households before the reform period was very high.

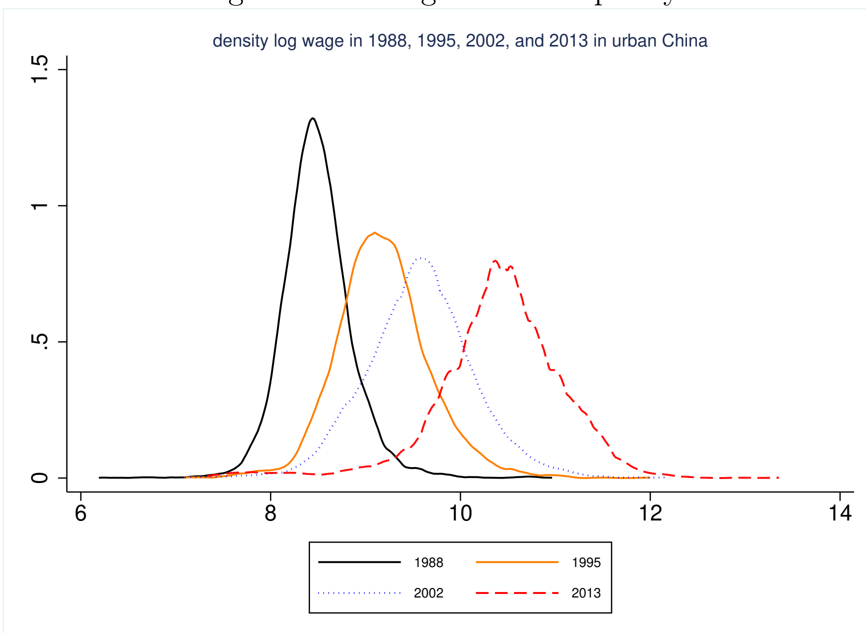
[117] I provide a representative agent model that makes this point in a stylized way in the appendix. Quantitative work supporting this claim can be found in Coeurdacier, Rey, and Winant (2019).

[118] The work by Buera and Shin (2017) or Z. Song, Storesletten, and Zilibotti (2011) are helpful for understanding high saving rates for entrepreneurs but the evidence suggests that pretty much everyone is saving at a relatively high rate in urban China, compared to the United States for instance.

[119] Wage inequality is more easily measured than total household income, especially because property reforms and privatization changed the kind of benefits workers used to obtain from their employers (meals, in kind transfers, housing), and the money-equivalent of these benefits is prone to measurement error. If one were to use household income instead, the fanning out of the distribution would be even more extreme and a fatter right tail of the distribution would emerge in 2013. Results available upon request.

are income realizations that are below the mean or the median of the wage distribution in the previous waves. A central insight of the model in the next section is that a necessary condition for bufferstock savings during the growth miracle is that households need to face risk that could leave them worse off in terms of household income, at least for some time.

Figure 17: Rising Urban Inequality



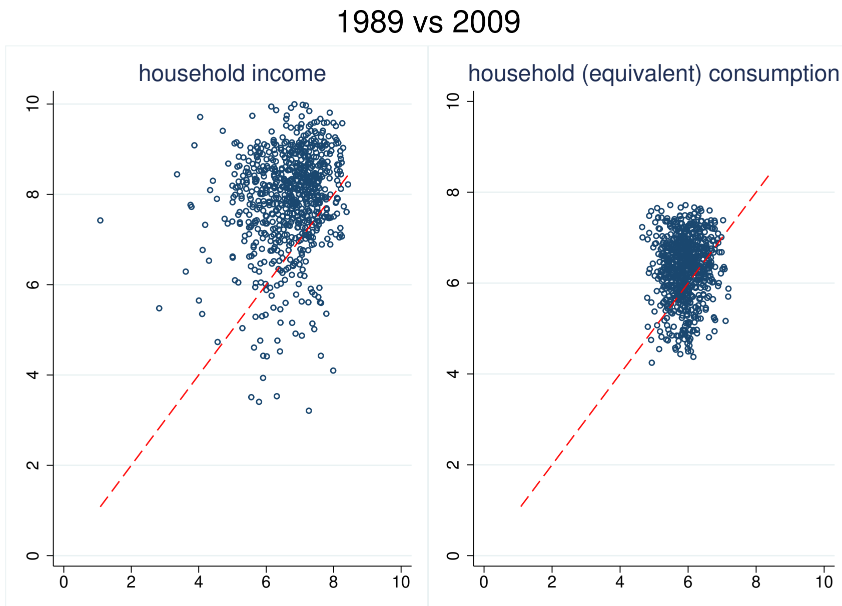
Data based on Chinese Household Income Project (CHIP), see Shi (2009). Density plot of log of real wage income for fulltime male household heads between the age of 23 and 60 in urban China in 1988, 1995, 2002, and 2013.

While the cross-sectional plot isn't really informative, I use the panel dataset constructed by Santaeulalia-Llopis and Zheng (2018) in figure 2.3.4 to show that even when focusing on the same household, a substantial share of the population is in fact experiencing income losses over time. The dashed red line is the 45 degree line, indicating that all households below the 45 degree have experienced an income loss in 2009 relative to 1989. When looking at household equivalent consumption a similar picture emerges. Importantly, the fact that a number of households land below the 45 degree line does not seem to be driven by household compositional effects.[120] The same holds true for household consumption including expenditures on food, utilities, health, and semidurable supplies.

It seems indeed possible that some households are worse off than before the growth miracle, despite rates of aggregate growth close to ten percent for more than two decades.

[120]You can drop household heads older than 55 years, or you can check income per capita or income per adult, which gives similar results.

Figure 18: Long-Run Differences in Income and Consumption



Log of household income and household food consumption in equivalent units based on the China Health and Nutrition Survey (CHNS). Data directly taken from Santaeulalia-Llopis and Zheng (2018), see their paper for details.

The second key ingredient in the model is that growth itself is unevenly distributed. In subsection 2.4 I am more precise about this, but loosely speaking I introduce heterogeneity in the growth rate itself. The fast rise of top income inequality in China suggest that heterogeneous growth rates are important.[121] While the fast rise in inequality can in principal be modeled by a rising variance of shocks to permanent income as in Santaeulalia-Llopis and Zheng (2018), Gabaix et al. (2016) show that such a setting gives rise to very slow transitional dynamics.[122] A growth process that is inherently uneven, as proposed by Gabaix et al. (2016) and used for instance in Jones and Kim (2018), is able to match such fast changes in income inequality.

[121] Piketty, L. Yang, and Zucman (2019) document how top income inequality has shot up during the Chinese growth miracle, at a rate that is unprecedented in modern history.

[122] It seems that the estimate of the variance of the random shock to the permanent component of household income of Santaeulalia-Llopis and Zheng (2018) is on the higher end of available estimates, compared for instance to M. Chamon, K. Liu, and Eswar Prasad (2013) that are much closer to estimates in the US.

2.4 Theoretical Framework

Here I show how a simple model with a stochastic process that combines random convergence growth with "type" draws from a distribution can account for the patterns in the data displayed in section 2.3. In particular, the model generates strong precautionary savings which can give rise to a hump shaped saving rate, and capital outflows, together with a realistic income distribution.

The trajectory of households in this model economy is characterized by three stages: First, households optimally decide whether to stay in the agriculture sector or move on to the non-agricultural sector. A diminishing returns-to-scale technology on the countryside combined with productivity growth of the constant-returns urban technology gives rise to structural change out of agricultural activity. Second, after entering the non-agricultural sector, the household's income starts growing at a higher growth rate compared to the industrialized world (RWO). This feature delivers catch-up growth relative to the ROW. The time agents spend in the high-growth regime is random. Agents leave the high-growth regime according to a Poisson arrival process, after which their income grows at a lower "normal" rate that is the same as the growth rate in the rest of the world. Note that this formulation makes growth itself risky and uneven on the household level. Third, once the agents' income growth slows down, they have to draw their "type" from a distribution with positive support centered around one., i.e. an inequality shock. This inequality shock is important as it creates the possibility of households being worse off, at least for some time. The theoretical analysis will show that without this additional shock it would be impossible to generate precautionary savings. From then on, all uncertainty is resolved and the household grows along a balanced growth path with constant consumption, income, and asset growth.

Time & Sectors:

Time is continuous and indexed by $t \in R_+$. There are two sectors in the economy, a rural and an urban sector. These sectors are endowed with different technologies, and households are allowed to switch from the rural area to the urban area, but not the other way around.

Production Technology, and Market structure:

Urban and rural firms produce a single final good with prices normalized to one, $Y_t = Y_t^u + Y_t^r$, where the superscript u and r stand for urban and rural, respectively. The urban technology is constant-returns-to-scale with free entry and labor as only factor of production, ensuring that there are zero profits in equilibrium. There is a Solow-neutral productivity shifter A_t that will grow over time. The firm problem, after substituting in the technological constraints, reads

$$\max_{[H_t^u]} A_t H_t^u - w_t H_t^u, \tag{2.2}$$

where H_t^u is the effective labor supply of households in the city. In equilibrium under perfect competition the wage rate in the city equals $w_t = A_t$.

The technology on the country side displays diminishing-returns due to a fixed factor land which is normalized to one. The parameter $\alpha \in (0,1)$ governs the curvature of this production function,

$$Y_t^r = A_0^r \left(L_t^r\right)^{\alpha}. \tag{2.3}$$

I assume that all workers on the country side collectively own the land, i.e. rural household income is total rural output divided by the number of rural households.[123] There is no depreciation. Hence, the compensation for the worker w_t^r in the rural sector is output divided by the rural labor force (so that it includes the return to land) and reads

$$w_t^r = A_0^r \left(L_t^r\right)^{-(1-\alpha)}.$$

Note that the diminishing-returns-technology on the country side implies that the rural wage increases as workers leave the rural sector. Instead, the urban sector can accommodate an unlimited amount of workers while maintaining a constant marginal product of labor. In combination with productivity growth in the urban-technology, this setting will give rise to structural change where productivity growth in manufacturing pulls out workers from agricultural production.

Storage Technology

In order to simplify, I assume that only households in the city have access to an internationally traded risk-free bond that pays a constant interest rate r^* determined by the balanced-growth equilibrium of the industrialized world. In contrast, rural households live hand-to-mouth similar to the setup by Moll (2014). This assumption is qualitatively consistent with the low built up of assets documented in section 3 as well as work by De Magalhães and Santaeulàlia-Llopis (2018b).

Convergence, Type Space, and Stochastic Processes:

In the "city" workers grow relatively fast for some time, but are also exposed to human capital risk in the form of a bad type draw, leading to additional inequality. I introduce

[123] This assumption fits the Chinese context well. See Tombe and X. Zhu (2019) for a model of internal migration and trade where collective ownership of rural land in China induces an additional frictions.

convergence in a very tractable way to solve the model in closed form. First, I assume that the technology A_t^u grows exponentially at the industrialized world growth rate g^*.

$$A_t^u = A_0^u \exp(g^* t) \tag{2.4}$$

Since I want the model to relate to the growth experience of miracle economies, the best way to think about A_0 is as a state that prevails for some time before the country introduces policy reforms and begins its catch-up process. In that sense, one can think of the economy before t_0 as a stagnant one, where productivity in the city is constant, i.e. $A_s^u = A_0^u, \forall s < 0$. Time zero is in that sense a normalization and really marks the time that reforms begin. I do not need to keep track of what happens before time zero since the equilibrium is stationary and summarized in the (old) steady state at time zero. The process of reforms, then, causes continued per capita growth, unique to the capitalist system (R. E. J. Lucas 2018).[124] I model this catch-up process as a Poisson process where individual households get to catch-up at a very high growth rate with the rest of the world for some random time. When a household enters the city, it also gets to enter a Luttmer (2011)-"growth rocket", grow at a very high growth rate g_h until they are randomly pulled out at time T^i, based on a Poisson process with arrival rate λ. This income growth reflects a rising effective labor endowment which could be micro-founded by models of learning by doing or human capital accumulation. When this growth spurt is over, households are hit by a type shock φ that parameterizes inequality in the market-economy. The type draw itself is centered around one and does not generate aggregate catch-up. Note that t_m^i stands for the time of migration of household i and a household's effective labor supply before entering the urban economy is normalized to unity. Adding up the growth rate of technology together with growth of the effective labor supply yields the following expression for income of household i for $t < T_i$

$$\begin{aligned} y_t^i &= w_t h_t^i \\ y_t^i &= w_t \exp((g - g^*)[t - t_m^i]). \end{aligned} \tag{2.5}$$

[124] Because of the convergence growth, there is a discontinuity in the agricultural employment share in the model at time zero.

Then, using (2.2), the log of income for household i equals

$$\log(y_t^i) = \begin{cases} \log(A_t) + [t - t_m^i](g - g^*) & \text{if } t \leq T^i \\ \log(A_t) + [T^i - t_m^i](g - g^*) + \log(\varphi^i) & \text{if } t > T^i \end{cases} \tag{2.6}$$

which, in terms of growth rates reads

$$\frac{d \log(y_t^i)}{dt} = \begin{cases} g_h & \text{if } t < T^i \\ g^* & \text{if } t > T^i \end{cases} \tag{2.7}$$

for agents in the city with

$$F(T^i - t_m^i) \sim exponential(\lambda). \tag{2.8}$$

At time T the derivative is not well defined as income jumps up or down, depending on the type draw. Recall that there is no technological change on the country side. Note that the type draw is entering multiplicatively so as to augment effective human capital.

Now I can characterize the budget constraint for households in this economy. Let the sets S_r and S_u form a partition of the unit interval of agents into rural and urban activity. Agents on the country side have to consume all their income

$$w_t^r = c_t^r \quad \forall t, \forall i \in S_r. \tag{2.9}$$

In the city, I allow for a meaningful intertemporal consumption-saving choice with a standard budget constraint

$$\dot{a}_t = \begin{cases} [r^* a_t + y_t^0(t_m) - c_t]\, dt & \text{if } t < T \\ [r^* a_t + \varphi^i y_t^1(t_m, T) - c_t]\, dt & \text{if } t \geq T. \end{cases} \tag{2.10}$$

where I dropped the i superscript. Agents with superscript 0 have not drawn their type yet, and grow at the faster rate g_h. Agents with superscript 1 did draw their type, and grow at the world growth rate g^*.

This paragraph contains the key assumptions of the model that end up delivering an income process similar to the one displayed in figure 19, where W_{gap} denotes a potential urban-rural wage gap. Let's discuss these assumptions in turn. I employ the most simple method to induce a process of urban-rural structural change that is consistent with the importance of pull-factors during early stages of the development process (Alvarez-Cuadrado and Poschke 2011; Hnatkovska and Lahiri 2018).[125] In the same vein, the high income growth rate in the city is qualitatively consistent with faster income growth in urban areas in China (Santaeulalia-Llopis and Zheng 2018). The Poisson process that governs the average time spent in the high-growth regime is extremely useful here as it leads to exponential and hence memoryless waiting time, allowing me to solve key aspects of the dynamic model as well as the evolving income distribution in closed form. The income process in the city combines risky growth with an additional inequality shock. We will see shortly that only the latter piece can potentially generate precautionary savings. The main idea captured by this stochastic income process is that human capital differences are mostly absent on the country side, while they are of first-order importance in urban production. That is, seemingly similar workers can earn massively different salaries in human capital intensive industries while they would earn the same income if they had to toil on the field.

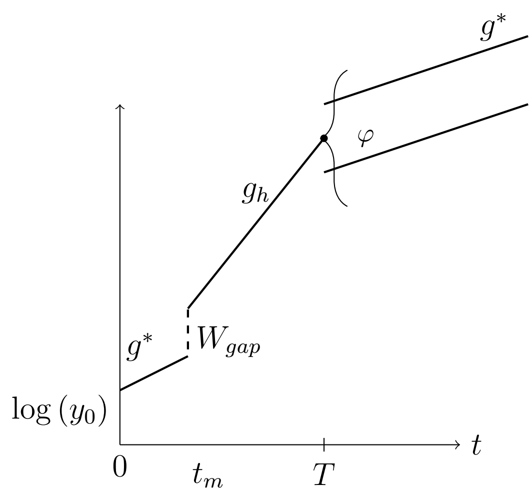


Figure 19: Income process

Preferences:

I assume a flow utility function of the CRRA form with coefficient of relative risk aversion η for a unit measure of infinitely lived dynastic household, indexed by $i \in [0, 1]$. Households discount utility exponentially at rate ρ. The labor supply of each household is inelastic and normalized to unity. There is no population growth. In what follows, I omit the i subscript

[125] See Herrendorf, Rogerson, and Valentinyi (2014) for a discussion of a variety of models that give rise to structural change, which depends on both the preferences structure of the agents (homothetic vs non-homothetic and complements vs substitutes) as well as the direction of technological change (productivity growth in the city vs productivity growth in the rural sector).

but in principal all household variables should be indexed by i

$$\max_{[c_t, a_t, t_m]} \mathbb{E}_{\varphi,T} \left[\int_0^\infty \exp(-\rho t) \frac{(c_t)^{1-\eta}}{1-\eta} dt \right]. \tag{2.11}$$

The agent maximizes expected utility, where the expectations are taken against the random arrival time $T \in R_+$, and the type of agent $\varphi \in R_+$, both of which represent a source of risk. Note that I keep the preference structure as simple as possible. Needless to say, solving the capital flow puzzle (which is ultimately a consumption-smoothing puzzle) becomes easier when introducing relative consumption preferences as in Kogan, Papanikolaou, and Stoffman (2020) or Epstein-Zin preferences to separate risk aversion from intertemporal elasticity (Epstein and Zin 1991).

Migration decision:
Migration in this framework is only allowed from the rural region to the city, and not the other way around. This may seem to be a strong assumption, given recent work on this topic. Young (2013) finds that migration in developing economies is "two-way".[126] Many of the countries considered in the analysis, however, are stagnant economies. They do not resemble countries like China or Japan, where the agricultural share as displayed in figure 14 declines at a stunning pace. In fact, Chinese data from the Chinese Family Panel Study (CFPS),, show that the fraction of households who change their hukou from urban to rural is virtually zero, while a change in the other direction is common. The hukou system is regulating migration flows within China, and essentially prevents most rural households from moving to urban areas. A household with a rural hukou in an urban area has a similar status as an illegal immigrant in the United States (Piketty, L. Yang, and Zucman 2019), although this depends partially on the federal province in question. See Chan and Buckingham (2008) for further information, and a discussion of reforms in the hukou system in the 2000s.[127]

To understand the migration decision, I write down the discrete-time equivalent and take the limit as the time interval Δ goes to zero. Since agents are allowed to leave the country side whenever they want to, this problem ends up being an arbitrage condition that keeps households indifferent between staying or leaving. In equilibrium, a sufficient amount of agents will leave the country side so as to preserve this indifference conditions for the stayers

[126] Lagakos and Waugh (2013) explain this finding in a Roy model of labor market sorting. See also Hicks et al. (2017) for evidence from long run panel data on this topic.

[127] The hukou system is complex and has seen multiple reforms since 1980. In the appendix in section B.9 I discuss the hukou system in a little more depth.

at every point in time. To avoid counterfactual implications for the urban-rural wage gap W_{gap}, I introduce a migration cost.[128] The cost is paid in utility terms and proportional to the utility associated with moving to the city.[129] I introduce this wedge in the form of the parameter $\tau^{\eta-1}$ where $\tau > 1$. Formally, let V_t denote the value function of moving to the city at time t. There is no other state variable than t since agents on the country side do not have access to a storage technology. The arbitrage condition that has to hold in equilibrium in discrete time reads

$$\Delta \frac{(w_t^r)^{1-\eta}}{1-\eta} + \tau^{\eta-1}(1-\Delta\rho)V_{t+\Delta} = \tau^{\eta-1}V_t. \tag{2.12}$$

Taking the limit as $\Delta \to 0$ yields the continuous-time equivalent

$$\frac{(w_t^r)^{1-\eta}}{1-\eta} = \tau^{\eta-1}\left(\rho V_t - \dot{V}_t\right). \tag{2.13}$$

The intuition is that a household on the country side is always indifferent between moving today, or waiting another period. Since this must hold every period in equilibrium, iterating equation (2.12) forward shows that the value of staying on the country side forever is equal to moving to the city at every point in time. Note that because of the curvature on the rural technology, there will always be agents that remain (optimally) on the country side. Moreover, I assume that the technology in the city is sufficiently productive to ensure an interior solution.

Competitive Equilibrium in Small Open Economy:
I define a competitive equilibrium based on Buera and Shin (2017). In order to do so, I need to introduce the joint distribution $G(t_m, T, a_t, \varphi; t)$ which keeps track of the migration decision, catch-up growth, and the type draw φ of each household and allows me to go from household choices to aggregate outcomes.

A competitive equilibrium in the small open economy consists of a sequence of joint distributions $\{G(t_m, T, a_t, \varphi; t)\}_{t \in R}$, household asset, consumption, and migration decisions

[128] For reasonable parameters of risk aversion, the rural wage would be higher in a frictionless environment because rural workers have to be compensated for the forgone opportunity of high-growth in the city. All empirical evidence, however, suggest that urban wages are much higher than rural once, albeit partially driven by selection (Young 2013; Hicks et al. 2017).

[129] The proportionality assumption is important to obtain a simple law of motion for the flow of workers out of agriculture. While frictions between urban and rural areas are well documented, the urban-rural wage gap is not at the center of my model. Accordingly, I chose the simplest possible way to correct for the counterfactual implication of higher wages on the countryside.

$\{c_t^i, a_t^i, t_m^i\}_{t\in R, i\in[0,1]}$, as well as wages $\{w_t^r, w_t^u\}_{t\in R}$ such that

- households maximize utility given (2.11),(2.13), the exogenous income process y_t^i, the type draw φ^i or the distribution $F(\varphi)$ (if $t < T^i$), and the world interest rate r^*
- urban and rural firms maximize profits given technological constraints (2.2),(2.3)
- the joint distribution G_t evolves consistent with agent's migration, and consumption decisions, as well as the arrival rate of drawing your type, the distribution of types $F(\varphi)$, and the labor resource constraint $(B12), (2.17)$
- labor markets $(2.16), (2.17)$ clear, and goods markets are consistent with asset markets (2.18)
- the no-Ponzi-scheme condition (2.14) is satisfied, i.e.

$$\lim_{t\to\infty} \exp(-r^* t) a_t^i \geq 0, \forall i \in [0,1]. \tag{2.14}$$

Labor and Goods Market clearing:
The labor market clearing condition has to hold for each sector separately, and needs to be consistent with a law of motion that governs the influx of farmers into the urban centers. Let L_t^r be the mass of agents on the country side. Define $M_{t,0}$ and $M_{t,1}$ as the measure of households that are in the high growth regime, or have already drawn their type, respectively. Hence, the measure of urban households reads $L_t^u = M_{t,0} + M_{t,1}$. Since labor is supplied inelastically within each sector I can immediately compute total sectoral output

$$Y_t^r = (L_t^r)^\alpha \tag{2.15}$$

$$\begin{aligned} Y_t^u = A_t &\int_{i\in M_0} \int_{t_{m,i}}^{t} \exp((g-g^*)[s-t_{m,i}])dsdi \\ &+ A_t \int_{i\in M_1} \varphi_i \int_{t_{m,i}}^{T_i} \exp((g-g^*)[s-t_{m,i}])dsdi. \end{aligned} \tag{2.16}$$

Labor market clearing then requires that the wage is such that produces break even. Importantly, the mass of agents in the city is not the same as the effective labor supply. Finally, there is an adding-up constraint that connects the two sectors with each other

$$L_t^u + L_t^r = 1. \tag{2.17}$$

Goods market clearing in the small open economy allows for a surplus or a deficit, which constitutes international capital flows and leads to changes in the net foreign asset position. Simply integrating over the individual budget constraints gives this aggregate market clearing condition

$$\int_{i\in S_u} \dot{a}_i di = Y_t^u + r^* A_t^b - C_t^u, \tag{2.18}$$

where A_t^b denotes aggregate bond holdings (b for bond).

2.4.1 Solution of the Household Problem Model

Before turning to the household problem, I need to define how the ROW grows. This matters since the economy is catching up with the industrialized economy. Second, the interest rate, while exogenous to the small open economy, is endogenously determined by the ROW and reads $r^* = \eta g^* + \rho$. Implicit here is the assumption that the ROW grows along a balanced growth path of rate g^* with no uncertainty and identical CRRA preferences. Next, I focus on the consumption problem of urban households which can be solved backwards. First, note that agents that have learned their type grow their income at the same rate as the industrialized world, and there is no additional source of uncertainty. Hence, the standard Euler equation holds

$$\frac{\dot{c}_s}{c_s} = \frac{1}{\eta}\left(r^* - \rho\right). \tag{2.19}$$

That means that consumption has to be equal to $c_t = y_t + \left[g^*\left(\eta - 1\right) + \rho\right] a_t$, which follows from the household budget constraint after imposing $\frac{\dot{c}}{c} = g^*$ along the balanced growth path. A consequence of this is that I can derive the value function in closed form for agents who know their type

$$\begin{aligned} V\left(a_T; \varphi, T\right) &= \int_T^\infty \exp\left(-\rho\left(s - T\right)\right) \frac{c_s^{1-\eta}}{1-\eta} ds \\ &= \int_T^\infty \exp\left(-\rho\left(s - T\right)\right) \frac{\left(y_s + \left[g^*\left(\eta - 1\right) + \rho\right] a_s\right)^{1-\eta}}{1-\eta} ds \\ &= \frac{\left(y_T + \left[g^*\left(\eta - 1\right) + \rho\right] a_T\right)^{1-\eta}}{1-\eta} \left\{\frac{1}{g^*\left(\eta - 1\right) + \rho}\right\}. \end{aligned} \tag{2.20}$$

Note that the value function is concave in assets a, and negative for values of risk aversion above unity. I focus on the empirically relevant case with $\eta > 1$ but the model is valid for any positive coefficient of risk aversion.

Due to the Poisson arrival one can show that the household problem in the high-growth regime simplifies to 2.21

$$V_{t_m} = \max_{c_s} \int_{t_m}^{\infty} exp\left(-\left(\lambda+\rho\right)\left[s-t_m\right]\right)\left[\frac{c_s^{1-\eta}}{1-\eta}+\lambda\mathbb{E}_{\varphi}\left[V\left(\varphi y_s, a_s\right)\right]\right] ds, \tag{2.21}$$

which is a version of the perpetual youth model of Blanchard and Yaari. Instead of dying at rate λ, however, households transition into a "stable life" of balanced growth.

Together with the budget constraint and the transversality condition one can use a standard Hamiltonian to solve the problem. The concavity of the utility function together with a compact budget constraint ensures that the solution to the household problem is unique. Then, let q_s be the co-state variable and define the present-value Hamiltonian:

$$H = \exp\left(-\left(\lambda+\rho\right)s\right)\left\{\frac{c_s^{1-\eta}}{1-\eta}+\lambda\mathbb{E}_{\varphi}\left[V\left(\varphi y_s, a_s\right)\right]\right\} + q_s\left[r^* a_s + y_s - c_s\right]. \tag{2.22}$$

After plugging (2.20) into (2.22), the optimality conditions with respect to c_s and a_s read

$$dc: \quad q_s = \exp\left(-\left(\lambda+\rho\right)s\right)\left(\frac{1}{c_s}\right)^{\eta} \tag{2.23}$$

$$da: \quad -\dot{q} = \exp\left(-\left(\lambda+\rho\right)s\right)\lambda\mathbb{E}_{\varphi}\left[\left(\frac{1}{\left[g^*\left(\eta-1\right)+\rho\right]+\varphi y_s}\right)^{\eta}\right] + q_s\left[r^*\right]. \tag{2.24}$$

Taking logs of (2.23) and differentiating with respect to s, plugging in (2.24), and using $g^* = \frac{r^*-\rho}{\eta}$ yields the law of motion for consumption for households in the city that are on the fast growth path and have not drawn their type yet

$$\frac{\dot{c}_s}{c_s} = \frac{\lambda}{\eta}\left\{\mathbb{E}_{\varphi}\left[\left(\frac{\left[g^*\left(\eta-1\right)+\rho\right]a_s+\varphi y_s}{c_s}\right)^{-\eta}\right]-1\right\}+g^*. \tag{2.25}$$

This simple law of motion of consumption of households in the high-growth regime contains the key argument proposed in this paper. Note the slight inconsistency of notation. I add the type φ in front of income y. That is meant to make explicit the type risk. One might also put the following expressions into the denominator of (2.25) where $\lim_{\Delta\downarrow 0} y_{s+\Delta} = \varphi y_s$ and $\lim_{\Delta\downarrow 0} a_{s+\Delta} = a_s$. Income is continuous except at the point in time when the household draws their type. Now I discuss several propositions that can be derived from this simple model, especially equation (2.25).

2.4.2 Theoretical Results

The first result, although well known (Schechtman 1976; Gourinchas and Parker 2002), is worth pointing out again. For CRRA preferences, the marginal utility of consumption at zero is infinity. Therefore, agents will never borrow. An important caveat is in order. The differential equations, and especially the law of motion of consumption (2.25) are derived under the implicit assumption that the value function is differentiable. This need not be the case around an asset position that is zero.[130]

Let $\underline{\varphi}$ be the lower bound of the state space of φ. Then, agents' borrowing decisions in the high growth regime will always respect the following inequality $\frac{a_t}{y_t}\left[(\eta-1)\, g^* + \rho\right] > -\underline{\varphi}$. By contradiction, suppose an agent borrows above the borrowing level. Then there is a range of values $\varphi \in [\underline{\varphi}, -\frac{a_t}{y_t}\left[(\eta-1)\, g^* + \rho\right]]$ where the agent would have to consume weakly below zero. Furthermore, assuming that $\int_{\underline{\varphi}}^{-\frac{a_t}{y_t}[(\eta-1)g^*+\rho]} dF(\varphi) = \epsilon > 0$, then such a borrowing position would yield an expected continuation value of minus infinity since, loosely speaking, $\epsilon * -\infty = -\infty$. This is strictly worse than a consumption profile where income equals consumption at all points in time. Hence this cannot be a solution to the household problem. Proposition 2.4.2 is a well known result, that has received little consideration in the context of the capital flow puzzle. If we let the lower bound of the state space of φ go to zero, even an extremely small level of risk in the economy is sufficient to prevent fast-growing households from borrowing. Of course, this does not help us understand why there are capital outflows, i.e. strong savings pressure.[131] Consumption growth for households that converge at the high growth rate g_h is strictly larger than consumption growth of high-growth households in a world without human capital risk, which in turn is strictly larger than consumption growth in the industrialized world g^*. See appendix. First, consumption growth and hence precautionary savings are higher in a world where there is a non-degenerate inequality distribution which can be shown by using Jensen's inequality.[132] This result is a necessary but not sufficient condition to generate capital outflows. The reason that consumption growth is higher in the small open economy is solely due to risk. It is easy to show that once there is no income risk in the form of random time spent in the high growth

[130]Implicit in the proof is the assumption that zero or negative levels of consumption yield a utility of minus infinity. The marginal utility is also not continuous at zero.

[131]Note that in the phase diagram analysis that I perform in the appendix I focus on the case where the long-run steady state is such that it automatically respects the inequality in proposition 2.4.2. If it doesn't, it is clear that the solution to the household problem will be at a corner with an asset position of zero.

[132]The positive link between consumption growth and precautionary savings arises due to the budget constraint. High consumption growth that also respects the budget constraint, means relatively small initial consumption. In turn, this implies relatively high savings at early periods.

regime, and the type draw, consumption growth in the small open economy is $\frac{r^*-\rho}{\eta} = g^*$. If households in the small open economy are still growing faster than the rest of the world for some deterministic time, then this would lead to initial borrowing and persistent trade balance deficits during the catch-up phase.[133] A model without income inequality, i.e. $\varphi_i = 1, \forall i$, cannot generate capital outflows. See appendix. Proposition 2.4.2 is a key result and shows that random convergence cannot generate the capital flow patterns we observe in the data. It highlights the need for an additional source of risk in order to explain the puzzle. The intuition for this result is as follows: If there is no human capital risk in the form of the type draw, then the only risk that households are exposed to is the random time they spend in the high growth regime. This type of risk, however, only represent upside risk and is therefore not sufficient to induce precautionary savings. At any point in time, a household will be better off in the future, no matter how long they are in the high growth regime and thus will want to borrow against future income. In contrast, in a model with human capital risk, some households can actually be worse off, at least for some time, despite strong convergence growth. Only this type of risk can leads to a precautionary savings motive that can dominate the consumption-smoothing motive.

The next proposition deals with the case when expectations about convergence growth are biased. One way to shut down the consumption smoothing motive of households is to make them believe that there is no convergence growth. My model allows me to consider this case effortlessly. Let $\tilde{\lambda}$ be the households' belief about the arrival rate of the low-growth regime, i.e. the relevant parameter for the household Euler equation. A high $\tilde{\lambda}$ relative to the correct λ represents "pessimistic" expectations in terms of convergence growth since households think their convergence period is, on average, shorter than it actually is.[134] Without human capital risk, there are no capital outflows, even if expectations about convergence are downward biased, $\tilde{\lambda} > \lambda$. See appendix. This proposition makes clear that biased expectations per se are not a solution to the capital flow puzzle.[135] As long as the future looks always brighter than today, even if the household underestimates "how bright" it looks, a forward-looking agent will want to borrow against future income. Mathematically, as long as $\tilde{\lambda} > 0$, household income is bounded below by continuous growth in the low-growth regime, and households will consume slightly above that consistent with the consumption smoothing

[133] This is consistent with the simple models in chapter 2 in Uribe and Schmitt-Grohé (2017). The trade balance usually follows a unit root process in these types of modes – this would also be true in the context of this model without any risk.

[134] Recall that average time spend in the high growth regime is inversely related to the arrival rate.

[135] Thanks to Pablo Ottonello for drawing my attention to this point.

motive. Continued presence in the high growth regime will then seem a bit like a surprise, and the household increases consumption enough to remain a borrower. Once human capital risk is incorporated, however, downward biased expectations help to generate capital outflows. The reason is that convergence growth provides a consumption smoothing motive counteracting the precautionary savings motive. The less convergence a household expects, the less powerful is the motive to smooth consumption. For a sufficient amount of human capital risk, parameterized here in the form of the distribution $F(\varphi)$, there exists a unique equilibrium with capital outflows driven by households' precautionary asset accumulation in the high-growth regime. A necessary and sufficient condition for this equilibrium to obtain is

$$\frac{g_h - g^*}{\lambda}\eta < \mathbb{E}_{\varphi}\left[\left(\frac{1}{\varphi}\right)^{\eta}\right] - 1 \tag{2.26}$$

See appendix. Proposition 2.4.2 is the main result of the model. It shows that there is a set of parameter values that can generate capital outflows despite convergence growth. The left-hand side of the inequality in proposition 2.4.2 represents the consumption smoothing force, while the right hand side reflects human capital risk. Stronger convergence growth governed by the convergence rate $g - g^*$ as well as the average time spent in the high-growth regime $\frac{1}{\lambda}$ counteract capital outflows, while greater human capital risk induces outflows. Note that by Jensen's inequality and the assumption that the type draw is centered around unity, the right hand side of 2.26 is always larger than zero but not necessarily larger than the left hand side. Moreover, note the ambiguous role played by the coefficient of relative risk aversion η, pushing up both the left hand side and right hand side of the inequality. This reflects that curvature in the utility function induces both inter- and intratemporal smoothing.[136] In the next subsection I provide a simple calibration of the model that assumes a log normal distribution with $\log(\varphi) \sim N(-\frac{\sigma^2}{2}, \sigma^2)$. In that case the right hand side equals $\exp\left(\frac{\eta\sigma^2}{2}[1+\eta]\right) - 1$. While a common assumption, the proposition generalizes beyond the log normal case. In the appendix I show that the result holds for arbitrary distribution with sufficient dispersion. The only restrictions are that the distribution takes on non-negative values and has a mean of one.

There exist no closed form solution for the transitional dynamics during the high-growth phase of a household, a feature shared with the baseline neoclassical model. It is possible, however, to study the general qualitative properties of the non-linear system using a properly normalized version of the household Euler equation in the high-growth regime (2.25), similar

[136]Under reasonable parameter restrictions a higher coefficient of relative risk aversion η also induces capital outflows.

to the analysis of the neoclassical growth model in continuous time. This normalization allows me to solve for a "pseudo" steady state, which is the steady state that households converge to if they were to stay in the high growth regime forever.[137] Dividing by household income turns out to do the trick. Then, uniqueness and the qualitative properties of convergence to the steady state can be characterized using a phase diagram approach.

The qualitative analysis reveals that household consumption and assets grow at a rate higher than income initially, and converge from above to the growth rate of income (in the high-growth regime), assuming that the inequality in proposition 2.4.2 is satisfied. This leads to a constant consumption-to-income and asset-to-income ratios. If the type risk is not large enough to generate capital outflows, then consumption growth converges from below to the growth rate of income. Of course, agents are pulled out of the high-growth regime randomly according to the Poisson Process. Hence, they never fully reach the steady state. The qualitative predictions still hold as they are valid along the transition path. I obtain the following qualitative predictions from the phase diagram analysis carried out in the appendix:

- consumption growth and asset growth are above income growth while the household resides in the high-growth regime

- a higher level of human capital risk in the form of inequality in the low-growth regime induces a stronger precautionary savings motive and generates higher consumption and asset growth, and an initially lower consumption-to-income ratio for households that just entered the city, as well as a higher asset-to-income ratio in the long run.

Both predictions are standard in the literature on precautionary savings. Obtaining those results in the presence of catch-up growth without extreme business cycle or unemployment risk is not. The next section discusses the key differences to the canonical precautionary savings model that allow for this possibility.

2.4.3 Discussion of Income Process

In the presence of powerful income growth, business cycle risk is not sufficient to generate savings pressure that dominates the consumption smoothing force as discussed before. A similar result emerges when focusing on incomplete-market models with idiosyncratic household risk. The canonical model is build on an income process that consists of a transitory

[137] Normalizing by income was not a random conjecture. In fact, Carroll (1994) shows that asset-to-income ratios are stationary in incomplete market models with human capital risk in the form of a random walk in the log of income. His insight extends to the framework at hand.

shock and a persistent shock, together with a homogeneous trend growth rate. Let P be permanent income and Y be the current income, then agents in the economy face the following income process

$$Y_{t+1} = (1 + g_h)P_{t+1}u_{t+1}$$

with

$$P_{t+1} = P_t n_{t+1}$$

and u and n being iid random draws centered around unity, usually of the log normal type.[138] While this income process has been employed very successfully in various setting, see for example Kaboski and Townsend (2011), for reasonable parameter values it is very hard to generate capital outflows during an episode of fast growth because quantitatively the consumption smoothing force dominates (Coeurdacier, Rey, and Winant 2019).

One of the key differences between the canonical model and the approach chosen here is that growth itself is unevenly distributed across households. Of course, in an incomplete market model measured growth as the change in the log of income is always heterogeneous across households. But what I am concerned with here is the growth rate g_h that is assumed to be uniform in the benchmark models. This uniformly high growth rate is the quantitatively troubling piece as it induces households across the board to smooth consumption. In the appendix in section B.4 I simulate a version of the model where catch-up growth is evenly distributed. The variance of the type shock would have to be more than ten times larger than what I need in the baseline calibration with an uneven growth process. Moreover, one can show that if the type draw is the only source of risk all households would display optimal consumption growth at g^*, even the ones that experience income growth at g_h. Risky growth, then, is the key piece that delivers a realistic co-movement between income and consumption.

In contrast in the model economy here there is an important distinction between the mean and the median household. Note that much aggregate growth is directly related to an emerging thick right tail of the income distribution as depicted in a simulation exercise in figure 22 in the next section. From an individual household's point of view, landing anywhere on the right tail is a very unlikely outcome that, at time zero when the consumption plan is made, is also heavily "effectively" discounted by the curvature on the utility function. As a consequence, aggregate growth originating from a rising right tail induces much less

[138] Of course, more general shock processes can and have been used (Blundell, Pistaferri, and I. Preston 2008) as well as specifications with heterogeneous income profiles as in Guvenen (2007).

consumption smoothing pressure relative to a world where every household gets to participate in the average growth rate, just as in the canonical incomplete market model.

To see this formally, consider a decision maker with additive preferences over a consumption good of the following type

$$U = \mathbb{E}\left[\log(c)\right].$$

Now the lottery that the agent is facing is such that their initial endowment $c_0 = 1$ is growing exponentially at rate $g_h - g^*$ for some random time. The agent then consumes everything at once, ignoring the time dimension. This leads to a payoff that is following a Pareto distribution, and hence the expectation of the log is simply the average of an exponential distribution, given by $U_1 = \frac{g_h - g^*}{\lambda}$. In contrast, consider a lottery that is degenerate where the agent receives the average over all outcomes of the previous lottery. This means that the utility of the agent is given by $U_2 = \log\left(\frac{\lambda}{\lambda - [g_h - g^*]}\right)$. It immediately follows that $U_1 < U_2$, trivially so, since the utility function is concave. The more interesting aspect, however, is what happens as the tail-coefficient converges to unity. In that case, expected utility for any individual household is still well-defined by U_1. On the other hand, the average outcome is shooting off to infinity, and so does U_2. We can see in this simple example how we can construct and arbitrarily large growth miracle with infinite catch-up. The effect of this growth miracle on time zero expected utility is still quite modest, precisely because the household heavily "discounts" the possibility of ending up on the right tail. This analogy carries over to our agents in the fast growth regime that are able to smooth consumption, if they want to. Growth in the tail induces much less smoothing compared to evenly distributed deterministic household income growth, keeping the aggregate growth miracle fixed.

Clearly, this income process is rather stylized. It cannot match the micro household income data that are characterized by persistent period-by-period shocks (Blundell, Pistaferri, and I. Preston 2008). Conceptually, it is easy to add a source of noise to the income process, which would leave all conclusions unchanged and only raise overall savings pressure due to higher risk. Of course, nothing could be solved in closed form any longer. What is necessary, however, and less standard, is that there exists a multiplicative type draw that can shift fast-growing households' income up or down substantially. This stylized structure cleanly separates out growth from risk, and leads to closed form expressions for key statistics in an infinite-horizon forward-looking economy that experiences structural change and catch-up growth. In the next section we will see that this income process leads to aggregate growth and saving dynamics that look very much like an actual growth miracle.

2.4.4 Productive Capital

I abstract away from capital accumulation in the baseline model. In principal, the current account flows could be driven by either relatively high saving rates or relatively low rates of investment. It is well known, however, that the rate of capital accumulation tend to be very high for miracle economies (Young 1995). Consequently, one needs to look for an explanation why saving rates exceed already high investment rates along the transition path, a point developed carefully in Gourinchas and Jeanne (2013). Hence, I abstract away from capital accumulation in the baseline model to focus on the two key forces at play: consumption smoothing and precautionary savings. Note that the puzzling household saving rates emerge independently of the supply side of the economy, simply because the PIH suggests that households should be smoothing consumption along the transition path. Yet, numerous studies using Chinese micro data have documented rising saving rates of urban households (M. D. Chamon and E. S. Prasad 2010; M. Chamon, K. Liu, and Eswar Prasad 2013). Nonetheless, in Appendix B.5, I sketch out a version of the model with productive capital that leaves the main theoretical insights unchanged. From a quantitative point of view, however, it is fair to admit that the capital flow puzzle in a model with productive capital is harder to solve since capital and labor are complements. If human capital grows fast, then the rate of return to capital increases as well, ceteris paribus. The standard fix here would be to introduce additional financial frictions as in Z. Song, Storesletten, and Zilibotti (2011) or Buera and Shin (2017) so that domestic entrepreneurs are cut off from financial markets. I hope to offer a complementary view to the large literature on financial frictions that highlights the importance of urban-rural differences and uneven growth in urban labor markets to understand the demand for safe assets of ordinary non-capitalist households along the transition path.

2.5 Calibration

To conclude this theoretical section, I solve for the household equilibrium dynamics, as well as the aggregate trajectory of the economy. While the transitional saving and consumption dynamics of households during the high-growth phase need to be simulated, the flow of workers out of the agricultural sector, the share of agents that have already drawn their type can be solved in closed form, as well as the evolving income distribution.

The goal of this calibration is to show that the elements introduced in the model can give rise to realistic "miracle growth dynamics". As a starting point, I set $g = 7\%$, $g^* = 2\%$, and

$\lambda = \frac{7}{100}$ which implies an average time spent in the high-growth regime for each household of a bit less than 15 years. Expected income growth after moving to the city, after netting out the effect of the urban-rural wage gap, equals

$$\begin{aligned}
\mathbb{E}_i\left[\exp\left([g-g^*]s_i\right)\varphi_i\right] &= \mathbb{E}_i\left[\varphi_i\right]\int_1^\infty \frac{\lambda}{g-g^*} y^{-\frac{\lambda}{g-g^*}} dy \\
&= 1 + \frac{g-g^*}{\lambda-(g-g^*)}
\end{aligned}$$

where I use both the independence of the type draw, as well as the fact that the Poisson process leads to Pareto-distributed income due to catch-up growth.[139] For the parameters picked, this amounts to convergence growth of 250%, which is multiple times larger than the growth miracle in Buera and Shin (2017). After setting the urban rural wage gap to one, this would imply measured GDP per capita growth of 500%![140]

For the coefficient of relative risk aversion (that simultaneously pins down the elasticity of inter-temporal substitution) I pick $\eta = 2$. The results are sensitive to this number. Proposition 2.4.2 precisely shows how the parameters of the model, and in particular η, pin down the direction of capital flows.[141] I model the type distribution as a draw from a Log-Normal distribution, i.e.

$$log(\varphi) \sim N(-\frac{\sigma^2}{2}, \sigma^2). \tag{2.27}$$

This particular representation ensures that $\mathbb{E}[\varphi] = 1 \quad \forall\sigma$, so as to isolate the effect of higher inequality due to a larger variance from first-order effects that would otherwise shift up the mean and obscure analysis. The specific distributional assumption is not essential, but leads to an empirically plausible stationary income distribution.[142] In order to calibrate the variance of the log of the type draw, I try to match the level of household inequality in the United States, implicitly assuming that the miracle economy converges to this long-run

[139] See Jones and Kim (2018) for a derivation.

[140] The GDP gains are larger than the welfare-gains of the growth miracle for two reasons. First, we would have to be properly discount future growth that only materializes far in the future. Second, there is a distributional cost of the growth miracle. Ex ante, the household risk embodied in the growth miracle leads to an even higher effective discount factor. The actual growth miracle is also slightly smaller because of the agents $M_{0,1}$ in the city that do not experience miracle growth. That is, the per capita growth rate after netting out the urban-rural wage gap would be roughtly 206% instead of 250%.

[141] Kaboski and Townsend (2011) and Gourinchas and Parker (2002) estimate the coefficient of relative risk aversion between 1 and 2 based on structural estimation. Regression evidence suggest a larger coefficient of relative risk aversion (Hall 1988).

[142] Log normality is a common assumption in the context of cross sectional wage distributions, albeit not innocuous (Guvenen, Karahan, et al. 2015). Note that due to uneven growth the right tail of the log of income will be dominated by the exponential distribution.

equilibrium. Noting the the variance of the log of income for the stationary distribution is $\left(\frac{g-g^*}{\lambda}\right)^2 + \sigma^2$, I pick $\sigma^2 = .49$ so that the log variance ends up being close to one. This is consistent with measures of household income inequality in the US (D. Krueger, Mitman, and Perri 2016).[143]

parameter	baseline value
discount factor	$\rho = .01$
coefficient of relative risk aversion	$\eta = 2$
log-variance of type draw	$\sigma^2 = .49$
Poisson arrival rate	$\lambda = 0.07$
industrialized growth	$g^* = 2\%$
miracle growth	$g = 7\%$
urban-rural wage gap	$W_{gap} = 100\%$
elasticity of agr. output with respect to labor	$\alpha = \frac{5}{9}$
initial agr. share	$L_0^r = 75\%$
initial share of agents that know their type	$M_{0,1} = 17.5\%$

$\hat{W_{gap}}$ denotes the wage gap between urban and rural individuals before the urban individual could accumulate additional human capital, i.e. $\frac{w^u_{t+\Delta} - w^r_t}{w^r_t}$. This wage gap is going to be another source of convergence. I set it to unity, which I view as a lower bound. Fan and Zou (2019) suggest that the wage of an unskilled urban worker is three times that of a rural worker in China, and they provide evidence that this ratio is relatively stable.[144] The agricultural share is set high at 75% which leads to powerful catch-up growth. Note that I assume that an initial share of households already has learned their type and is in group $M_{0,1}$. If I assumed that all agents that are in the city at time zero started growing fast, then this mass point would dominate the dynamics of aggregate savings completely.[145] Lastly, I need to set the parameter α which determines the curvature on the rural production function. This is a key parameter as it governs the speed at which households move out of the agricultural sector. The next subsection shows how to estimate α through the lens of the model.

[143] Actually, the relevant statistic here is provided by De Magalhães and Santaeulàlia-Llopis (2018a) who themselves rely on unpublished data by D. Krueger, Mitman, and Perri (2016). Alternatively, recent work by Guvenen, Kaplan, et al. (2017) measures the variance of the log of income from tax returns around .8.

[144] Because we have allowed for a wedge τ when deriving the migration arbitrage equation we can pick the wedge so that it delivers the empirically observed wage gap between urban and rural workers.

[145] An unpleasant side effect of that is that the aggregate saving rate would not deliver a hump-shaped pattern. Instead, it would mimic the individual saving rate, first shooting up and then monotonically declining.

2.5.1 Structural Change

The model features a transition of the economy from agricultural production towards non-agricultural production, consistent with the fast-paced structural change in miracle economies. In the appendix I go through all the steps in detail, while I report only the final result in form of a law of motion of agricultural employment here

$$L_t^r = L_0^r \exp\left(-\frac{g^*}{1-\alpha}t\right). \tag{2.28}$$

The intuition for this result is straightforward: at every point in time, the relative attractiveness of the city increases by g^* percent due to productivity growth in A_t^u. Of course, there is convergence growth as well, but this scales up household income by a constant factor in expectation, i.e. is fixed over time. As a consequence, households only remain on the country side if their income increases by g^* percent as well. For that to be the case the model requires a continuous inflow of workers into the urban sector, given by equation 2.28. This law of motion of agricultural employment leads to the following estimating equation,

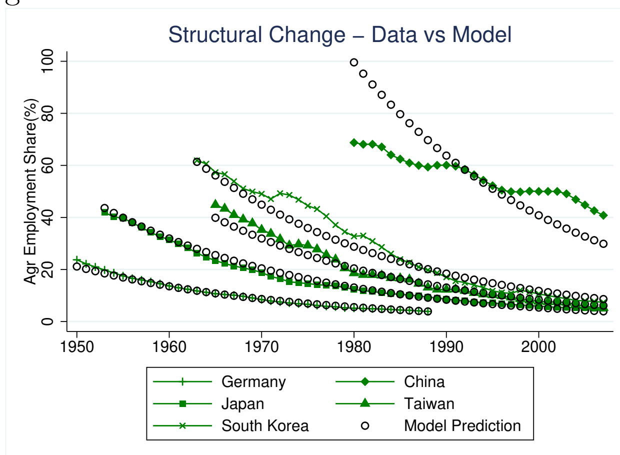
$$\log\left(L_{t,c}^r\right) = \beta_{0,c} + \beta_1 * t + \epsilon_{t,c}, \tag{2.29}$$

where I added a random error term. The estimating equation (2.29) can efficiently be estimated using a random effects model across a sample of miracle economies. In doing so, I acknowledge that different initial conditions lead to different initial agricultural shares while maintaining that the technology coefficient α is constant across economies. Table B17 in the appendix reports the regression results for a sample of five miracle economies (Germany, Taiwan, Japan, Korea, and China) beginning from the point in time when they started to reform, following Buera and Shin (2013). Maintaining that g^* is equal to 2 %, consistent with long-run growth in developed economies over the twentieth century (R. E. J. Lucas 2018), implies an estimate of $\hat{\alpha} \approx \frac{5}{9}$.[146] Figure 20 shows the fit of model-implied structural change relative to the observed agricultural employment. The fit is nearly perfect for all economies but China. One wonders whether this is a manifestation of the detrimental effects of the Hukou system potentially hampering the process of structural change, as discussed in

[146]Note that the sample of countries is too small for the fixed estimator to be consistent based on cross sectional variation. Nevertheless, the results of the random and fixed effects model are very similar. The fit of the model in a R-squared sense is excellent, and accounts for more than 95% of variation in the data.

Tombe and X. Zhu (2019).[147]

Figure 20: Prediction based on random effects model.



2.5.2 Growth dynamics

Let g_{agg} denote the aggregate growth rate. Since labor is normalized to one, and constant, this is also the per capita growth rate. We can start computing the aggregate growth rate. Despite idiosyncratic type draws and movers whose wage jumps up, one can show that this randomness washes out in the aggregate,[148]

$$g_{agg}(t) = \underbrace{(g - g^*)\frac{Y_{0,t}}{Y_t} + \left(\frac{g^*}{1-\alpha}\right)\frac{\hat{W}_{gap}L_t^r}{Y_t}}_{\text{catch-up growth}} + \underbrace{g^*}_{\text{long run growth}} \tag{2.30}$$

This derivation separates catch-up growth from long-run growth, and catch-up growth itself is generated by fast growth (first term) as well as the urban-rural wage gap (second term). The calibration is such that the urban-rural wage gap contributes to convergence growth. In the long run, the aggregate growth rate will of course be equal to g^*. It is worth pointing

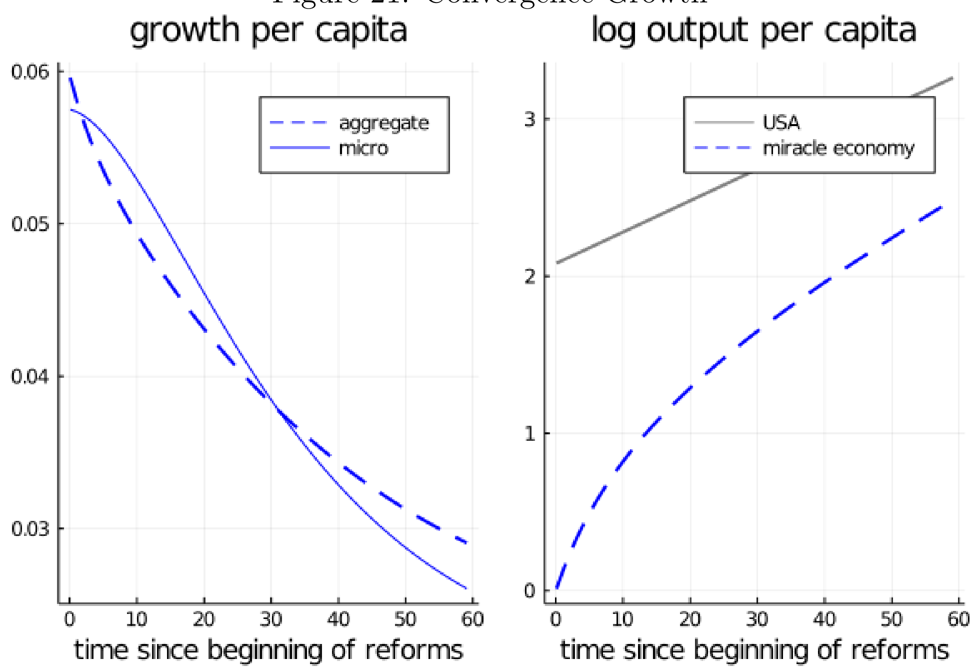
[147] The reader might wonder whether risk in the urban sectors adds to the rural-urban wage gap in the form of a compensating differential: this is indeed the case as I show in the appendix in subsection B.1.1.

[148] Derivation:

$$\begin{aligned} g_{agg}(t) = \frac{dY_t/dt}{Y_t} &= g\frac{Y_{0,t}}{Y_t} + g^*\frac{Y_{1,t}}{Y_t} + \frac{1}{Y_t}P\left(i \in \dot{M}_{0,t}\right)\frac{1}{w_t^r}\frac{w_{t+\Delta}^u - w_t^r}{\Delta} + \frac{L_t^\alpha}{Y_t}g^* \\ &= g\frac{Y_{0,t}}{Y_t} + g^*\frac{Y_{1,t}}{Y_t} + \frac{L_t}{Y_t}\frac{g^*}{1-\alpha}\hat{W}_{gap} + \frac{L_t^\alpha}{Y_t}g^* \end{aligned}$$

out that the part of convergence growth that is due to the urban-rural wage gap from a welfare point of view would be undone by the utility cost τ. Figure 21 plots the aggregate growth rate for the parameter values chosen in the calibration. The growth miracle is sizable, and comparable to the experience of Taiwan. Note that the aggregate growth rate can be larger than the growth rate measured in the micro data. In the data, aggregate growth is larger than average household growth rates (Santaeulalia-Llopis and Zheng 2018). The the canonical income process fails to capture this. In the context at hand, it arises for two reasons. First, note that the urban-rural wage gap contributes to higher aggregate growth. It is likely that this income jump is missed in the micro data, or even discarded on purpose as an outlier. Second, fast growing household achieve a relatively higher share in aggregate GDP in the long run, thus dominating aggregate dynamics and raising the growth rate relative to the simple average in the micro data. Figure 21 also shows the convergence in terms of log output per capita relative to the United States.[149]

Figure 21: Convergence Growth



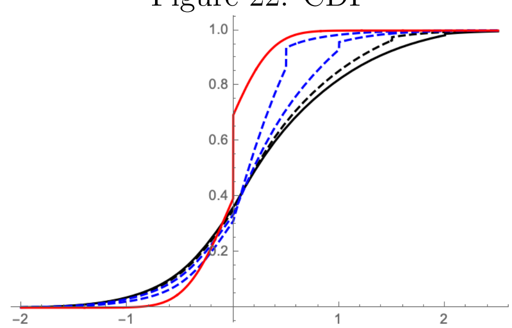
[149]Total output of the miracle economy at time zero is normalized to unity. The US is assumed to be 8 times as rich. This normalization together with an urban-rural wage gap of 100% implies technology coefficients $(A_0^u, A_0^r) = (1.600, 0.704)$. These values are consistent with the equilibrium definition only for the right value of τ. The implied value of τ can be backed out after simulating the model.

2.5.3 Income Inequality

The model delivers closed form solutions for the distribution of income along the transition path. The derivation can be found in the appendix in subsection B.2. It is well known that heterogeneous growth rates give rise to a fat-tailed income distribution (Luttmer 2011; Gabaix et al. 2016; Aoki and Nirei 2017). The Pareto-tail in the model at hand is given by $\frac{g-g^*}{\lambda}$.

Figure 22 shows the CDF of the log of normalized income for different decades. Importantly, while the distribution fans out overall, more and more weight is being shifted to the right tail that is composed of households that remained in the high growth regime for a relatively long time.

Figure 22: CDF



Simulated CDF of log of income, normalized by g^*t. The figure plots the distribution at time 0, 10, 20, 30, and 40 years in after the reforms started. Lines to the right are later periods compared to the left. The CDF displays jumps that stem from the initial share of households that start growing fast. This mass point disappears over time as more and more agents are pulled out of the high growth regime.

The normalized stationary distribution of the log of income is given by the exponentially modified Gaussian distribution. This emerges as the sum of two independent random variables, one of which is normal (what I call the type draw) and one of which is exponentially distributed (time spent in high growth regime scaled by the growth rates).[150]

Armed with this CDF I can compute the log variance of income for non-agricultural households displayed in table 6. The results are broadly inline with the fast rise of inequality in China.

[150]The economy starts with an initial distribution that at every point in time converges closer to the limiting distribution. This limiting stationary density reads $f(x;\mu,\sigma,\frac{\lambda}{g-g^*}) = \frac{\lambda}{2[g-g^*]}\exp\left(\frac{\lambda}{2[g-g^*]}\left(\sigma^2\left(\frac{\lambda}{g-g^*}-1\right)-2x\right)\right)\operatorname{erfc}\left(\frac{\sigma^2(\frac{\lambda}{g-g^*}-\frac{1}{2})-x}{\sqrt{2}\sigma}\right)$ where erfc is the complementary error function $\operatorname{erfc}(z)=\frac{2}{\sqrt{\pi}}\int_z^\infty \exp(-t^2)dt$.

Table 6: Inequality – Data and Model

year	data	model
1980	NA	0.07
1988	0.17	0.16
1995	0.37	0.26
2002	0.36	0.36
2013	0.5	0.5

The data moments come from the CHIP and concern non-agricultural occupations for household heads between 23 till 60 years of age, smallest 2 percent of income realizations dropped. Variance of log of income of the model in last column.

Two more points are noteworthy. First, and most obviously, there is a direct link between expected convergence growth ($\frac{\lambda}{g-g^*}$) and top income inequality. For finite aggregate convergence we also need $\lambda > g - g^*$. The reason is that the fraction of agents that experience fast growth, measured in terms of their share of GDP, converges to unity in the limit when $\lambda \leq g - g^*$. A smaller and smaller share of fast-growing agents would eventually account for 100% of GDP, leading to a growth rate of g_h forever. Second, even though ex-post inequality matters for ex-ante savings pressure since λ, g_h, and g^* all impact the Euler equation, only the uncertainty related to the type draw φ is able to generate precautionary savings that tilt the balance toward capital outflows during growth miracles. This relates directly to proposition 2.4.2 and suggest that the risk that matters for precautionary savings and capital outflows is the dispersion measured in the middle and the left tail of the income distribution.

Lastly, I would like to emphasize one important but subtle distinction: the model does no require that inequality is necessarily rising, and the relationship between growth and inequality is admittedly much more complex than in this stylized model. What is needed for the mechanism to go through is that human capital is more risky in the non-agricultural sector. Whether inequality increases or not also depends on the level of inequality that prevailed in the pre-reform period. For instance, it could have been that the pre-reform economy features a supremely uneven income distribution, and the reallocation that follows actually generates a more even distribution of income, on average.[151]

[151] Imagine the most extreme version of a feudal society. Inequality is as high as it could be as virtually everything, even the household's labor supply, belongs to the royal emperor. As a consequence, economic reforms that get rid of the special privileges of the ruling class are bound to reduce inequality.

2.6 Capital Outflows

Finally, I compute the capital flows of the economy along the transition path. To do so, I need to first solve for the transitional consumption and asset accumulation dynamics in the high-growth regime. This is simple, however, since the problem of the household in the high growth regime always looks the same (up to some linear scaling factor), no matter if a household enters the urban economy at time zero or a thousand years in. Intuitively, households always go through the same dynamics, albeit at different starting wages w_t^u. All choice variables are then simply scaled by income but otherwise unchanged. Using the Euler equation in 2.25 and the budget constraint, I employ a simple shooting algorithm to solve the household problem.[152] Figure B1 in the appendix plots the phase diagram. There I also prove uniqueness of the optimal path, and I show that the solution for cohorts entering the city at different points in time, up to a level shift, is identical. It suffices to solve for the consumption-to-income ratio of one single household. This ratio is always the same, no matter when the household enters the high-growth regime. Obtaining the actual solution then amounts to simply shifting up savings and consumption choices by the income level A_t^u for later cohorts.

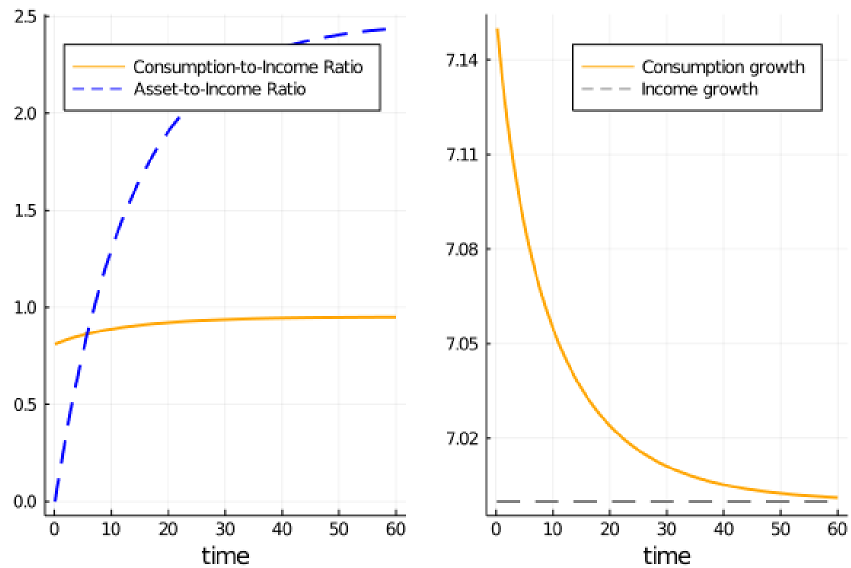
Figure 23 shows the optimal asset-to-income and consumption-to-income ratio. Unsurprisingly, households in the high-growth regime behave like buffer-stock savers (Carroll 1997). Time 0 here stands in for the time the household entered the urban sector. In other words the time line can be read as $t - t_m$. What I am simulating here is convergence towards a "Pseudo-steady-state". "Pseudo" because households are pulled out of this path by the Poisson process. When computing aggregate savings, then, I use the consumption function of figure 23 together with knowledge about how much time households spent in the high-growth regime, to get the right aggregate asset position. Put differently, the dynamics are valid for a very lucky household that happens to stay in the high-growth regime for a very long time. The pseudo-steady-state asset-to-income ratio is 2.47, and the consumption-to-income ratio is 0.95. The dynamics reflect the precautionary motive – fast consumption growth and a quick build up of "bufferstock" assets.

Going back to the empirical exercise in section 2.3, the asset-to-income ratio is informative about the precautionary motive, at least through the lens of the model. A larger variance of the type draw leads to a larger asset-to-income ratio. Persistent urban-rural differences

[152] This problem is almost identical to the neoclassical model in continuous time. In fact, the solution is slightly simpler because the rate of interest is exogenously fixed in this small open economy model, the rest is the same.

in the asset-to-income ratio, then, are indicative of greater human capital risk in modern production. The parameterization already reveals that the growth miracle is going to be accompanied by capital outflows, since households accumulate assets in the high growth regime. This is by no means guaranteed, and for more "even" growth miracles this would not be the case.

Figure 23: Household's Precautionary Savings



Transitional dynamics in high-growth regime using parameters from table 2.5. Time here is measured as time passed since the household switched sectors.

In a final step, I use the distribution of income together with the solution to the optimal consumption path in the high-growth regime to back out what the aggregate saving rate of our model economy would be. Recall that this aggregate saving rate is comparable to the current account since there is no capital in the model. To do so, define the mapping $Z : R^+ \to R^+$ that takes as input household income in the high-growth regime $y\left(t - t^i_m | T^i > t\right)$ and gives as output the asset-to-income ratio plotted in figure 23. Note that there is a one-to-one mapping between time spent in the high growth regime and income growth, and we can regard Z as a policy function where income is the state variable. This is of course only valid in the high growth regime. But for households in the low growth regime we know that assets grow at the balanced growth rate of 2%. I approximate Z by using a higher order polynomial where I suppose that after 50 years in the high growth regime the household has reached their long-run asset-to-income ratio. Aggregate asset holdings in the economy A^b_t can then be computed using the following accounting identity:

$$
\begin{aligned}
A_t^b &= \int a_i di \\
&= \int y_t^i \frac{a_t^i}{y_t^i} di \\
&= \int_{i \in M_{t,0}} y\left(t - t_m^i\right) Z\left(y\left(t^i - t_m^i\right)\right) di + \int_{i \in M_{t,1}} \exp\left(g^*[t - T^i]\right) y\left(T^i - t_m^i\right) Z\left(y\left(T^i - t_m^i\right)\right) di \\
&= A_0^u \exp\left(g^* t\right) \int_{i \in M_{t,0}} y_0\left(t - t_m^i\right) Z\left(y_0\left(t^i - t_m^i\right)\right) di \\
&+ A_0^u \exp\left(g^* t\right) \int_{i \in M_{t,1}} y_0\left(T^i - t_m^i\right) Z\left(y_0\left(T^i - t_m^i\right)\right) di
\end{aligned}
$$

This derivation uses the fact that, for households on the balanced growth path, assets grow at a rate of 2%. The asset position is therefore fully pinned down by the asset-to-income ratio last observed while in the high-growth regime, times income purged of the type draw. The type draw does not change the asset position – it's a permanent income shock that pushes up or down lifetime consumption but it does not induce additional savings. Using a change of variable we can now compute aggregate asset holdings in the economy using the income distribution that I have derived in the appendix.

$$
\frac{A_t^b}{A_0^u \exp\left(g^* t\right)} = M_{t,0} \int_1^{\exp((g-g^*)t)} y_0 Z\left(y_0\right) dF_0\left(y_0\right) + M_{t,1} \int_1^{\exp((g-g^*)t)} y_0 Z\left(y_0\right) dF_1\left(y_0\right) \quad (2.31)
$$

The conditional densities for households in the low-growth regime reads

$$
f_1\left(k\right) = \frac{1}{M_{1,t}} \frac{\lambda}{\left(g - g^*\right)} k^{-\frac{\lambda}{(g-g^*)} - 1} \left\{1 - M_{1,0} - L_t^r \left[k^{\frac{g^*}{(g-g^*)(1-\alpha)}}\right]\right\} \quad (2.32)
$$

Note that f_1 is technically not the distribution of income since it ignores the random type draw φ. It really is the distribution of income that accumulates due to convergence growth. I provide the result here because the expression is quite intuitive: in the long run the distribution converges to a Pareto distribution. This is no surprise since I introduced heterogeneous growth rates using a Poisson process. For non-zero agricultural shares L_t^r, there is a negative drag on the expected value. This reflects the fact that selection improves over time: At very early periods everyone in the pool $M_{t,1}$ only experienced small amounts of convergence growth. Over time, there is more potential for convergence and the expression converges to a Pareto distribution with a correction term $1 - M_{0,1}$ since I assumed that a fraction of households at time zero in the city do not participate in convergence growth. If one computes the expectation then, the only thing left to do is to account for the mass point at 1 with probability $M_{0,1}$.

The conditional density for households in the high growth regime reads

$$f_0\left(k\right) = \frac{1}{M_{0,t}} \frac{g^*}{\left(g - g^*\right)\left(1 - \alpha\right)} L_t k^{-\left(\frac{\lambda}{g-g^*} - \frac{g^*}{(1-\alpha)(g-g^*)}\right)-1} \tag{2.33}$$

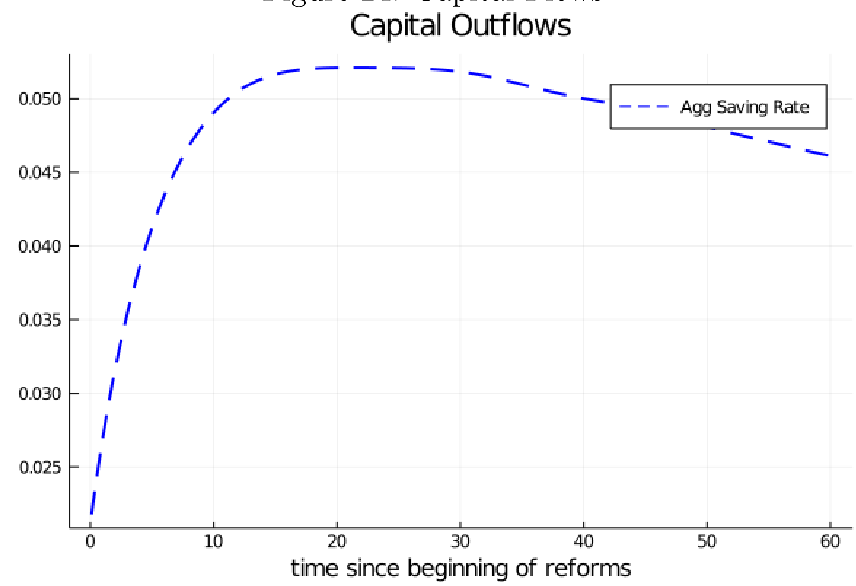
where the probability mass at $y = \exp\left((g - g^*)t\right)$ is equal to $\frac{M_{0,0}\exp(-\lambda t)}{M_{t,0}}$. That is to say, there is a positive mass of agents who start growing fast at time zero, and this mass point shrinks exponentially over time.

Putting the pieces together we get a trajectory for the aggregate saving rate displayed in figure 24. A success of the model is that it can replicate the hump-shaped saving rate that is characteristic of growth miracles. The magnitude of the current account flows is also broadly consistent with the level of capital outflows observed in China or Taiwan. The timing is off as the saving rate shoots up too fast. Note that I did not include a force that pushes the current account toward balance, as is usually done in small open economy models. Accordingly, the aggregate saving rate inherits the unit-root of the household consumption problem. Two hundred years in, the saving rate stabilizes around 3.5%.

In order to understand why the model delivers a hump-shaped saving rate it is helpful to look at the movements from households out of agriculture, and from the high growth to the low growth regime. Figure 25 shows the declining share of agriculture. The hump-shaped saving rate can only emerge as a compositional effect. That is, there needs to be an increasing share of precautionary savers relative to total output for some time. This is precisely what happens as figure 25 shows.The share of agents in the fast-growth regime is hump-shaped, and the aggregate saving rate can inherit those dynamics. For that to be the case we need the "right" values for g^*, α, and λ as these govern inflow and outflow into $M_{t,0}$ as well as $M_{0,1}$.

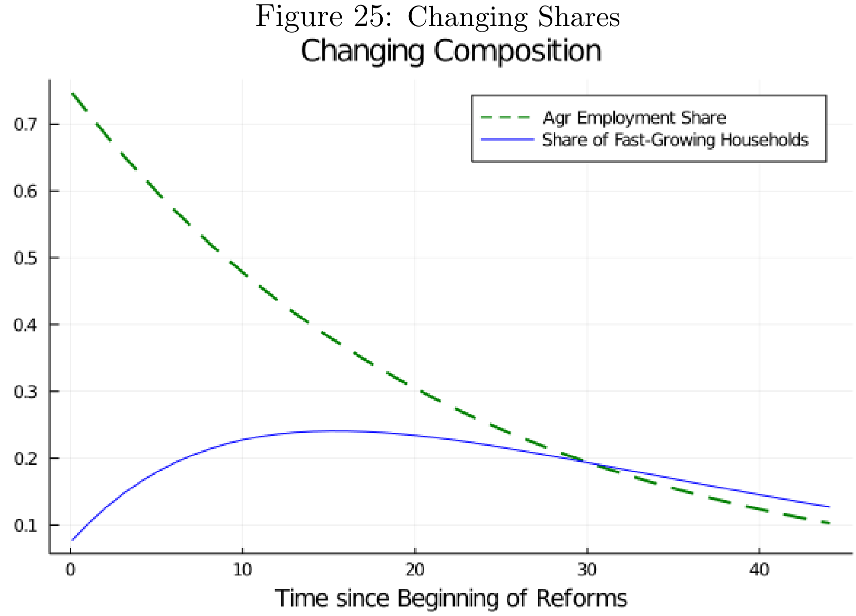
Figure 24 shows that this model economy can solve the Lucas puzzle for the right parameter values. The heterogeneity in the growth process and urban-rural differences in human capital risk are key deviations from the representative agent neoclassical model that make this feasible.

Figure 24: Capital Flows



Aggregate saving rate over time. Since there is no capital in the model this coincides with the Current Account in the small open economy.

Figure 25: Changing Shares



2.7 A final look at the household data: Evidence from Hukou-Switchers

After developing the theory, there is a non-trivial prediction that relates to the bufferstock savings behavior of households displayed in figure 23. Household safe asset growth is highest for fast-growing switchers that entered the urban sector recently. One can see this by noting that the consumption-to-income ratio is strictly increasing, and the saving rate is inversely related to this statistic. The Chinese data offer an opportunity to test this prediction to lend additional credibility to the model.

Specifically, I focus on Chinese households that were able to switch their Hukou status from rural to urban. The Hukou systems in China is a household registration system that assigns individuals into agricultural and non-agricultural households, based on their mother's Hukou at birth. Non-agricultural Hukous offer better public services and opportunity but it is very difficult for households to change their Hukou, although the system has been influx since the 1990s.[153] The model focuses on infinitely-lived households that that enter a life of human capital intensive production. While stylized, this is most consistent with households that were born with a rural household registration and have been able to obtain an urban one throughout their life. Note that these households are very different from temporary migrants who tend to take up low-skill labor intensive work and return to their rural homes eventually.

I compare Chinese households where the household head was born with a rural Hukou but has an urban Hukou in 2012. The appendix provides a set of descriptive statistics for each group, *agr_agr*, *agr_urban*, *urban_urban*. Income per capita is about $18k$ vs $23k$ Yuan and household heads have, on average, 9 and 11 years of schooling for switchers and urban Hukou holders, respectively. The key takeaway is that for any measure of development, say income per capita or years of schooling, the switchers (*agr_urban*) fall between the rural and urban Hukou holders.[154] Clearly, switchers are selected, but they are selected in a way that we can make some sense of. Through the lens of a model with human capital risk in urban production, we expect switchers to have lower financial asset-to-income ratios. On the other hand, we would expect them to display faster asset growth, precisely because they are below their long-run desired bufferstock savings position, which leads to fast accumulation. Both

[153] Overall, households that were able to change their registration status are positively selected on educational achievement, business achievement, or successful military or political careers. I provide additional information on the Hukou system in section B.9 in the appendix.

[154] While the Hukou status used to be tightly correlated with overall urban-rural status, fast urbanization and a number of reforms of the Hukou system have lowered the correlation between urban-rural status and urban-rural Hukou.

predictions are born out by the data. In the appendix the reader can verify that the median financial-assets-to-income ratios are systematically higher for households that always held an urban Hukou.[155]

In table 7 I report mean differences between households that always held an urban Hukou and switchers in terms of the growth rate of financial assets.

Table 7: Linear Regression for CFPS 2012 – 2016

	(1)	(2)	(3)	(4)	(5)
	g_fin_asset	g_fin_asset	g_fin_asset	g_fin_asset	g_fin_asset
hukou_switcher	0.0504*	0.0449	0.0465	0.0478	0.0393
	(0.0296)	(0.0297)	(0.0307)	(0.0307)	(0.0324)
_cons	0.271***	0.242***	0.565**	0.544**	0.515**
	(0.0189)	(0.0212)	(0.235)	(0.237)	(0.250)
income growth	No	Yes	Yes	Yes	Yes
age & demographics	No	No	Yes	Yes	Yes
education	No	No	No	Yes	Yes
province fe	No	No	No	No	Yes
N	1115	1110	1110	1110	1110

Note: The dependent variable is growth in nominal financial household wealth. *, **,*** denote statistical significance at 1, 5, and 10 percent level based on heteroscedasticity-robust standard errors. Rural households as well as the largest 1% of asset growth rates are dropped.

I run a simple OLS regression based on equation 2.1 but now with the growth of financial assets as the dependent variable.[156] I report the more conservative estimates here based on the geometric growth rates, which turns out to be, on average, 5 percentage points higher for switchers. If one uses the arc-percentage growth rate,[157] the differences becomes even larger. Given the small sample size, and potential measurement error, I interpret these results as supporting the main argument of the paper. Urban-rural differences matter for aggregate demand for safe assets, and households that join the urban economy have a strong incentive to build up buffer-stock savings.

[155]While the results hold qualitatively, the differences stop being statistically significantly different at the 1 % level after I start controlling for more variables. One issue here is the much smaller sample size of around 1500 households which makes detecting differences harder compared to before.

[156]This relates to the fact that growth rates are more stable than asset-to-income ratios or saving rates. It is important to note, though, that I drop the largest 1 % of outliers both for total wealth as well as financial wealth to improve the precision of my estimates from notoriously noisy household survey.

[157]This might be a sensible thing to do as some households hold zero financial assets in the base period which means that they are dropped in the baseline regression.

2.8 Conclusion

I have argued that the transition of households out of traditional agricultural production during episodes of fast catch-up growth is important to understand capital outflows in miracle economies. Empirically, rural (agricultural) households hold significantly less safe assets compared to urban households, conditional on their income and other observables. Taken together with the observation that households move out of traditional agricultural production very fast suggests that the interplay of urban-rural differences and structural change play an important role for the puzzling capital outflows of miracle economies.

I rationalize this finding in a simple model that highlights how structural change and human capital risk can give rise to strong demand for safe assets for urban households, ultimately leading to capital outflows. The main assumption underlying the model is that ex post inequality represent ex ante human capital risk in urban production. The model allows for an analytical characterization of the trade-off between consumption smoothing on the one hand, and the precautionary motive on the other. It endogenously generates structural change out of agriculture, and features a growth miracle that is multiple times larger than what is usually considered in the literature, roughly equal to the per capita growth rates of Taiwan from 1968 to 2000. Households face massive human capital risk as the economy ushers into a market-based system and workers move out of traditional agricultural production. Combined with uneven catch-up growth, this can generate a precautionary savings motive that is powerful enough to dominate the permanent income hypothesis – in spite of miraculous per capita growth.

A representative agent model would not be able to account for a growth miracle of that size without additional financial frictions because the consumption smoothing force is so dominant. This does not necessarily happen in the model at hand because growth itself is uneven and risky. This twist is central to quantitatively accounting for the capital flow puzzle and hopefully will be useful to other researchers as well.

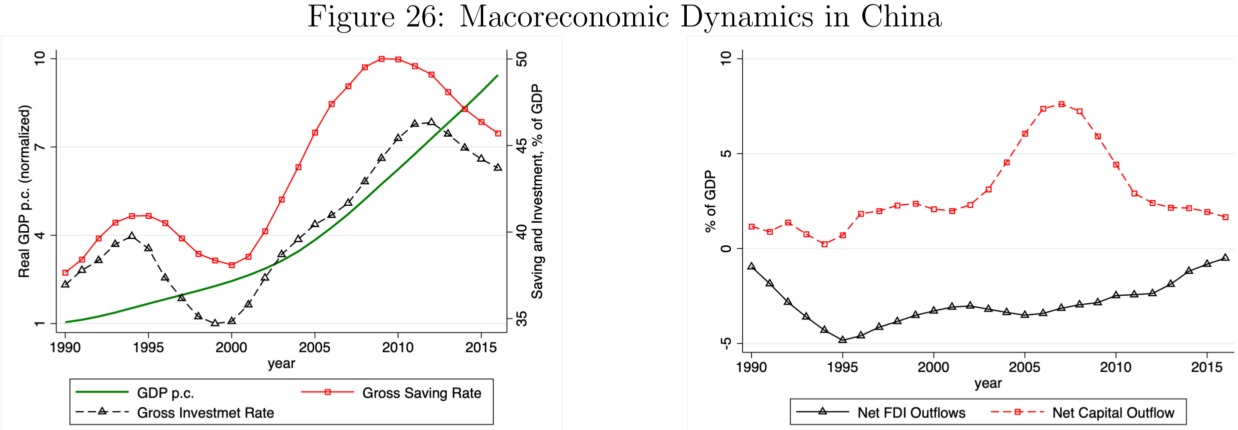
The framework is purposefully stylized to shed light on the main forces at play. The next step in this research agenda is to document income processes in fast-growing economies more carefully, while paying attention to urban vs rural (agricultural vs non-agricultural) differences. A carefully measured income process, then, lends itself to a more quantitative approach.

Chapter III. Uneven Miracle Growth – Modeling Income, Consumption, and Savings in Fast-Growing Emerging Markets

3.1 Introduction

Episodes of exceptionally fast national development in East Asia, and more recently China, are characterized by fast technological change, and high rates of capital accumulation. While aggregate investment is high, aggregate savings are even higher, leading to capital outflows and persistent current account surpluses. Demand for safe assets plays an outsized role in the current account surplus, while capital in the form of FDI flows into the emerging market. Figure 26 displays these patterns for the case of China. . An tenfold increase in real per capita GDP is what more than two decades of annual per capita growth of ten percent add up to. The characteristic hump shape in savings and investment rate emerges, where the gap between savings and investment constitutes the current account flow. While the current account is positive, the component of the current account that relates to foreign direct investment exhibits the opposite pattern with persistent net inflows.

Figure 26: Macroeconomic Dynamics in China



The data are taken from the World Development Indicators (June 2023). To smooth out short-run dynamics, I plot 5 year moving averages.

In this paper, I model these aggregate dynamics with a focus on the distribution of catch-up growth across households during an episode of very fast transitional growth growth.

Focusing on household income dynamics, and their implication for risk, inequality, and ultimately consumption-saving choices, connects the paper to a vast literature on precautionary savings in heterogenous agent models of the Ayagari-Bewly-Hugget-Imrohoroglu type. Yet, incomplete market models have been rather unsuccessful in squaring aggregate consumption-saving dynamics in miracle economies with standard economic theory.[158] The reason is that measured income risk is to small relative to fast aggregate catch-up growth, which induces strong consumption smoothing pressure for forward looking households. Given standard elasticities of inter-temporal substitution, this leads to counter-factually low domestic saving rates and capital inflows. That is to say, saving rates in the left panel of figure 26 should tank, and the current account should be negative.

I propose two key departures from the benchmark incomplete markets model and study theoretically and quantitatively how these additional features shape the evolution of consumption, savings, and wealth in fast-growing emerging markets observed. First, I assume income growth is unevenly distributed within cohorts across households, with a small share of extraordinarily fast-growing types as the growth miracle unfolds. At the same time, the rest of the population grows at a substantially lower rate. This growth rate heterogeneity is essential to match income and consumption dynamics in developed economies (Guvenen 2007; Huggett, Ventura, and Yaron 2011) as well as the evolution of top income inequality (Gabaix et al. 2016). I will show that growth rate heterogeneity is even more important in the context of fast growing emerging markets. Specifically, growth rate heterogeneity modeled such that household income grows very fast for some random time – a so-called Luttmer rocket (Luttmer 2011; Jones and Kim 2018) – reduces consumption smoothing pressure by an order of magnitude relative to a representative agent benchmark with the same rates of aggregate growth. The reason is that households' risk-aversion together with risky growth makes households heavily discount stellar outcomes, thereby lowering consumption smoothing pressure. At the same time, a thick right tail emerges as a combination of the rate of miracle growth and the average time spent in the fast growth regime, which fuels aggregate growth. I have made this argument before in a more stylized setting that also included urban-rural structural change, see Trouvain (2021). Here, I show in a quantitative incomplete market model that growth rate heterogeneity is essential to understand savings behavior in fast-growing economies.

Second, I introduce intergenerational human capital risk on top of the standard persistent income shocks, where the term human capital is used tantamount to earnings potential or

[158] An exception is the unpublished paper of Carroll and Jeanne (2009).

ability. I model this tractably using a perpetual youth model of the Blanchard-Yaari type (Yaari 1965; Blanchard 1985), where dynastic households die randomly, and are replaced by their offspring. At replacement, the new generation draws their human capital type from a distribution that is imperfectly correlated with the parent's human capital type. Building on a recent idea flow literature (R. E. J. Lucas and Moll 2014; Perla and Tonetti 2014a) I propose a simple specification where the distribution as a mixture between a draw from the cross-sectional income distribution, and the parent's human capital type. I incorporate this intergenerational risk into the continuos-time heterogenous agent model of Achdou et al. (2022a). I state a more general version of intergenerational human capital risk which might be of interest to researchers focusing on stationary equilibria, but I focus on a special case in my main application to preserves the linearity of the differential operator. This reduces the computational burden of notoriously hard-to-study transition dynamics.

As long as earnings potential is imperfectly correlated across generations, growth rate heterogeneity and cross-section inequality interact to generate strong precautionary savings. Importantly, the precautionary savings motive here is much stronger than in the standard incomplete market model with only small but persistent human capital shocks. Growth rate heterogeneity induces a fast increase in inequality as argued in Gabaix et al. (2016), with a thick tail of top income earners. Given that there is a chance that a high-income parent is replaced by a low ability child, forward looking households will increase precautionary savings to project the success of the dynasty into the future in spite of statistical mean reversion in income due to imperfect transmission of human capital.

The model makes a sharp distinction between "even" and "uneven" growth, where the former can be though of as a rise in productivity that leaves the income distribution unchanged,[159] whereas uneven growth pushes up inequality quickly. Both forces show up in aggregate statistics as fast GDP per capita growth, but they will have very different implications for consumption-saving choices. Evenly distributed income growth leads to the standard consumption smoothing force, which, under mild regularity conditions, does not interact with the degree of idiosyncratic household income risk. The saving dynamics of the standard incomplete market model along the transition path are qualitatively no different form the benchmark neoclassical representative agent model. Uneven growth, on the other hand, can play a very different role: As some high growth households get richer and richer relative to ordinary household types, the gap between high and low income households increases. In combination with intergenerational human capital risk, where there type of the

[159] In the sense of a Gini measure of inequality, or the log variance of income.

next generation is to some extent random, a strong self-insurance motive arises that is quantitatively much more powerful than the standard insurance motive in standard incomplete market models.

This logic finds a direct empirical counterpart in the explosion of fortunes of a small share of Chinese households. These households generated extremely large incomes that compare to the ultra wealthy in advanced economies in less than two decades, while the vast majority is still substantially poorer than the median household in industrialized countries. It is unlikely that the children of these superstar households will be able to maintain this level of income, let alone replicate the same growth spurt. Ramping up savings to raise consumption of these future generations is thus a natural response of a dynastic household in the presence of inequality in combination with intergenerational human capital risk. I support this argument with hand-collected data documenting the growth of the wealthiest Chinese households, who come form diverse backgrounds and build their fortunes rapidly while China was liberalizing its economy. I will argue why this supports the type of growth rate heterogeneity that I build into the framework. The rate of growth of these superstar households, and their share in the population, which I argue is particularly high during the early stages of the convergence process, will be crucial to generate appropriate transitional dynamics, which will be studied in detail.

Empirically, I will demonstrate that both growth rate heterogeneity and intergenerational risk find strong empirical support in the data. I will use Chinese household level data to discipline these key parameters, as well as more standard source of income risk. I will also provide empirical evidence in favor of the importance of inequality and high-income status as most important drivers of aggregate household savings, as opposed to competing approaches focusing on life cycle motives or competitive sex motives. Additional empirical moments that are crucial for my model is the evolution of income and wealth inequality, building on the recent work in Gabaix et al. (2016).

The focus on China is mainly due to data availability, but the uneven nature of growth miracles applies to other East Asian growth miracles as well. Whenever the growth spurt allows a subset of the population to get rich quickly, these wealthy households will have a strong incentive to build up large asset positions as they anticipate mean reversion due to the presence of intergenerational human capital risk. Given a fixed intergenerational income elasticity, which is meant to capture all sorts of institutional details and household activity that aim to perpetuate status across generations, asset accumulation is a tool to transmit economic prowess to following generations.

I close the model in general equilibrium by allowing for two types of assets, a safe internationally traded bond, and a risky asset that represents physical capital. The contrast between positive, large and persistent capital outflows and net FDI inflows as seen in the right panel of figure 26 calls for a two-asset structure. In order to match this fact, I incorporate time varying aggregate disaster risk a la Rietz (1988), Barro (2009), and Gabaix (2012), which drives a wedge between the returns of risky physical capital, and safe assets. The special role of safe assets for global current account imbalances has been highlighted by Bernanke et al. (2005), Caballero, Farhi, and Gourinchas (2008) and Mendoza, Quadrini, and Rios-Rull (2009). An important difference between their work and my approach is that I allow for stronger catch-up growth, which in their setup would create a force against holding safe assets from the emerging market's point of view.

I thus arrive at a general-equilibrium heterogenous agent model to study the coexistence of fast growth, high aggregate savings, current account surpluses and high demand for safe assets, despite high returns to capital and net FDI inflows. The key contribution of the paper is conceptual, namely that the distribution of income growth is quantitatively important to understand macroeconomic dynamics during an episode of fast convergence growth. Growth rate heterogeneity and intergenerational human capital risk are the key ingredients to break the otherwise dominant consumption smoothing force. This resurrects the importance of precautionary household savings during a growth miracle, which eludes benchmark incomplete market models. Needless to say, enriching my framework with additional financial frictions on the would render the problem easier. I find it useful to draw out the implications of my model while maintaining the neoclassical structure on the supply side to obtain a clean assessment of the strength of the proposed channel.

Lastly, I make a methodological contribution that will be useful for researchers working with grid-based global solution methods in environments where the economy transitions from a low to a high productivity state. To make the model operational, I develop a novel normalization procedure where the effective grid from the household's point of view is moving in line with movements in TFP. To see why this is useful, note that In the benchmark neoclassical model capital and wages would scale one-for-one in labor-augmenting TFP, while the interest rate would be unchanged in high-productivity steady state. Under mild regularity conditions, this is also true for Aiyagari-style models with household heterogeneity. However, an approximate solution based on global solution methods is unlike to replicate these homoetheticity properties. The reason is that too many households will hit the right most boundary of the grid as they earn higher wages and are wealthier. The same grid that is appropriate for a low productivity economy may be very problematic for an almost identical economy where the only difference is that TFP increased substantially. One solution is to make the grid very large. Yet, this is rather unattractive for two reasons. First, a very large grid is computationally expensive. And second, it would still not replicate the homotheticity properties exactly unless the grid is not only extremely large but also extremely dense. In contrast, my method preserves the homotheticity properties independent of the number of grid points, given mild regularity conditions. This means that the quantitative implications are not contaminated by finite-grid issues when transitioning from one steady state to another.[160]

The version of this paper that forms part of my dissertation is restricted to studying special cases that admit closed form solutions. A serious quantitative evaluation is work in progress, and will complement the research agenda that I have set out in the introduction. The rest of the paper is structure as follows. Section 3.2 situates the paper within a large literature, section **??** offers motivating empirical evidence, section 3.3 introduces the formal model, section 3.4 describes the computational innovation, i.e. the moving grid, section 3.5 applies the model to special cases, and section **??** concludes.

3.2 Literature Review

A large literature has focused on the link between economic growth, aggregate savings, and optimal household consumption. One strand of the literature focuses on life-cycle models a la Modigliani (1986) and Modigliani and Cao (2004). Deaton and Paxson (1994) and Deaton and Paxson (2000) explicitly study the link between growth and household savings in life cycle models with mixed success. The reader should note that much of the analysis is based on steady-state comparisons with permanent changes in the long-run growth rate. In the Chinese context, Curtis, Lugauer, and Mark (2015) study the effect of demographic change on savings in a closed economy setting. KEJU uses a life cycle model to study differences in households saving rates between the USA and China and account for global current account imbalances. In their model, growth matters because of a compositional effect. The economy that has a lower autarky interest rate, i.e. China, becomes larger in terms of its share of world GDP, which pushes down the world interest rate through a compositional effect. A key assumption is that China is stuck on its low level of financial development, which causes the low real rate in autarky. In contrast, I am interested in transitional dynamics where the

[160]It is clearly still problematic to have a grid that is too small, but my point is that it would be equally to small in the initial and final steady state, since form the household's point of view the grid scales in TFP.

emerging market ends up looking like the advanced economy in the long run, but displays very different dynamics during the transition.

Another strand of the literature studies incomplete market models – following Bewley (1977), **imrohorouglu1989**, Huggett (1993), Aiyagari (1994) – in emerging markets. See Townsend (1994), Kaboski and Townsend (2011), M. Chamon, K. Liu, and Eswar Prasad (2013), or He et al. (2018) among many others. Much of this work is either cast in partial equilibrium, or abstracts away from temporary fast income growth to generate realistic savings behavior. Relative to this work, I focus on more layers of household heterogeneity, and explicitly consider transition dynamics.

Household human capital dynamics and intergenerational persistent appears in the seminal work of G. S. Becker and Tomes (1986).[161] Intergenerational persistence features in De Nardi (2004), although bequests are modeled in a reduced form way through non-homothetic preferences. My approach also delivers strong bequest incentives, but they are endogenous to the degree of inequality in society, intergenerational persistence, and income state of the parent generation. Intergenerational human capital risk also appears in Laitner (1992), RIUSRULL, and Straub (2019) although their focus is quite different, and my specification is particularly tractable. The key point that I aim to make is that inequality, growth rate heterogeneity, and intergenerational risk interact to create precautionary savings that reflect an intergenerational motive. I abstract away from strategic interactions between parent and offspring generation, which is the focus in Boar (2020).

Seminal works on growth rate heterogeneity are Guvenen (2007), Huggett, Ventura, and Yaron (2011), as well as Gabaix et al. (2016). I propose a relatively simple model where homogenous Poisson processes create a speedy evolution of inequality (Gabaix et al. 2016), as well as a thick-tailed income distribution. In the spirit of the Aiyagari (1994) all risk is associated with household human capital. An alternative approach could involve capital return heterogeneity and a more explicit role for entrepreneurs, see the discussion in Benhabib, Bisin, and S. Zhu (2015) and Benhabib and Bisin (2018). It should be noted, however, that at least in the US high returns of private businesses seem in most part to be due to high levels of human capital, see Smith et al. (2019b), so the specification is perhaps not too outlandish, even if some aspect of the returns to human capital are mismeasured as returns to physical capital.

The perhaps most successful literature combines financial frictions on the supply side

[161] I abstract away from matters of human capital investment in the presence of borrowing constraints as in Loury (1981) and Galor and Zeira (1993).

with misallocations across firms. In contrast to Z. Song, Storesletten, and Zilibotti (2011), the focus on my paper rests on household saving dynamics, which usually is the largest component of aggregate savings in emerging markets. Their reliance on an overlapping generation model reduces the consumption smoothing force by construction, as old agents are not able to participate in future wage growth. I focus on an infinite horizon model which will capture this force, but I simultaneously allow for more realistic heterogeneity in income growth. This turns out to be empirically relevant during the growth takeoff, and quantitatively important through the lens of my model. The work of Buera and Shin (2017) is perhaps closest to my setting as they employ a quantitative heterogenous agent model with an entrepreneurial sector. My approach is complementary to their work as I offer a sharp distinction between even and uneven growth, and document under what conditions uneven growth and inequality lead to precautionary savings. In addition, my model features aggregate disaster risk and two types of assets, which allows me to speak to divergent trends in capital flows relating to safe assets and foreign direct investment, which are an important and understudied feature of the aggregate dynamics.

3.3 A Model of Uneven Catch-up Growth

I set up a simple heterogenous agent continuous time model. I set up the general economic environment here. After that, I will impose special restrictions to build intuition, before I turn to the quantitative implications of the full model.

3.3.1 Economic Environment

Households. There is a measure L of infinitesimal dynastic households indexed by i, which I normalize to unity, i.e. $\int_i di = N := 1$. Household supply their labor inelastically and make optimal consumption and portfolio choices so as to maximize utility given CRRA preferences, subject to a flow budget constraint and borrowing limits:

$$\max_{\{c_{i,s},a_{i,s}\}_{t\geq 0}} \int_0^\infty e^{-\rho s} \frac{c_{i,t}^{1-\gamma}}{1-\gamma} ds \tag{3.1}$$

$$\begin{aligned}\text{subject to: } & da_{i,t} + db_{i,t} = \left(r_{B,t} b_{i,t} + r_{K,t} a_{i,t} + w_t h_{i,t} - c_{i,t}\right) dt + g_a\left(a_{i,t}\right) dN_t^a \\ & b_{i,t} \geq 0 \\ & a_{i,t} \geq 0 \end{aligned}\tag{3.2}$$

where $b_{i,t}$ denotes a safe bond, $a_{i,t}$ denotes a risky asset, w_t is the wage rate, $h_{i,t}$ is household-specific human capital, $c_{i,t}$ consumption, and $r_{B,t}$ and $r_{K,t} + \mathbb{E}_t \left[\frac{g_a(a_{i,t}) dN_t^a}{a_{i,t}} \right]$ safe and expected risky assets returns, respectively.

Idiosyncratic Human Capital Risk. Households are heterogeneous in terms of their human capital, which evolves according to the jump-diffusion process

$$dh_{i,t} = \mu_{i,t} h_{i,t} dt + \sigma_{i,t} h_{i,t} dZ_{i,t} + \omega_{i,t} dN_{i,t}^{IGR}, \tag{3.3}$$

where $Z_{i,t}$ is a Wiener process. I assume there are two types of human capital growth regimes $j \in \{\mathrm{L}, \mathrm{H}\}$, low and high, with drift $\mu_t^{\mathrm{H}} \geq \mu_t^{\mathrm{L}}$ and variance $\sigma_t^{\mathrm{H}} \geq \sigma_t^{\mathrm{L}}$, and individuals draw their type at birth. The probability of drawing a high type equals χ_t^{H}. High-growth individuals switch into low-growth types with intensity δ_t^{L}, and the low-growth is an absorbing state within each generation. The time subscripts are essential as these parameters will change along the transition path. $N_{i,t}^{IGR}$[162] is a jump process with intensity $\lambda_{i,t}$. If a jump occurs, the evolution of human capital depends on the realization of $\omega_{i,t}$ drawn form a household-specific distribution, giving rise to intergenerational human capital risk explained below.

Intergenerational Human Capital Risk. Households are exposed to intergenerational human capital shocks.[163] Consistent with a large literature documenting income growth patterns over the life cycle, new generations start, on average, with a lower level of human capital than the parent generation. This gives rise to "intergenerational precautionary savings", where in particular high-income households build up savings in anticipation of the next generation's expected mean-reversion.[164] This insight appears in Dynan, Skinner, and Zeldes (2004), which introduces an intergenerational risk component. Relative to their model, I propose a more general stochastic process that explicitly allows for a varying degree of intergenerational persistence, and takes into account the entire cross-sectional income distribution in the population. This is important, as rising inequality will interact with the bequest motive in my framework.[165]

[162] IGR stands for intergenerational risk, explained below.

[163] In Trouvain (2023) I show how intergenerational human capital risk constitutes a crucial difference between urban and rural households, and in turn shapes differences in wealth accumulation between these groups.

[164] High saving rates of high-income households have recently been modeled using non-homothetic preferences in life-cycle models (De Nardi 2004; Straub 2019). I will show to what extent my model with standard preferences generates non-homotheticity.

[165] In Dynan, Skinner, and Zeldes (2004) households assume that their offspring will earn exactly the average income in the economy. Given reasonable measures of household risk-version, the mean outcome

Specifically, assume that with intensity λ_t, a new generation replaces the current household ($N^{IGR}_{i,t} = 1$).[166] As shown in equation (3.3), the human capital type of the offspring depends on the realization of the random variable $\omega_{i,t}$ with CDF $\Omega\left(h, j; i, t\right)$. This random variable should be correlated with parents' human capital $h_{i,t}$, as in the classic study of G. S. Becker and Tomes (1986), but not perfectly so. I model this tractably by combining a Bernoulli random variable $X_{i,t} \in \{0, 1\}$ with mean β_t with random sampling from some density ψ. With probability β (when $X_{i,t} = 1$) the next generation gets the exact same human capital type as the current generation, generating intergenerational persistence. In particular β has a structural interpretation as intergenerational elasticity, i.e. the slope coefficient resulting from a regression of child log income on parent log income.[167] Otherwise, the household randomly draws from the distribution $\Psi\left(y, j\right)$, so

$$\left(h_{i,t+\Delta}, j_{i,t+\Delta} | N_{i,t} = 1\right) = \begin{cases} \left(h_{i,t}, j_{i,t}\right) & \text{if } X_{i,t} = 1 \\ \left(u_{i,t}, k_{i,t}\right) & \text{else: random draw from distribution } \Psi\left(h, j\right). \end{cases} \tag{3.4}$$

Consistent with a large literature documenting income growth patterns over the life cycle, one would expect that the next generation starts, on average, with a lower level of human capital than the parent generation. Conceptually, offspring's initial income should be a draw from a distorted income distribution of the overall population. Denote with $F\left(h|j\right)$ the condition distribution of human capital given growth type j at time t, and assume that the distribution $\Psi\left(h|j, \kappa\right)$ is a tilted version of this income distribution,

$$\Psi\left(h|j, \kappa\right) = \left[F\left(h|j\right)\right]^{\kappa}, \kappa \in \left[0, 1\right], \tag{3.5}$$

in the sense that for $\kappa < 1$, mass is shifted to lower human capital outcomes within each growth regime. Consequently, children's income at birth is, on average, lower than parents' pervious income. This setting is inspired by a recent idea flow literature (R. E. J. Lucas

of a lottery is very different from its expected utility, a basic insight in expected utility theory that also matters quantitatively in my setting by raising the incentives for high-income households to save for the next generation. Boar (2020) provides empirical evidence in favor of the generational interdependence of household consumption choices. Intergenerational risk also appears in the works of Loury (1981) and Laitner (1992).

[166] This structure implies that household age is not a state variable, a useful simplification. I will address life-cycle issues later.

[167] See appendix See Blanden, Doepke, and Stuhler (2022) for an overview of the recent literature on intergenerational inequality.

and Moll 2014; Perla and Tonetti 2014a; Benhabib, Perla, and Tonetti 2021b) where agents sample from the population distribution. Consistent with a large literature on human capital accumulation, setting $\kappa < 1$ implies that the offspring generation starts, on average, with a lower level of human capital than the parent generation. Moreover, a newborn gets to sample with probability $\chi \in (0, 1)$ from the high-growth distribution and enters as a high growth type, i.e. χ is the probability of being born into and sampling from the high growth regime.

To keep the distributional dynamics tractable, I consider a special case of equation (3.5) where households draw with Bernoulli probability $\hat{\delta}$ from the distribution $\Psi(h|j, \kappa = 0)$, in which case the offspring will have a lower human capital level than the parent generation. With probability $1 - \hat{\delta}$, the household draws from the distribution $\Psi(h|j; \kappa = 1)$, which represents a random draw from the conditional income distribution in which case human capital stays, on average, constant. It follows that expected income of a new cohort is a weighted average between worst outcome and population average

$$\mathbb{E}_{t+\Delta}\left[w_{t+\Delta}h_{t+\Delta}|N_{i,t}^{IGR} = 1\right] = w_{t+\Delta}\left[\hat{\delta}\underline{h}_t + \left(1 - \hat{\delta}\right)\mathbb{E}_i\left[h_{i,t}\right]\right].$$

Note that only these two special cases preserve the linearity of the Kolmogorov Forward equations (KFE), which I develop below, in the sense that the evolution of the income distribution can be summarized using a linear differential operator, see Gabaix et al. (2016) for details. This leads to substantial speed gains that come in handy when studying notoriously hard-to-solve transition dynamics in heterogenous agent models. I explain this aspect in more detail in the appendix C.2.2 but intuitively, the linear system involves inverting a large matrix which is relatively fast compared to a non-linear system that requires iterating on a large probability vector potentially thousands of times.

Technology. A final goods sector produces GDP Y, which serves as the numeraire, according to a Cobb-Douglas production function combining effective labor $\widetilde{Lh} := \int_i h_i di$ and physical capital $K := \int k_i di$. Since the size of the population is normalized to unity and I abstract away from population growth, average human capital on the household level $\tilde{h}$ is identical to total labor $\widetilde{Lh}$, so aggregate output equals

$$Y = K^{\alpha}\left(A\tilde{h}\right)^{1-\alpha}, \tag{3.6}$$

where A is total factor productivity (TFP). Final output can be turned one-for-one into physical capital K and depreciates at rate δ_K so that the production side of the economy follows the benchmark neoclassical model.

Aggregate Risk. Economic transitions are risky. Growth spurts are often followed by stagnation or even regress.[168] Concretely, suppose that a possible change in policy in a fast-growing emerging market undermines the convergence process. I model this in a simple way by assuming that with time-varying arrival rate ϕ_t^{agg} a change in policy leads to i) a full and immediate reversal to the pre-reform state characterized by i) productivity A_0, ii) a complete nationalization of the physical capital stock with returns evenly shared across households, iii) a loss of human capital back to h_0. This simplifying assumption means that households fall back to their initial level of income. This is an absorbing state. In the appendix I show that these extreme assumptions are very similar to a model where aggregate risk is higher but only a fraction of capital is nationalized and TFP does not fall all the way back to A_0.[169] Ultimately, the reader should think of this source of aggregate risk as a residual needed for the model to match capital return differentials between countries in an open economy equilibrium, which I will discuss in more depth later on. This setup builds on a recent literature on rare disaster risk, originally due to Rietz (1988) and reinvigorated by Barro (2009). I follow Gabaix (2012) in introducing time-varying disaster risk, see also Farhi and Gabaix (2016) for an application in international finance.

Equilibrium. I focus on a competitive equilibrium where final goods producers maximize profits leading to the standard demand for labor and capital,

$$w = (1-\alpha) A \left(\frac{K}{A\tilde{h}}\right)^{\alpha} \tag{3.7}$$

$$r_K = \alpha \left(\frac{K}{A\tilde{h}}\right)^{\alpha-1} - \delta_K, \tag{3.8}$$

households maximize (3.1) subject to constraints and the evolution of their human capital (3.3), and markets for labor, physical capital, bonds, and final output clear, and in particular

$$K = \int a_i di \tag{3.9}$$
$$0 = \int b_i di.$$

I focus on a closed economy for now, where the risk-free asset is in zero net supply, and

[168] A famous example that arguable falls into this category is Argentina. Acemoglu and Zilibotti (1997) provide a micro-foundation for fickle growth in emerging markets.

[169] You gotta do that, also the distributional thingy interacts a little with the capital thing.

the risky asset equals physical capital. Later, I will consider an integrated equilibrium with trade in risk-free assets and physical capital.

3.3.2 Economic Growth & Inequality

The focus of the project rests on the distribution of growth, which I split into two parts. One "even" component represents standard catch-up growth in productivity, which I describe first. An "uneven" component concerns the evolution of inequality, holding productivity fixed, which is the novel aspect of the model.

Even Growth. At time zero, the economy starts out at a low level of TFP compared to its long-run productivity. I assume that a set of unspecified reforms in the emerging market induce productivity growth, which starts a process of economic convergence whereby the technology gap $\frac{A^F}{A_t}$ declines over time.[170] I use a simple logistic growth model to parameterize this process

$$\dot{A} = \zeta A \left[1 - \frac{A}{\tau A_F} \right], \tag{3.10}$$

where catch-up growth is initially fast and peters out eventually as the economy approaches its long-run productivity level τA_F where $\tau < 1$ concerns the empirically relevant case in which countries don't travel all the way up to the technological frontier. This is a convenient modeling choice that admits a closed form solution

$$A_t = \frac{\tau A_F}{1 + Ce^{-\zeta t}},$$

where an initial condition pins down $C = \frac{\tau A_F}{A_0} - 1$.

I coin this type of growth "even" because it lifts up household labor income evenly, independent of whether households possess a high or low amount of human capital. Improvements in A raise the marginal product of labor, which leads to rising wages as can be inferred from (3.7). Household inequality due to differences in human capital, measured in terms of the log variance of household income, or any other scale invariant measure of inequality, does not respond as aggregate productivity A catches up with the rest of the world.

It is precisely this type of catch-up growth that leads to strong consumption smoothing pressure, which pushes up the interest rate in the closed economy, while leading to capital inflows in the open economy. Forward-looking households anticipate that they will earn

[170] See Buera and Shin (2013) for a model where the evolution of productivity is endogenous.

higher wages in the future. The concavity of the flow utility function then implies that households raise their level of spending instantly to smooth consumption over time. Since this mechanism in the paper is standard, any non-standard results must come from what I call "uneven" growth.

Uneven Growth & Idiosyncratic Risk. I assume that convergence growth itself is to some extent unevenly distributed across households. The intuition is that growth miracles offer a unique opportunity for households to get very wealthy in a short amount of time. Yet, only a relatively small share of households is able to seize this extraordinary opportunity. While few, these households can play a disproportionate role in aggregate dynamics as an evolving thick-tailed income distribution puts a large weight on these "superstar" households so that aggregate growth is substantially larger than income growth of the median household.

Moreover, as the growth miracle comes to an end, the opportunities for such extraordinary growth become rarer and the share of households growing at extraordinary rates recedes, thereby contributing to the decline in the aggregate growth rate. Since the system converges to a stationary one in the long-run, and since a stationary system exhibits mean-reversion by construction, the reader might anticipate that superstar households have an incentive to accumulate assets to smooth consumption over the long-run. This is the sense in which "uneven" growth and household heterogeneity can turn the relationship between temporary income growth and aggregate consumption upside down.

3.3.3 Solving the Model

HJB equation and Komolgorov Forward Equation With this additional structure, and after combining the two types of assets to net worth $W := b + a$ and $\pi := \frac{a}{W}$ defined as the share of wealth in risky assets [171], one can reformulate the household problem (3.1) using an HJB equation

[171] This reduces the state space, and leads to the classic portfolio choice problem of Merton (1969).

$$\rho v(W,h,j) - \dot{v} = \max_{c,\pi} \frac{c^{1-\gamma}}{1-\gamma} + v_W\left[(r+(R-r)\pi)W + wh - c\right]$$

$$+ \underbrace{v_y y \mu_t^j + v_{yy} y^2 \frac{\left(\sigma_t^j\right)^2}{2}}_{\text{drift-diffusion}} \tag{3.11}$$

$$+ \mathbf{1}_{\{j=\mathrm{H}\}} \delta_{\mathrm{L}} \underbrace{\left[v(W,h,\mathrm{L}) - v(W,h,\mathrm{H})\right]}_{\text{leave high growth rate regime}} \tag{3.12}$$

$$+ \underbrace{\lambda(1-\beta_t) \int \left[v(W,u,k) - v(W,h,j)\right] d\Omega(u,k)}_{\text{inter-generational human capital risk}}$$

$$+ \underbrace{\phi_t^{agg}\left[v(W(1-\pi)+k_0,h_0,j_0) - v(W,h,j)\right]}_{\text{aggregate risk}}, \tag{3.13}$$

where the first row concerns standard elements, including the flow utility of consumption, the partial derivative w.r.t. wealth v_W, the wealth accumulation constraint, a drift term, and a volatility term using Ito's Lemma. The second row includes intergenerational human capital risk, as well as changes in the growth type in case the household is in the fast-growth regime where $\mathbf{1}_{\{j=\mathrm{H}\}}$ is an indicator function. The last line captures aggregate risk in the following sense.

The unconditional income distribution, defined as $\int p(W,h,j)\,dW = p(h,j)$, evolves according to the Kolmogorov Forward Equation (KFE)

$$\frac{\partial p(h,\mathrm{H})}{\partial t} = -\frac{\partial}{\partial h}\left[\mu_t^{\mathrm{H}} h p(h,\mathrm{H})\right] + \frac{\partial^2}{\partial h \partial h}\left[\frac{\left(\sigma_t^{\mathrm{H}}\right)^2}{2} h^2 p(h,\mathrm{H})\right] - \lambda(1-\beta_t)\hat{\delta} p(h,\mathrm{H}) - \delta_{\mathrm{L}} p(h,\mathrm{H}) \tag{3.14}$$

$$\frac{\partial p(h,\mathrm{L})}{\partial t} = -\frac{\partial}{\partial h}\left[\mu_t^{\mathrm{L}} h p(h,\mathrm{L})\right] + \frac{\partial^2}{\partial h \partial h}\left[\frac{\left(\sigma_t^{\mathrm{L}}\right)^2}{2} h^2 p(h,\mathrm{L})\right] - \lambda(1-\beta_t)\hat{\delta} p(h,\mathrm{L}) + \delta_{\mathrm{L}} p(h,\mathrm{H}) \tag{3.15}$$

together with the boundary conditions

$$0=-\left[\mu_t^{\mathrm{L}}\underline{h}p\left(\underline{h},\mathrm{L}\right)\right]+\frac{\partial}{\partial h}\left[\frac{\left(\sigma_t^{\mathrm{L}}\right)^2}{2}\underline{h}^2p\left(\underline{h},\mathrm{L}\right)\right]+\lambda\left(1-\beta_t\right)\hat{\delta}\left(1-\chi\right) \tag{3.16}$$

$$0=-\left[\mu_t^{\mathrm{H}}\underline{h}p\left(\underline{h},\mathrm{H}\right)\right]+\frac{\partial}{\partial h}\left[\frac{\left(\sigma_t^{\mathrm{H}}\right)^2}{2}\underline{h}^2p\left(\underline{h},\mathrm{H}\right)\right]+\lambda\left(1-\beta_t\right)\hat{\delta}\left(\chi\right).$$

This set of equations constitutes a system of coupled partial differential equations, which I solve following the finite difference method in Achdou et al. (2022a). The boundary conditions are somewhat non-standard due to growth rate heterogeneity, and I derive them step-by-step in the appendix C.1. Moreover, the equivalence between the infinitesimal generator and the transpose of the operator of the KFE breaks down in my application, due to the intergenerational risk component, and I derive appropriate generators in the appendix.

Relative to this seminal paper, I offer four extensions, three conceptual and one computational, all of which are important for studying consumption and wealth dynamics in transition economies. First, I incorporate growth rate heterogeneity, which is essential to match income and wealth dynamics, building on Gabaix et al. (2016). Second, I introduce intergenerational risk, which is crucial in generating saving pressure for fast-growing high-income individuals. I will show how these two features interact in a way to overcome the standard consumption smoothing force that otherwise dominates along the transition path. Third, time-varying aggregate capital risk and internationally-traded safe assets will be useful to reflect the high demand for safe assets in spite of high returns to capital in fast-growing economies. It will allow me to match FDI inflows and "safe asset" capital outflows. Fourth, I offer a simple normalization procedure that allows me to mimic a time-varying wealth grid from the household's point of view, which I view as a simple and useful computational innovation explained next.

3.4 Normalized Asset Grid

Before delving into the normalization, let's first discuss the problem with the standard procedure. The standard procedure employs a global solution methods that requires specifying a fixed asset grid on which the finite difference scheme operates. Given that TFP is increasing substantially when studying fast growth episodes in emerging markets, this is not a good assumption. When TFP is increasing, and the asset grid is fixed, large computational errors are likely, especially in applications with fat-tailed income and asset distributions. The reason is that an increase in TFP shifts up the entire distribution of household savings, i.e households pick higher individual asset stocks since their real income went up permanently. Given the finite nature of the grid, too many households end up being constrained by the upper bound of the grid. A simple way to gauge the severity of the problem is to work with a fixed asset grid, and compare two steady states where in the second steady state TFP is twice as high as in the initial equilibrium. Given regularity conditions, wages and physical capital should be twice as high in the new steady state, while the interest rate remains unchanged, which are well-known homotheticity properties of standard incomplete market models.[172] I experimented with a fixed asset grid and found a doubling of TFP in my model, while qualitatively correct, was quantitatively inconsistent and violated the aforementioned homotheticity properties substantively.[173] One solution is to choose a very large and dense grid but this is computationally costly. I propose an alternative strategy.

Instead of fixing an asset grid in real terms $\{a_{0,}a_{1,...,}a_{max}\}$, I set up the grid space using TFP as a normalizing factor. Define the normalized asset level $\tilde{a} := \frac{a}{A_{TFP}}$. It is clear that the household problem can be solved using $\tilde{a}$ as state variable as long as TFP is an exogenous scalar, which is the case in my application. Abstracting away from growth, risk and multiple assets for expositional purposes, consider the stationary HJB equation using the normalized asset level

$$\rho v\left(\tilde{a}, A_{TFP}\right) = \max_{c} \frac{c^{1-\gamma}}{1-\gamma} + v_{\tilde{a}}\dot{\tilde{a}}, \tag{3.17}$$

where the law of motion for normalized assets follows from differentiating $\tilde{a}$ with respect to time

$$\dot{\tilde{a}} = \left(\left(r - g_{TFP}\right)\tilde{a} + \frac{y-c}{A_{TFP}}\right), \tag{3.18}$$

and $\frac{\dot{A}}{A} := g_{TFP}$ is the TFP growth rate. The differential operator $\mathcal{A}$ encodes the evolution of income and type space over time, using Ito's lemma and including intergenerational risk and aggregate capital risk. In the steady state TFP growth is zero, while during transition dynamics a changing exogenous TFP growth rates appears in the budget constraint (3.18).

[172] These homotheticity properties are well-known, see Aiyagari (1994) footnote 12, and work with standard CRRA utility and borrowing constraints that scale in TFP, for instance when they are proportional to the wage rate.

[173] For example, I found interest rates jump from 0.007 to -.005, almost a one percentage point difference.

The first order condition arising from (3.17),

$$\frac{v_{\tilde{a}}}{A_{TFP}} = c^{-\gamma},$$

is no different from the familiar formula $v_a = c^{-\gamma}$. Note that $v(a) = v(\tilde{a})$ since household consumption choices are independent of the normalization so that the present discounted value of expected utility v remains unchanged. Differentiating both sides with respect to a, and using $\frac{\partial \tilde{a}}{\partial a} = \frac{1}{A_{TFP}}$, it follows that $v_a = \frac{v_{\tilde{a}}}{A_{TFP}}$. It is straightforward to prove that the value function scales such that $v(\tilde{a}, A_{TFP} = 2) = 2^{1-\gamma} v(\tilde{a}, A_{TFP} = 1)$.[174]

Computationally, I apply the fixed point algorithm in Achdou et al. (2022a) to the normalized grid, i.e. the value function $v(\tilde{a})$ is a function of $\tilde{a}$. However, I compute consumption choices and boundary conditions using the normalizing factor A_{TFP}. For example, $c = \left(\frac{v_{\tilde{a}}}{A_{TFP}}\right)^{-\frac{1}{\gamma}}$, or for the boundary condition at the lower end of the state space I have $v_{\tilde{a}} \geq A_{TFP} \left(w\underline{h} + rA_{TFP}\underline{\tilde{a}}\right)^{-\gamma}$. Similarly, the stationary distribution is defined on $(\tilde{a}, h)$, which is a fixed grid, but I compute aggregate asset demand by adjusting for TFP,

$$K = A_{TFP} \int \tilde{a} \cdot g(\tilde{a}, h)\, dh d\tilde{a},$$

which works in and out of steady state. In summary, while the grid is fixed, from the household's point of view the effective grid is moving and scaling linearly in TFP. With this method, which as of now I have not found elsewhere in the literature, a moving grid ensures that the central homotheticity properties of incomplete market models hold automatically. This allows for accurate comparisons across equilibria with different long-run TFP levels, and reduces the computational burden as less grid points are needed to approximate the asset choice. This method should be more broadly useful to researcher using grid-based solution methods in heterogenous agent models with TFP dynamics.

[174] For this to be true, it needs to be the case that the borrowing constraint takes the form $a \geq -\kappa w$ where $\kappa \geq 0$ is a constant, i.e. the constraint is proportional to the wage rate. In that case, a scaled-up but otherwise unchanged consumption sequence is a solution to the household problem whenever TFP is scaled up, see Aiyagari (1994).

3.5 Special Cases

3.6 Hugget Economy and Risk-Free Rate

After outlining the general setup, it is useful to analyze some special cases of the model to provide intuition for the quantitative results. First, I consider a simplified version of the model, with the goal in mind to show how the distribution of growth matters in combination with intergenerational human capital risk counteracts the standard consumption-smoothing pressure in representative agent models. Suppose $\sigma^{\mathrm{H}} = \sigma^{\mathrm{L}} = 0$, i.e. there are no random income fluctuations within each generation.

Moreover, I focus on a Hugget economy with labor as only factor of production ($\alpha = 0$) and risk-free assets in zero net supply. This shuts down aggregate capital risk, so the application is about pricing a risk-free bond in autarky as the economy transitions from stagnation to growth. Aggregate disaster risk still matters in that with time-varying rate ϕ_t all agents are pulled back to their pre-reform income state denoted by y_0 and c_0 is the associated consumption level. Further, suppose that $\hat{\delta} = 0$ so the offspring's human capital distribution is drawn from the sample distribution. I make two final simplifying assumptions, $\delta_{\mathrm{L}} = 0$ and $\chi = 1$ iff $y_t' > \min\{y_t\}$ where y_t' is the draw of the offspring generation, and $\chi = 0$ otherwise. This means that there are only two types of households, high income-high growth types, and low income-low growth types. Being hit by an intergenerational shock means that the household draws from a bimodal distribution, and whenever the high type is drawn, fast growth follows. These assumptions lead to an analytically tractable no-trade equilibrium similar to Constantinides and Duffie (1996) and Heathcote, Storesletten, and Violante (2014b). I follow Krusell, Mukoyama, and Smith Jr (2011) in that unconstrained agents price the bond, which allows me to provide closed-form solutions of the risk-free rate for non-standard income processes. Derivations are in appendix C.3.

Using the HJB equation together with the envelope condition, I derive the standard Euler equation which applies to whichever household type is at an interior solution

$$\frac{\dot{c}}{c} = \frac{r-\rho}{\gamma} + \frac{\lambda(1-\beta)}{\gamma}\mathbb{E}_{c'}\left[\left(\frac{c'}{c}\right)^{-\gamma} - 1\right] + \frac{\phi_t}{\gamma}\left[\left(\frac{c_0}{c}\right)^{-\gamma} - 1\right]. \tag{3.19}$$

Note that while the first order condition in a continuos-time heterogenous agent model always binds, i.e. $v_a = u'(c)$, as argued in Achdou et al. (2022a), the envelope condition and the household Euler equation only applies to the unconstrained agent who is ultimately pricing the bond.

To derive the risk-free rate, I next specify the income process. Suppose at time zero, all households are identical and $\mu_{\mathrm{H}} = \mu_{\mathrm{L}} = 0, y_{\mathrm{H}} = y_{\mathrm{L}} = y_0 = 1$,i.e. the disaster outcome is equivalent to pulling households back to their time zero income level. I know partition time into an early $\left(t < \hat{t}\right)$ and a late $\left(t \geq \hat{t}\right)$ growth phase with $\hat{t}$ marking the point in time when changing from one regime to the other. The early growth regime is associated with the subscript 1 and the late growth regime with 2. In the early regime, $\mu_{\mathrm{H},1} > \mu_{\mathrm{L},1} = 0$. In the late regime, $\mu_{\mathrm{H},2} = \mu_{\mathrm{L},2} > 0$, and $\phi_2 = 0$ while $\phi_{t|t<\hat{t}}$ is a strictly decreasing function of time, which means disaster risk is higher at early stages of development.

In this setup with maximally tight borrowing constraint, high-growth households are initially borrowing constrained as their faster expected income growth means that they want to consume more than their low-growth counterparts. Using the fact that $c = y$ in this no-trade equilibrium, and taking account of the intergenerational risk, I can compute the risk-free rate in the regime 1 for the case with the low-growth households pricing the bond

$$r = \rho + \lambda \left(1 - \beta\right) \chi \left[1 - \left(\frac{y_{\mathrm{L}}}{y_{\mathrm{H}}}\right)^{\gamma}\right], \tag{3.20}$$

note that the disaster risk drops out because low-growth households experience no improvement in their income and consumption level relative to time zero. The interest rate is above the discount factor as low-growth households internalize that with probability $\lambda \left(1 - \beta\right) \chi$ they will transition into the high-growth high-income regime, which induces consumption smoothing pressure.

I denote with $\Gamma := \frac{y_{\mathrm{H}}}{y_{\mathrm{L}}}$ the income gap between high and low-income households, which is directly proportional to the degree of inequality in the economy, and increasing at rate μ_{H} in the early growth regime.[175] In the early stage of the growth takeoff, inequality rises and aggregate growth is fueled by growth form fast-growing high-income households. In growth regime 2 income inequality is stable and equals $\left(\mu_{\mathrm{H}}\hat{t}\right)^2 \chi \left(1 - \chi\right)$. Aggregate growth is now evenly distributed in the sense that each household type experiences the same growth rate. Given that $\hat{t}$ is large enough, the high growth type will eventually be become the unconstrained agent. Using (3.19) I can show that when the high-growth agent is unconstrained, the interest rate equals

[175]The log variance of income in growth regime 1 equals $\left(\mu_{\mathrm{H}} t\right)^2 \chi \left(1 - \chi\right)$ and is an increasing function of time.

$$r = \rho + \gamma\mu_{\mathrm{H}} - \underbrace{\lambda\left(1-\beta\right)\left(1-\chi\right)}_{\text{intg. risk}}\left(\underbrace{\Gamma^{\gamma}-1}_{\propto\text{inequality}}\right) - \underbrace{\phi}_{\text{agg risk}}\underbrace{\left(\left(\frac{c_0}{c}\right)^{-\gamma}-1\right)}_{\text{disaster depth}}. \tag{3.21}$$

A few remarks are in order. Note that the standard consumption smoothing force shows up in the term $\gamma\mu_{\mathrm{H}}$. In the absence of intergenerational risk and aggregate risk, fast growth households will always be the constrained households. The following proposition states this key result formally. In the case without intergenerational risk, and without aggregate risk, i.e. $\beta = 1$ and $\phi = 0$, respectively, fast-growing high-income household will always be borrowing-constrained with zero savings. However, for sufficiently high intergenerational risk β, this is no longer true. In particular, given high rates of inequality Γ, it is highly likely that the high growth households will eventually want to accumulate savings. To see this, compute t^* – the point in time at which both households are at an interior solution – as a function of exogenous parameter by setting (3.20) equal to (3.21)

$$\gamma\mu_{\mathrm{H}} \hat{=} \lambda\left(1-\beta\right)\left(\Gamma^{\gamma}-1\right)\left[1-\chi\left(1+\Gamma^{-\gamma}\right)\right] + \phi\left(\Gamma^{\gamma}-1\right), \tag{3.22}$$

where the existence of t^* depends on exogenous parameters. The standard consumption smoothing force appears on the left-hand side, pushing against positive saving rates for fast-growing households. The more novel aspect of equation (3.22) is that intergenerational risk, together with rising inequality, pushes the other way and exerts intergenerational precautionary savings pressure. The higher inequality in the population, the stronger the intergenerational human capital risk, which pushes up the precautionary motive. As pointed out in previous work Trouvain (2021), ex post inequality can appear as ex-ante risk when households face uncertainty as to where they will land on the income distribution.

In the same vein, aggregate risk leads to additional precautionary savings pressure for fast-growing high-income households, while it plays no role for low-income households. Since I have modeled aggregate risk as falling back to the pre-growth state, only high-income households, who have improved their income a lot, have something to lose.

Lastly, note that the left hand side is linearly increasing in the growth rate, while the right hand side is exponential increasing in the growth rate for a given t since $\Gamma^{\gamma} = e^{t\mu_{\mathrm{H}}\gamma}$. That means that especially for very high household growth rates, paradoxically, high-income high-growth households will try to build up savings early. This turns upside down the relationship between savings and growth, even within an infinite horizon model with forward-looking households. Aggregate growth reads

$$g_Y = \mu_{\mathrm{H}} \frac{\chi y_{\mathrm{H}}}{\chi y_{\mathrm{H}} + (1-\chi)\, y_{\mathrm{L}}}, \tag{3.23}$$

and is a simple weighted average between high and low growth rates, the latter being conveniently zero.

PLOT against R.

At this point, we have only focused on the consumption-smoothing problem, i.e. the Euler equation. However, another crucial aspect somewhat less obvious is how "rich" in terms of present discounted value of life-time income households feel. I argued that growth rate heterogeneity a la Luttmer (2011), used in Jones and Kim (2018) and Gabaix et al. (2016), is useful because it reduces substantially how "rich" households feel. The reason is that these growth spurts are over randomly, and the randomness interacts with the present discounted value just as it does for standard financial asset.

3.7 Luttmer Rocket and Consumption Smoothing

Another crucial feature is that uneven income growth in the form of random time $\frac{1}{\lambda}$ spend in a fast growth regime leads to fundamentally less consumption smoothing pressure compared to a scenario where all households converge using a deterministic trend. I will evaluate this claim quantitatively later on, but it is instructive to go through a simple case to understand the underlying mechanics. For this purpose, I simplify further by assuming that some household gets access to the growth rocket ones, and when they are pulled out of the fast growth regime, they are stuck with their current level of income forever, i.e. there is no income risk other than the random time spent in the fast-growth regime. Formally, I set $\beta = 0$ and λ simply governs the average time spent in the high growth regime.[176]

The HJB equation reads

$$(\rho + \lambda)\, v(y) = \frac{y^{1-\gamma}}{1-\gamma} + v_y \mu_{\mathrm{H}} y + \lambda v(y, \mathrm{L}) \tag{3.24}$$

where the value in the low-growth regime reads

[176] This a simplified version of the growth process in Trouvain (2021).

$$v\left(y,\mathrm{L}\right)=\frac{1}{\rho}\frac{y^{1-\gamma}}{1-\gamma}. \tag{3.25}$$

Combining (3.24) and (3.25) allows me to solve for the value function in the high-growth regime in closed form,[177]

$$v\left(y\right)=\frac{y^{1-\gamma}}{\gamma-1}\left(\frac{\gamma-1}{\rho+\lambda}\mu_{\mathrm{H}}-\frac{1}{\rho}\right),$$

which is negative for $\gamma>1$ and strictly increasing in the growth rate μ_{H}. Now I can ask, how much lower is the present discounted value compared to a scenario where all households get the average growth rate, i.e. markets are complete and risk-averse households insure one another against relatively short growth spurts. Given the following parameterization $\{\mu,\lambda,\gamma,\rho\}=\{5\%,7\%,2,4\%\}$, it turns out that the household would be willing to reduce its initial income level by roughly 40% if allowed to trade the aggregate path for the individually risky growth path.

[177] Here is the derivation:

$$\begin{aligned}
\left(\rho+\lambda\right)v\left(y\right)&=\frac{y^{1-\gamma}}{1-\gamma}+v_{y}\mu_{\mathrm{H}}y+\lambda v\left(y,\mathrm{L}\right)\\
\left(1-\gamma\right)\left(\rho+\lambda\right)v\left(y\right)&=y^{1-\gamma}\left(1+\frac{\lambda}{\rho}+\left(1-\gamma\right)\mu_{\mathrm{H}}\right)\\
v\left(y\right)&=-\frac{y^{1-\gamma}}{\left(1-\gamma\right)}\left(\frac{\gamma-1}{\rho+\lambda}\mu_{\mathrm{H}}-\frac{1}{\rho}\right).
\end{aligned}$$

Aggregate growth can be computed as follows,

$$
\begin{aligned}
dY &= P\left(\mathrm{H}\right) dy_{\mathrm{H}} + \left(1 - P\left(\mathrm{H}\right)\right) dy_{\mathrm{L}} \\
dY &= P\left(\mathrm{H}\right) y_{\mathrm{H}} \frac{dy_{\mathrm{H}}}{y_{\mathrm{H}}} \\
&\Leftrightarrow \\
\frac{dY}{Y} &= \frac{P\left(\mathrm{H}\right) y_{\mathrm{H}}}{Y} \frac{dy_{\mathrm{H}}}{y_{\mathrm{H}}} \\
\frac{dY}{Y} &= \frac{\chi y_{\mathrm{H}}}{Y} \mu_{\mathrm{H}} \\
\frac{dY}{Y} &= \frac{\chi y_{\mathrm{H}}}{\chi y_{\mathrm{H}} + \left(1 - \chi\right) \mathbb{E}\left[y | y < y_{\mathrm{H}}\right]} \mu_{\mathrm{H}} \\
\frac{dY}{Y} &= \frac{1}{1 + \left(\frac{1-\chi}{\chi}\right) \mathbb{E}\left[\frac{y}{y_{\mathrm{H}}} | y < y_{\mathrm{H}}\right]} \mu_{\mathrm{H}}
\end{aligned}
$$

Note that the distribution of the households in the low growth regime is a truncated Pareto, so $\mathbb{E}\left[\frac{y}{y_{\mathrm{H}}} | y < y_{\mathrm{H}}\right]$ has a closed form solution of the type

$$
\mathbb{E}\left[\frac{y}{y_{\mathrm{H}}} | y < y_{\mathrm{H}}\right] = \frac{1}{y_{\mathrm{H}} - y_{\mathrm{H}}^{1-\alpha}} \cdot \left(\frac{\alpha}{\alpha - 1}\right) \cdot \left(1 - \frac{1}{y_{\mathrm{H}}^{\alpha-1}}\right)
$$

where I used the fact that $y_0 = 1$ which is the the worst income any household can have in this setting. Putting the pieces together, after noting that $\chi = \exp\left(-\lambda t\right)$, top income is a function of time $y_{\mathrm{H}} = \exp\left(\mu_{\mathrm{H}} t\right)$, and the tail parameter of the Pareto distribution is $\alpha = \frac{\lambda}{\mu_{\mathrm{H}}}$, I get

$$
\begin{aligned}
g_Y\left(t\right) &= \frac{1}{1 + \frac{\left(\exp(\lambda t) - 1\right)}{\exp(\mu_{\mathrm{H}} t) - \exp\left(\left(\left(\frac{\mu_{\mathrm{H}} - \lambda}{\mu_{\mathrm{H}}}\right)\right) \mu_{\mathrm{H}} t\right)} \cdot \left(\frac{\lambda}{\lambda - \mu_{\mathrm{H}}}\right) \cdot \left(1 - \frac{1}{\exp\left(\left(\frac{\lambda - \mu_{\mathrm{H}}}{\mu_{\mathrm{H}}}\right) \mu_{\mathrm{H}} t\right)}\right)} \mu_{\mathrm{H}} \\
g_Y\left(t\right) &= \mu_{\mathrm{H}} \frac{1}{1 + \frac{\left(\exp(\lambda t) - 1\right)}{\exp(\mu_{\mathrm{H}} t) - \exp\left(\left(\mu_{\mathrm{H}} - \lambda\right) t\right)} \cdot \left(\frac{\lambda}{\lambda - \mu_{\mathrm{H}}}\right) \cdot \left(1 - \exp\left(\left(\mu_{\mathrm{H}} - \lambda\right) t\right)\right)}
\end{aligned}
$$

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

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 (1997). “Was Prometheus unbound by chance? Risk, diversification, and growth”. In: *Journal of political economy* 105.4, pp. 709–751.

— (2001). “Productivity differences”. In: *The Quarterly Journal of Economics* 116.2, pp. 563–606.

Achdou, Yves et al. (2022a). *Income and wealth distribution in macroeconomics: A continuous-time approach*. Tech. rep. Review of Economic Studies.

Achdou, Yves et al. (2022b). “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, Antonin Bergeaud, Timo Boppart, et al. (2019). *A theory of falling growth and rising rents.* Tech. rep. National Bureau of Economic Research.

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

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

Aiyagari, S Rao (1994). “Uninsured idiosyncratic risk and aggregate saving”. In: *The Quarterly Journal of Economics* 109.3, pp. 659–684.

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

Alfaro, Laura, Sebnem Kalemli-Ozcan, and Vadym Volosovych (2007). “Capital flows in a globalized world: The role of policies and institutions”. In: *Capital controls and capital flows in emerging economies: Policies, practices and consequences.* University of Chicago Press, pp. 19–72.

— (2008). “Why doesn’t capital flow from rich to poor countries? An empirical investigation”. In: *The review of economics and statistics* 90.2, pp. 347–368.

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.

Alvarez-Cuadrado, Francisco and Markus Poschke (2011). “Structural change out of agriculture: Labor push versus labor pull”. In: *American Economic Journal: Macroeconomics* 3.3, pp. 127–58.

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 (2001). “Transitional dynamics of the savings rate in the neoclassical growth model”. In: *V manuscript*.

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.

Aoki, Shuhei and Makoto Nirei (2017). “Zipf’s Law, Pareto’s Law, and the Evolution of Top Incomes in the United States”. In: *American Economic Journal: Macroeconomics* 9.3, pp. 36–71.

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 (2010). “Market penetration costs and the new consumers margin in international trade”. In: *Journal of political economy* 118.6, pp. 1151–1199.

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, Natalia Ramondo, 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.

Attanasio, Orazio and Miguel Székely (2000). “Household Saving in East Asia and Latin America: Inequality Demographics and All That”. In: *Annual World Bank Conference on Development Economics*, pp. 393–438.

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.

— (2009). “Rare disasters, asset prices, and welfare costs”. In: *American Economic Review* 99.1, pp. 243–264.

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, Gary S and Nigel Tomes (1986). "Human capital and the rise and fall of families". In: *Journal of labor economics* 4.3, Part 2, S1–S39.

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

Benhabib, Jess and Alberto Bisin (2018). "Skewed wealth distributions: Theory and empirics". In: *Journal of Economic Literature* 56.4, pp. 1261–91.

Benhabib, Jess, Alberto Bisin, and Shenghao Zhu (2015). "The wealth distribution in Bewley economies with capital income risk". In: *Journal of Economic Theory* 159, pp. 489–515.

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.

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

— (2021b). "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.

Bernanke, Ben S et al. (2005). *The global saving glut and the US current account deficit*. Tech. rep.

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.

Bewley, Truman (1977). "The permanent income hypothesis: A theoretical formulation". In: *Journal of Economic Theory* 16.2, pp. 252–292.

Blanchard, Olivier J (1985). "Debt, Deficits, and Finite Horizons". In: *The Journal of Political Economy* 93.2, pp. 223–247.

Blanden, Jo, Matthias Doepke, and Jan Stuhler (2022). *Educational inequality*. Tech. rep. National Bureau of Economic Research.

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

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

Blundell, Richard, Luigi Pistaferri, and Ian Preston (2008). “Consumption inequality and partial insurance”. In: *American Economic Review* 98.5, pp. 1887–1921.

Boar, Corina (2020). *Dynastic precautionary savings*. 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.

Bryan, Gharad, Shyamal Chowdhury, and Ahmed Mushfiq Mobarak (2014). “Underinvestment in a profitable technology: The case of seasonal migration in Bangladesh”. In: *Econometrica* 82.5, pp. 1671–1748.

Buera, Francisco J, Joseph P Kaboski, Martí Mestieri, et al. (2020). *The Stable Transformation Path*. Tech. rep. National Bureau of Economic Research.

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

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

Buera, Francisco J and Yongseok Shin (2013). “Financial frictions and the persistence of history: A quantitative exploration”. In: *Journal of Political Economy* 121.2, pp. 221–272.

— (2017). “Productivity growth and capital flows: The dynamics of reforms”. In: *American Economic Journal: Macroeconomics* 9.3, pp. 147–85.

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

Caballero, Ricardo J, Emmanuel Farhi, and Pierre-Olivier Gourinchas (2008). “An equilibrium model of “global imbalances” and low interest rates”. In: *American Economic Review* 98.1, pp. 358–93.

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.

Carroll, Christopher D (1994). "How does future income affect current consumption?" In: *The Quarterly Journal of Economics* 109.1, pp. 111–147.

— (1997). "Buffer-stock saving and the life cycle/permanent income hypothesis". In: *The Quarterly Journal of Economics* 112.1, pp. 1–55.

Carroll, Christopher D, Karen E Dynan, and Spencer D Krane (2003). "Unemployment risk and precautionary wealth: Evidence from households' balance sheets". In: *Review of Economics and Statistics* 85.3, pp. 586–604.

Carroll, Christopher D and Olivier Jeanne (2009). *A tractable model of precautionary reserves, net foreign assets, or sovereign wealth funds.* Tech. rep. National Bureau of Economic Research.

Carroll, Christopher D and Miles S Kimball (1996). "On the concavity of the consumption function". In: *Econometrica*, pp. 981–992.

Carroll, Christopher D, Jody Overland, and David N Weil (2000). "Saving and growth with habit formation". In: *American Economic Review* 90.3, pp. 341–355.

Carroll, Christopher D and Andrew A Samwick (1997). "The nature of precautionary wealth". In: *Journal of Monetary Economics* 1.40, pp. 41–71.

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.

Chamon, Marcos, Kai Liu, and Eswar Prasad (2013). "Income uncertainty and household savings in China". In: *Journal of Development Economics* 105, pp. 164–177.

Chamon, Marcos D and Eswar S Prasad (2010). "Why are saving rates of urban households in China rising?" In: *American Economic Journal: Macroeconomics* 2.1, pp. 93–130.

Chan, Kam Wing and Will Buckingham (2008). "Is China abolishing the hukou system?" In: *The China Quarterly* 195, pp. 582–606.

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.

Coeurdacier, Nicolas, Hélène Rey, and Pablo Winant (2019). “Financial integration and growth in a risky world”. In: *Journal of Monetary Economics*.

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.

Constantinides, George M and Darrell Duffie (1996). “Asset pricing with heterogeneous consumers”. In: *Journal of Political economy* 104.2, pp. 219–240.

Coppola, Antonio et al. (2020). *Redrawing the map of global capital flows: The role of cross-border financing and tax havens*. Tech. rep. National Bureau of Economic Research.

Costinot, Arnaud, Dave Donaldson, 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.

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.

Curtis, Chadwick C, Steven Lugauer, and Nelson C Mark (2015). “Demographic patterns and household saving in China”. In: *American Economic Journal: Macroeconomics* 7.2, pp. 58–94.

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 Magalhães, Leandro and Raül Santaeulàlia-Llopis (2018a). “The consumption, income, and wealth of the poorest: An empirical analysis of economic inequality in rural and urban Sub-Saharan Africa for macroeconomists”. In: *Journal of Development Economics* 134, pp. 350–371.

— (2018b). “The consumption, income, and wealth of the poorest: An empirical analysis of economic inequality in rural and urban Sub-Saharan Africa for macroeconomists”. In: *Journal of Development Economics* 134, pp. 350–371.

De Nardi, Mariacristina (2004). “Wealth inequality and intergenerational links”. In: *The Review of Economic Studies* 71.3, pp. 743–768.

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

Deaton, Angus (1991). “Saving and liquidity constraints”. In: *Econometrica* 59, pp. 1221–1248.

— (2003). “Adjusted Indian poverty estimates for 1999-2000”. In: *Economic and political Weekly*, pp. 322–326.

Deaton, Angus and Christina Paxson (1994). “Intertemporal choice and inequality”. In: *Journal of Political Economy* 102.3, pp. 437–467.

— (2000). “Growth and saving among individuals and households”. In: *Review of Economics and Statistics* 82.2, pp. 212–225.

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

— (Dec. 2020). “Changing Business Dynamism and Productivity: Shocks versus Responsiveness”. In: *American Economic Review* 110.12, pp. 3952–90.

Diamond, Peter A and Joseph E Stiglitz (1974). “Increases in risk and in risk aversion”. In: *Journal of Economic Theory* 8.3, pp. 337–360.

Ding, Haiyan and Hui He (2018). “A tale of transition: An empirical analysis of economic inequality in urban China, 1986–2009”. In: *Review of Economic Dynamics* 29, pp. 106–137.

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.

Dynan, Karen E, Jonathan Skinner, and Stephen P Zeldes (2004). "Do the Rich Save More?" In: *Journal of Political Economy* 112.2, pp. 397–444.

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.

Epstein, Larry G and Stanley E Zin (1991). "Substitution, risk aversion, and the temporal behavior of consumption and asset returns: An empirical analysis". In: *Journal of political Economy* 99.2, pp. 263–286.

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.

Fagereng, Andreas et al. (2019). *Saving behavior across the wealth distribution: The importance of capital gains*. Tech. rep. National Bureau of Economic Research.

Fan, Jingting and Şebnem Kalemli-Özcan (2016). "Emergence of Asia: reforms, corporate savings, and global imbalances". In: *IMF Economic Review* 64.2, pp. 239–267.

Fan, Jingting and Ben Zou (2019). "Industrialization from Scratch: The" Construction of Third Front" and Local Economic Development in Chinaâs Hinterland". In.

Farhi, Emmanuel and Xavier Gabaix (2016). "Rare disasters and exchange rates". In: *The Quarterly Journal of Economics* 131.1, pp. 1–52.

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.

Fuchs-Schündeln, Nicola and Matthias Schündeln (2005). "Precautionary savings and Self-Selection: Evidence from the german reunification âExperimentâ". In: *The Quarterly Journal of Economics* 120.3, pp. 1085–1120.

Gabaix, Xavier (2012). "Variable rare disasters: An exactly solved framework for ten puzzles in macro-finance". In: *The Quarterly journal of economics* 127.2, pp. 645–700.

Gabaix, Xavier et al. (2016). "The dynamics of inequality". In: *Econometrica* 84.6, pp. 2071–2111.

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.

Galor, Oded and Joseph Zeira (1993). "Income distribution and macroeconomics". In: *The review of economic studies* 60.1, pp. 35–52.

Gan, Li et al. (2019). *Relocating or Redefined: A New Perspective on Urbanization in China.* 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.

Gollin, Douglas, David Lagakos, and Michael E Waugh (2013). "The agricultural productivity gap". In: *The Quarterly Journal of Economics* 129.2, pp. 939–993.

Gollin, Douglas, David Lagakos, and Michael E Waugh (2014). “The agricultural productivity gap”. In: *The Quarterly Journal of Economics* 129.2, pp. 939–993.

Gourinchas, Pierre-Olivier and Olivier Jeanne (2013). “Capital flows to developing countries: The allocation puzzle”. In: *Review of Economic Studies* 80.4, pp. 1484–1515.

Gourinchas, Pierre-Olivier and Jonathan A Parker (2002). “Consumption over the life cycle”. In: *Econometrica* 70.1, pp. 47–89.

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.

Guvenen, Fatih (2007). “Learning your earning: Are labor income shocks really very persistent?” In: *American Economic Review* 97.3, pp. 687–712.

Guvenen, Fatih, Greg Kaplan, et al. (2017). *Lifetime incomes in the United States over six decades*. Tech. rep. National Bureau of Economic Research.

Guvenen, Fatih, Fatih Karahan, et al. (2015). *What do data on millions of US workers reveal about life-cycle earnings risk?* Tech. rep. National Bureau of Economic Research.

Hall, Robert E (1988). “Intertemporal substitution in consumption”. In: *Journal of political economy* 96.2, pp. 339–357.

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

Harris, John R and Michael P Todaro (1970). “Migration, unemployment and development: a two-sector analysis”. In: *The American economic review* 60.1, pp. 126–142.

Hausmann, Ricardo, Lant Pritchett, and Dani Rodrik (2005). “Growth accelerations”. In: *Journal of Economic Growth* 10.4, pp. 303–329.

He, Hui et al. (2018). “Breaking the “iron rice bowl:” Evidence of precautionary savings from the chinese state-owned enterprises reform”. In: *Journal of Monetary Economics* 94, pp. 94–113.

Heathcote, Jonathan, Kjetil Storesletten, and Giovanni L Violante (2014a). “Consumption and labor supply with partial insurance: An analytical framework”. In: *American Economic Review* 104.7, pp. 2075–2126.

Heathcote, Jonathan, Kjetil Storesletten, and Giovanni L Violante (2014b). "Consumption and labor supply with partial insurance: An analytical framework". In: *American Economic Review* 104.7, pp. 2075–2126.

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.

Herrendorf, Berthold, Richard Rogerson, and Akos Valentinyi (2014). "Growth and structural transformation". In: *Handbook of economic growth*. Vol. 2. Elsevier, pp. 855–941.

Hicks, Joan Hamory et al. (2017). *Reevaluating agricultural productivity gaps with longitudinal microdata*. Tech. rep. National Bureau of Economic Research.

Hnatkovska, Viktoria and Amartya Lahiri (2018). "Urbanization, Structural Transformation and Rural-Urban Disparities in China and India". In.

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 (1999). "Productivity growth and factor prices in East Asia". In: *American Economic Review* 89.2, pp. 133–138.

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

Huggett, Mark (1993). "The risk-free rate in heterogeneous-agent incomplete-insurance economies". In: *Journal of economic Dynamics and Control* 17.5-6, pp. 953–969.

Huggett, Mark, Gustavo Ventura, and Amir Yaron (2011). "Sources of lifetime inequality". In: *American Economic Review* 101.7, pp. 2923–54.

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.

İmrohoroğlu, Ayşe and Kai Zhao (2017). "Household Saving, Financial Constraints, and the Current Account in China". In: *International Economic Review*.

— (2018). "The chinese saving rate: Long-term care risks, family insurance, and demographics". In: *Journal of Monetary Economics* 96, pp. 33–52.

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

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

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

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

Jones, Charles I and Jihee Kim (2018). "A Schumpeterian model of top income inequality". In: *Journal of Political Economy* 126.5, pp. 1785–1826.

Jordà, Òscar, Moritz Schularick, and Alan M Taylor (2017). "Macrofinancial history and the new business cycle facts". In: *NBER macroeconomics annual* 31.1, pp. 213–263.

Kaboski, Joseph P and Robert M Townsend (2011). "A structural evaluation of a large-scale quasi-experimental microfinance initiative". In: *Econometrica* 79.5, pp. 1357–1406.

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.

Kogan, Leonid, Dimitris Papanikolaou, and Noah Stoffman (2020). "Left behind: Creative destruction, inequality, and the stock market". In: *Journal of Political Economy* 128.3, pp. 855–906.

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.

Krueger, Dirk, Kurt Mitman, and Fabrizio Perri (2016). “Macroeconomics and household heterogeneity”. In: 2, pp. 843–921.

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.

Krusell, Per, Toshihiko Mukoyama, and Anthony A Smith Jr (2011). “Asset prices in a Huggett economy”. In: *Journal of Economic Theory* 146.3, pp. 812–844.

Lagakos, David, Samuel Marshall, et al. (2020). “Migration costs and observational returns to migration in the developing world”. In: *Journal of Monetary Economics* 113, pp. 138–154.

Lagakos, David, Ahmed Mushfiq Mobarak, and Michael E Waugh (2018). *The welfare effects of encouraging rural-urban migration.* Tech. rep. National Bureau of Economic Research.

Lagakos, David and Michael E Waugh (2013). “Selection, agriculture, and cross-country productivity differences”. In: *American Economic Review* 103.2, pp. 948–80.

Laitner, John (1992). “Random earnings differences, lifetime liquidity constraints, and altruistic intergenerational transfers”. In: *Journal of Economic Theory* 58.2, pp. 135–170.

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.

Loury, Glenn C (1981). “Intergenerational transfers and the distribution of earnings”. In: *Econometrica: Journal of the Econometric Society*, pp. 843–867.

Lucas Jr, Robert E (2004). “Life earnings and rural-urban migration”. In: *Journal of political economy* 112.S1, S29–S59.
— (1988). “On the mechanics of economic development”. In: *Journal of monetary economics* 22.1, pp. 3–42.
— (1990). “Why doesn’t capital flow from rich to poor countries?” In: *American Economic Review* 80.2, pp. 92–96.
— (1993). “Making a Miracle”. In: *Econometrica: Journal of the Econometric Society*, pp. 251–272.
— (2009). “Ideas and growth”. In: *Economica* 76.301, pp. 1–19.
— (2018). “What Was the Industrial Revolution?” In: *Journal of Human Capital* 12.2, pp. 182–203.
Lucas, Robert E Jr and Benjamin Moll (2014). “Knowledge growth and the allocation of time”. In: *Journal of Political Economy* 122.1, pp. 1–51.
Luttmer, Erzo GJ (2007). “Selection, growth, and the size distribution of firms”. In: *The Quarterly Journal of Economics* 122.3, pp. 1103–1144.
— (2011). “On the mechanics of firm growth”. In: *The Review of Economic Studies* 78.3, pp. 1042–1068.
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.
Machado, José António, Paulo Parente, and João Santos Silva (2020). “QREG2: Stata module to perform quantile regression with robust and clustered standard errors”. In.
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.
Mas-Colell, Andreu, Michael Dennis Whinston, Jerry R Green, et al. (1995). *Microeconomic theory*. Vol. 1. Oxford university press New York.
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.
Mendoza, Enrique, Vincenzo Quadrini, and Viktor Rios-Rull (2009). “Financial Integration, Financial Deepness and Global Imbalances”. In: *Journal of Political Economy* 117.3, p. 371.

Merton, Robert C (1969). "Lifetime portfolio selection under uncertainty: The continuous-time case". In: *The review of Economics and Statistics*, pp. 247–257.

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.

Modigliani, Franco (1986). "Life cycle, individual thrift, and the wealth of nations". In: *Science* 234.4777, pp. 704–712.

Modigliani, Franco and Shi Larry Cao (2004). "The Chinese saving puzzle and the life-cycle hypothesis". In: *Journal of Economic Literature* 42.1, pp. 145–170.

Moll, Benjamin (Oct. 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.

Morten, Melanie (2019). "Temporary migration and endogenous risk sharing in village india". In: *Journal of Political Economy* 127.1, pp. 1–46.

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 and Edward C Prescott (1991). *Technology Adoption and Growth*.

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

— (2014b). "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)*.

Piketty, Thomas, Li Yang, and Gabriel Zucman (2019). “Capital accumulation, private property, and rising inequality in China, 1978–2015”. In: *American Economic Review* 109.7, pp. 2469–96.

Prasad, E, R Rajan, and A Subramanaian (2007). “The paradox of capital, finance and development”. In: *IMF, March* 44.

Qin, Bo and Yu Zhang (2014). “Note on urbanization in China: Urban definitions and census data”. In: *China Economic Review* 30, pp. 495–502.

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.

Restuccia, Diego, Dennis Tao Yang, and Xiaodong Zhu (2008). “Agriculture and aggregate productivity: A quantitative cross-country analysis”. In: *Journal of Monetary Economics* 55.2, pp. 234–250.

Rietz, Thomas A (1988). “The equity risk premium a solution”. In: *Journal of monetary Economics* 22.1, pp. 117–131.

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.

Rosenzweig, Mark R and Oded Stark (1989). “Consumption smoothing, migration, and marriage: Evidence from rural India”. In: *Journal of political Economy* 97.4, pp. 905–926.

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.

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

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.

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

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

Sandri, Damiano (2014). "Growth and capital flows with risky entrepreneurship". In: *American Economic Journal: Macroeconomics* 6.3, pp. 102–23.

Santaeulalia-Llopis, Raul and Yu Zheng (2018). "The price of growth: consumption insurance in China 1989–2009". In: *American Economic Journal: Macroeconomics* 10.4, pp. 1–35.

Schechtman, Jack (1976). "An income fluctuation problem". In: *Journal of Economic Theory* 12.2, pp. 218–241.

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.

Shi, Li (2009). "Chinese household income project, 2002". In: *Ann Arbor, MI: Inter-university Consortium for Political and Social Research [distributor]*, pp. 08–14.

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

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

Solow, Robert M (1956). "A contribution to the theory of economic growth". In: *The quarterly journal of economics* 70.1, pp. 65–94.

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.

Song, Yang (2014). "What should economists know about the current Chinese hukou system?" In: *China Economic Review* 29, pp. 200–212.

Song, Zheng, Kjetil Storesletten, and Fabrizio Zilibotti (2011). "Growing like China". In: *American Economic Review* 101.1, pp. 196–233.

Storesletten, Kjetil, Christopher I Telmer, and Amir Yaron (2004). "Consumption and risk sharing over the life cycle". In: *Journal of Monetary Economics* 51.3, pp. 609–633.

Straub, Ludwig (2019). "Consumption, Savings, and the Distribution of Permanent Income". In: *Unpublished manuscript, Harvard University.*

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

Tombe, Trevor and Xiaodong Zhu (2019). "Trade, migration, and productivity: A quantitative analysis of china". In: *American Economic Review* 109.5, pp. 1843–72.

Townsend, Robert M (1994). "Risk and insurance in village India". In: *Econometrica* 62.3, pp. 539–591.

Trouvain, Florian (2021). *Structural Change, Inequality, and Capital Flows*. Tech. rep. Unpublished Manuscript, University of Michigan.

— (2023). *Urban-Rural Wealth Gaps, Human Capital Risk, and Aggregate Savings along the Development Path*. Tech. rep. Work in progress, University of Michigan.

Uribe, Martin and Stephanie Schmitt-Grohé (2017). *Open economy macroeconomics*. Princeton University Press.

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.

Wei, Shang-Jin and Xiaobo Zhang (2011). "The competitive saving motive: Evidence from rising sex ratios and savings rates in China". In: *Journal of Political Economy* 119.3, pp. 511–564.

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

Xie, Yu and Jingwei Hu (2014). "An introduction to the China family panel studies (CFPS)". In: *Chinese sociological review* 47.1, pp. 3–29.

Xu, Hongwei and Yu Xie (2015). "The causal effects of rural-to-urban migration on childrenâs well-being in China". In: *European sociological review* 31.4, pp. 502–519.

Yaari, Menahem E (1965). "Uncertain lifetime, life insurance, and the theory of the consumer". In: *The Review of Economic Studies* 32.2, pp. 137–150.

Young, Alwyn (1995). "The tyranny of numbers: confronting the statistical realities of the East Asian growth experience". In: *The Quarterly Journal of Economics* 110.3, pp. 641–680.

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

— (2013). "Inequality, the urban-rural gap, and migration". In: *The Quarterly Journal of Economics* 128.4, pp. 1727–1785.

APPENDICES

Appendices to Chapter I

A.1 Theory Appendix

A.1.1 Production Firm

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

$$\begin{array}{ll} \max & \int_0^A \left(\frac{x_j}{\alpha}\right)^\alpha dj \\ s.t. & \int p_j x_j dj \leq I. \end{array} \tag{A1}$$

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

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

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

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

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

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

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

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,

$$\begin{aligned} v\left(\rho + \delta_{ex}\right) &= \max_h \frac{\pi_t(z)}{w} - s_t h + v_z \dot{z} + \dot{v} \\ s.t. & \\ \dot{z} &= \zeta z^{\theta} h^{\beta} - \left(g_F + \delta_I\right) z. \end{aligned} \tag{A2}$$

A solution to the program (A2) needs to satisfy the following first order condition

$$\left\{\frac{v_z \beta \zeta z^{\theta}}{s}\right\}^{\frac{1}{1-\beta}} = h \cdot \tag{A3}$$

Equation (A3) 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

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

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

$$h^{1-\beta} = \frac{1}{s}\frac{\beta(1-\alpha)(\sigma-1)(g_F+\delta_I)}{\rho+\delta_{ex}+(1-\theta)(\delta_I+g_F)}\left[\frac{\pi}{w}\right]\frac{\zeta z^{\theta-1}}{(g_F+\delta_I)} \tag{A4}$$

$$z^{1-\theta} = \frac{\zeta h^{\beta}}{g_F+\delta_I} \tag{A5}$$

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.

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

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

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

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

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,

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

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 D. Romer (2012). Instead of using the HJB, I can define the current value Hamiltonian,

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

The optimality conditions are standard and read

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

and

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

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

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

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

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

In standard q-theory applications, θ equals one and π_z is falling in z.[178] The advantage of backwardness embodied in $1-\theta$ shows up in the firm problem and looks like an additional discount factor. The reader might note the strong resemblance to the neoclassical growth model where $\alpha k^{\alpha-1} = \rho + g + \delta$. Faster growth requires a higher return to capital, but the effect of this is attenuated for large θ as the diminishing returns in the accumulation of knowledge stock A disappear.

Another intuitive implication of the theory is that

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

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

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

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

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

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 1.2.3 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 (A4) establishes a negative link between z and h, while (A5) 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

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

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

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

A.1.2 Innovation problem and market clearing conditions for high skilled labor

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 (1.19) in the future by taking account of the waiting time δ leads to

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

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

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

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

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

$$(r+\delta_I)\,V_I - \dot{V}_I = \exp\left(-\int_t^{t+\tau}(r+\delta_I)\,dx\right)\frac{\alpha L_{t+\tau}^P w_{t+\tau}}{A_{F,t+\tau} z_{t+\tau}}\left[1+\tau'\right]. \tag{A6}$$

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 (A6) 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

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

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

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

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

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

and now combine the two to get

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

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

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

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

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

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

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

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.

Waiting time for innovator The waiting time for an innovator can be derived as follows. Recall equation (1.20). Use an integrating factor and note that on the balanced growth path with a constant adoption gap, $g_A = g_F$. Then,

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

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

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

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

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

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

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

which generalizes and nests the steady state result.

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

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

I totally differentiate this expression to obtain

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

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.

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.1.3 Nesting Jones (1995)

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 A.1.3 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.[179]

A.1.4 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,

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

[179] 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.[180]

A.1.5 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

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

It follows that

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

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

[180] 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.

yields

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

$$\Leftrightarrow$$

$$
\begin{aligned}
\left(\frac{\gamma}{\gamma^*}\right)^{\frac{1}{1-\lambda}}\left(\frac{s}{s^*}\frac{z}{z^*}\right)^{-\frac{\lambda}{1-\lambda}} &= \left(\frac{z}{z^*}\right)\frac{sh^{tot}-\Lambda^D}{s^*h^{tot,*}-\Lambda^D} \\
\left(\frac{\gamma}{\gamma^*}\right)^{\frac{1}{1-\lambda}}\left(\frac{s}{s^*}\left(\frac{s}{s^*}\right)^{-\frac{\beta}{1-\theta}}\right)^{-\frac{\lambda}{1-\lambda}} &= \left(\frac{s}{s^*}\right)^{-\frac{\beta}{1-\theta}}\frac{sh^{tot}-\Lambda^D}{s^*h^{tot,*}-\Lambda^D} \\
\left(\frac{\gamma}{\gamma^*}\right)\left(\left(\frac{s}{s^*}\right)^{\frac{1-\theta-\beta}{1-\theta}}\right)^{-\lambda} &= \left(\frac{s}{s^*}\right)^{-\frac{(1-\lambda)\beta}{1-\theta}}\left(\frac{sh^{tot}-\Lambda^D}{s^*h^{tot,*}-\Lambda^D}\right)^{1-\lambda} \\
\left(\frac{\gamma}{\gamma^*}\right) &= \left(\frac{s}{s^*}\right)^{\frac{(1-\theta-\beta)\lambda-\beta(1-\lambda)}{1-\theta}}\left(\frac{sh^{tot}-\Lambda^D}{s^*h^{tot,*}-\Lambda^D}\right)^{1-\lambda} \\
\left(\frac{\gamma}{\gamma^*}\right) &= \left(\frac{s}{s^*}\right)^{\frac{(1-\theta-\beta)\lambda-\beta(1-\lambda)}{1-\theta}}\left(\frac{sh^{tot}-\Lambda^D}{s^*h^{tot,*}-\Lambda^D}\right)^{1-\lambda} \\
\left(\frac{\gamma}{\gamma^*}\right) &= \left(\frac{s}{s^*}\right)^{\frac{(1-\theta-\beta)\lambda-\beta(1-\lambda)-(1-\lambda)(1-\theta)}{1-\theta}}\left(\frac{s}{s^*}\right)^{1-\lambda}\left(\frac{sh^{tot}-\Lambda^D}{s^*h^{tot,*}-\Lambda^D}\right)^{1-\lambda} \\
\left(\frac{\gamma}{\gamma^*}\right) &= \left(\frac{s}{s^*}\right)^{\frac{(1-\theta-\beta)\lambda+(1-\lambda)(1-\theta-\beta)}{1-\theta}}\left(\frac{h^{tot}-\frac{\Lambda^D}{s}}{h^{tot,*}-\frac{\Lambda^D}{s^*}}\right)^{1-\lambda} \\
\left(\frac{\gamma}{\gamma^*}\right) &= \left(\frac{s}{s^*}\right)^{\frac{1-\theta-\beta}{1-\theta}}\left(\frac{h^{tot}-\frac{\Lambda^D}{s}}{h^{tot,*}-\frac{\Lambda^D}{s^*}}\right)^{1-\lambda} \qquad \text{(A7)}
\end{aligned}
$$

Assumption that $\theta + \beta < 1$ is important because you want that skilled labor becomes more expensive in real terms when demand goes up. If not, the real wage of skilled labor would be higher in places with a lower skill premium. If that is desired the reader can flip the inequality but care must be taken that the relevant computational inequalities, especially 1.2.3, 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 A7 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/zˆ* 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 A7.

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.

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.

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

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

A.1.6 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 (2021a) where an entrant arrives with a small probability p right at the frontier $A_{K,\max}$. [181]

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 $\dot{v} = 0$)

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

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

[181] This is a technical assumption to ensure a stationary distribution emerges, similar to Benhabib, Perla, and Tonetti (2021a). 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 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

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

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

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

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

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

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 (2021a) and Perla and Tonetti (2014b). 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

$$V = \mathbb{E}_z [V(z)] \tag{A8}$$

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 (A8) 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.

A.2 Transition Dynamics

A.2.1 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, E. Liu, and S. J. 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

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

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

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

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.

A.2.2 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. (2022b). 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.[182]

[182] 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.

A.2.3 Production Firm Problem

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 (1.21) yields the following free entry condition that holds off and on the balanced growth path

$$f_e = \frac{\frac{\pi}{w} - sh}{r + \delta_{ex} - g_w} . \tag{A9}$$

Rearranging (A9) 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

$$\frac{\dot{h}}{h}(1-\beta) = \rho + \delta_{ex} + (1-\theta)(g_F + \delta_I) \tag{A10}$$

$$- \frac{\beta z^{\theta-1}\zeta h^{\beta-1}(1-\alpha)(\sigma-1)}{s}\left[\min\left\{f_e\,(r + \delta_{ex} - g_w) + sh, \frac{1}{m(1-\alpha)(\sigma-1)}\right\}\right] + \frac{\dot{s}}{s} \tag{A11}$$

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

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

A.2.4 Solving the model

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

The normalized firm problem reads

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

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 $\overrightarrow{X}$.[183] 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 $l^P=\min\left\{m(1-\alpha)(\sigma-1)\left[sh+f_e\left(r+\delta_{ex}-g_w\right)\right], 1\right\}$. 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. (2022b) 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 (A10) 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

[183] 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.

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

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

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

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

.

The derivation follows from setting

$$
\begin{aligned}
\underline{\dot{z}} &= 0 \\
&\Leftrightarrow \\
\underline{h} &= \left(\frac{\underline{z}^{1-\theta}}{\zeta}\left(\delta_I + g_{A_F}\right)\right)^{\frac{1}{\beta}} \\
v_{\underline{z_i}} &\geq \frac{s}{\underline{z_i}^{\theta}\beta\zeta}\left(\underline{h}\right)^{1-\beta} \\
v_{\underline{z_i}} &\geq \frac{s}{\underline{z_i}^{\theta}\beta\zeta}\left(\left(\frac{\underline{z}^{1-\theta}}{\zeta}\left(\delta_I + g_{A_F}\right)\right)^{\frac{1}{\beta}}\right)^{1-\beta} \\
v_{\underline{z_i}} &\geq \frac{s}{\beta\zeta}\underline{z_i}^{-\theta+(1-\theta)\left(\frac{1-\beta}{\beta}\right)}\left(\frac{\delta_I + g_{A_F}}{\zeta}\right)^{\frac{1-\beta}{\beta}} \\
v_{\underline{z_i}} &\geq \frac{s}{\beta\left(\zeta\right)^{\frac{1}{\beta}}}\left(\underline{z_i}\right)^{\frac{1-\theta-\beta}{\beta}}\left(\delta_I + g_{A_F}\right)^{\frac{1-\beta}{\beta}}
\end{aligned}
$$

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. (2022b) and use the Hamiltonian instead to pick the right drift, a strategy the authors discuss in their supplementary material.

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.[184] 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

[184]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.

labor in production but it is conceptually quite similar, which in turn shapes the stock of knowledge z which is the key state variable.

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 $h_t\left(m, z_i, z_{agg}\right)$. 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 $g\left(z_0, m_0\right) = 1$ which also has a unique choice $h_0\left(z_0, m_0\right)$. 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

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

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.

A.2.5 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

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

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 $w A_F^{-\phi}$. 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

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

Note that the recursion looks very similar in the open economy where the value function is split into a domestic and a foreign piece. To be precise,

$$
\begin{aligned}
v^{open} = &\underbrace{\Lambda_t\left(\{\psi\},t,t+\tau\right)\Delta\left(1+\tau'\right)\pi_{t+\tau}+\exp\left(-\Delta\psi_t\right)v_{t+\Delta}^{IN}}_{domestic} \\
&+\underbrace{\Lambda_t\left(\{\psi^*\},t,t+\tau^*\right)\Delta\left(1+\tau'^{*}\right)\pi_{t+\tau^*}^{*}+\exp\left(-\Delta\psi_t^*\right)v_{t+\Delta}^{IN,*}}_{foreign}.
\end{aligned}
$$

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

$$
v_{IN}=f_R s
$$

so

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

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

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

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

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

Moreover, note that the equilibrium solution to the computational solution reads $\mathbb{E}\left[hm\right]$ 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

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

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

$$\frac{-\log z}{\delta_I + \frac{\int_t^{t+\tau} g_A dx}{\tau}} = \tau \tag{A12}$$

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 (A12) 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.

A.2.6 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.

Here is the math:

First, compute the world value function of an innovation.

dition, and after some algebra

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

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

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

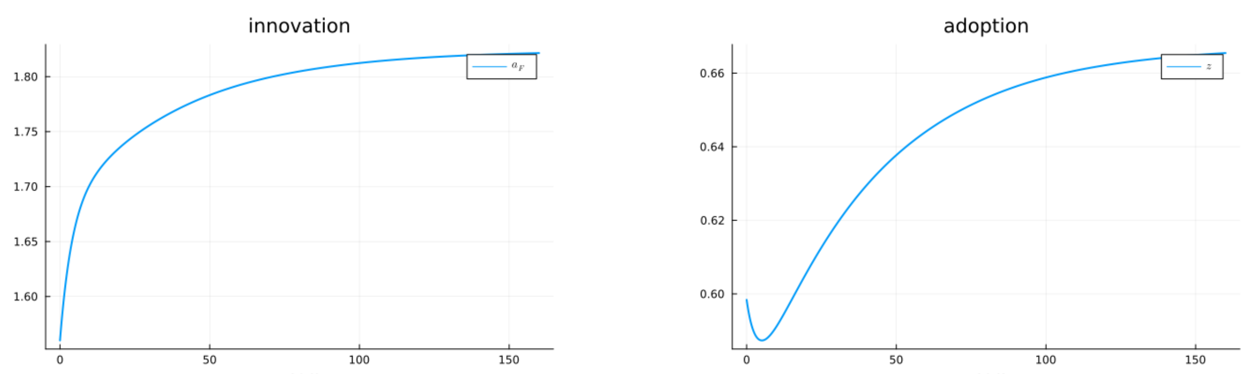
Focusing on the advanced economy, for a sequence of skill prices etc, one can invert the right hand side to obtain equilibrium demand for normalized skilled labor. Then update the price of skill gently to get at something reasonable. Note that this system depends on what happens in the foreign economy,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

A.2.7 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 A1 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.

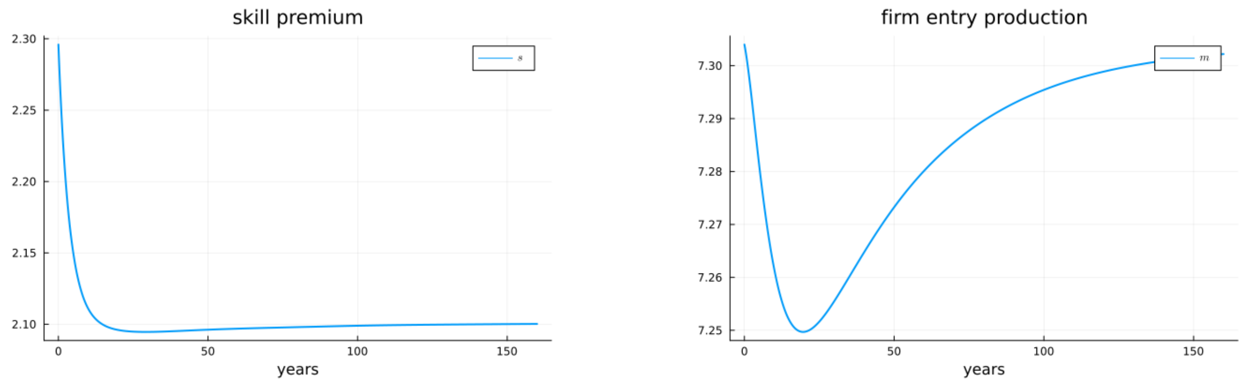
Figure A1: Innovation and Adoption in Closed Economy



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 A2: Innovation and Adoption in Closed Economy – Details



Time units are in $\Delta = 0.1$.

A.3 Extensions

A.3.1 Immigration

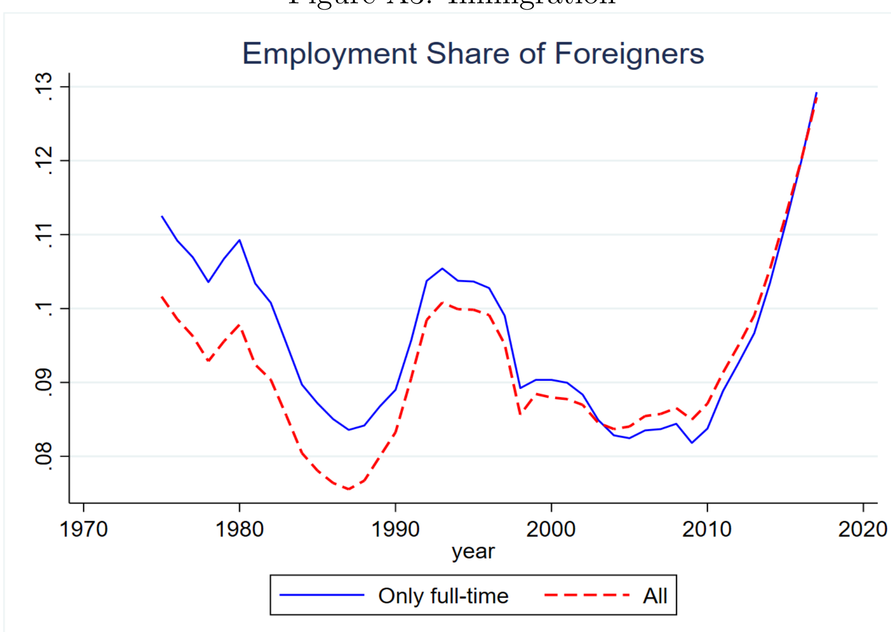
A fully integrated equilibrium behaves differently from the baseline model. Note that the factor price equalization theorem does not hold precisely because countries have different research productivities so goods market trade is no substitute for immigration. World output would be maximized by moving 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 (1.41). 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 A3 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 A3: Immigration



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

A.3.2 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 a symmetric equilibrium with same amount of research obtains.

A.3.3 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 (1.41) 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 $h(1) = 0$.[185] Again, such a model offers more scope for production labor to gain from market integration.

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

A.4 IAB DATA

The data I use is an establishment panel (BHP) provided by the IAB, which 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 information on educational attainment and average wage. High educational attainment refers to college or comparable (Meister degree) and I use this variable to stand in for skilled labor. Pre 1992 the data only contains West German establishments, while from 1992 East German establishments located in the former DDR are included as well.

A.4.1 Sample Selection and Weighting

I cut the establishment sample as follows. I drop all observations in the former DDR (East Germany) including all of Berlin. I use Kosfeld and Werner (2012)'s definition of local labor markets (excluding Berlin) which leaves me with 108 regions. I drop establishments with missing average wages or negative average wages as well as missing employment or negative employment. I then winsorize observations within each year at the lowest 1 percentile, and the highest 99.9 percentile.

I weight regressions by full time employment (az_vz), and I focus on stacked differences using the years 1975, 1985, 1995, 2007, 2015. Most of the time I focus on the episode from 1985 to 2007, which captures strong convergence dynamics in the earlier period, while a remarkable divergence occurs in the later period. Neither the period 1975 – 1985 nor the period 2007 – 2015 displays very strong con- or divergence although I note that regional convergence was a striking feature of early postwar growth in the US and Europe, see figure A6.

A.5 Information on Patent Data

The data is provided by Crios-Patstat Coffano and Tarasconi (2014) and contains patent data from the European Patent Office (EPO) from 1977 - 2014. Data on population and area by county is provided by Roesel (2022). I obtain the priority date of patent application, as well as cites of each patent within a 5-year window after the priority date, to compute county-level patenting activity.

County-level patenting activity is computed by summing over all patent applications within a year, using the priority date. I offer an alternative estimate where I apply weights w_i of the form $w_i = 1 + number\_cites$ where each patent is weighted by total cites within a 5-year period. On the local labor market level, these two measures are almost perfectly correlated ($\rho \approx .99$) and I focus on the raw patent count in the paper.

Use inventors or applicants?

rerun the simple thing with the specializaiton population growth exercise and check which one works best

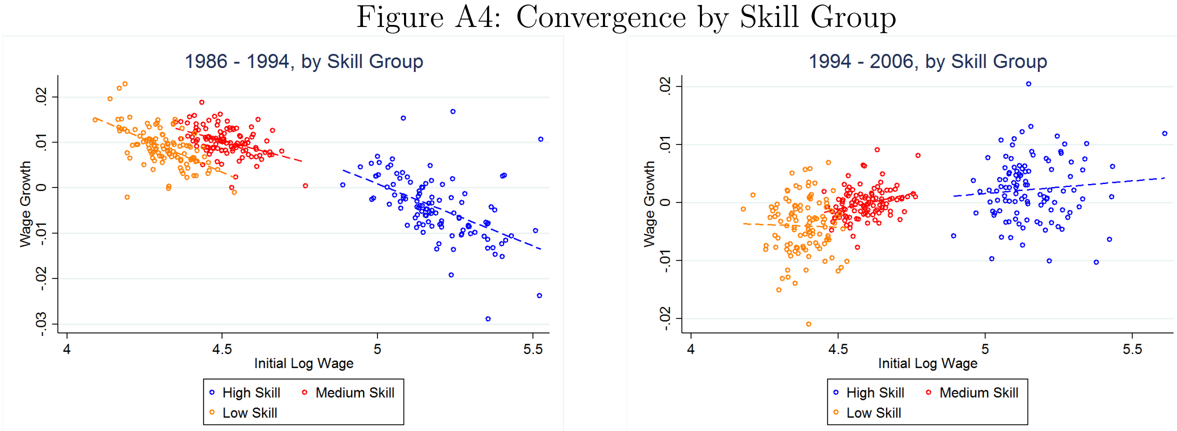
redo the shift share, also try trade and fdi, and try to leverage technology classes more 00 this must work!

the following files to build up the dataset:

- gen_regional_patent.do contains the code
- “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.

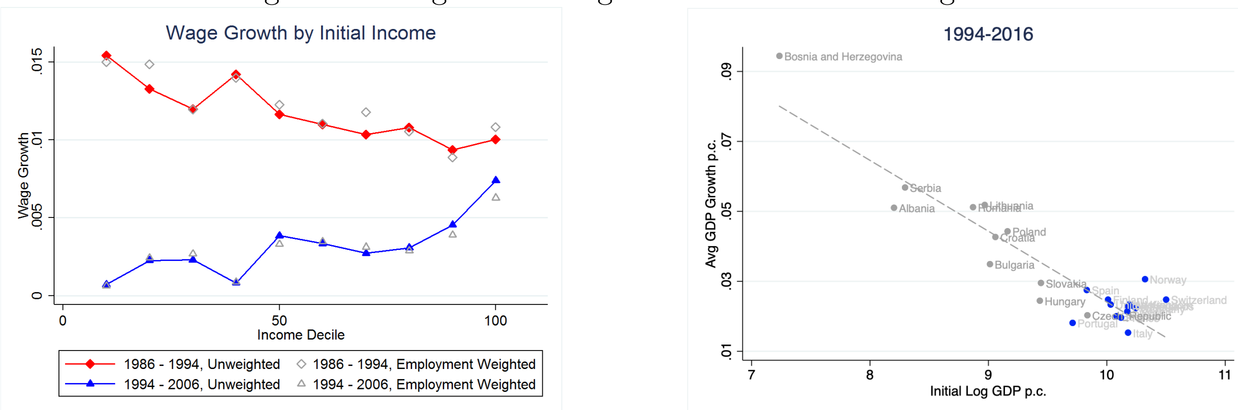
A.6 Changing Convergence Dynamics

Figure A4: Convergence by Skill Group



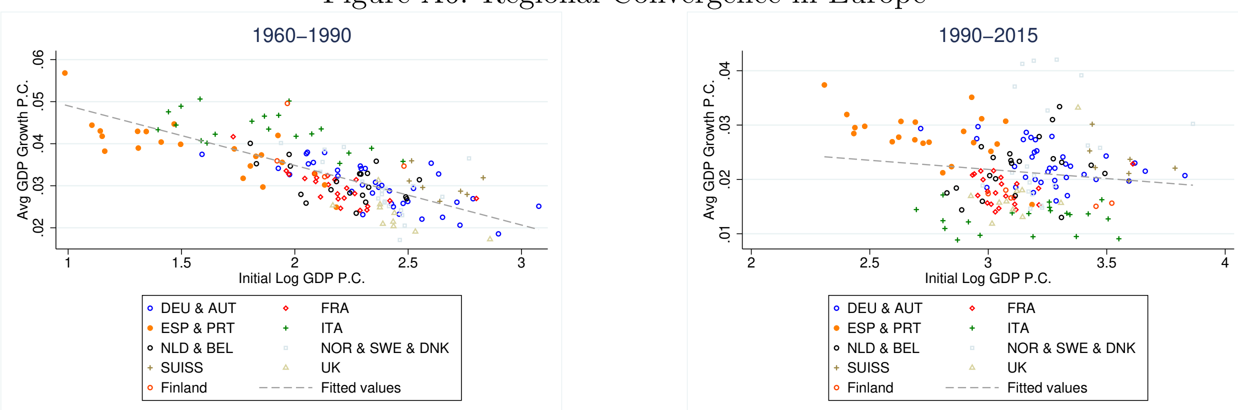
IAB BHP data. My plots.

Figure A5: 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.

Figure A6: 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.

A.7 Additional Results from Barro Catch-up Regression for Regions in Germany

Employment, High-Skill Firms, and Professional Occupations

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

for International Trade

Building on the work of D. 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, $\Delta\left(Import\, exp\right)_{j,t}^{East} = \sum_j \frac{E_{j,i,t}}{E_{i,t}} \frac{\Delta Im_{i,t}^{D \leftarrow East}}{E_{j,t}}$, where $\Delta Im_{i,t}^{D \leftarrow East}$ is the total increase in real imports (total value deflated in 1998 Euros) from the East, here including both China as well as Eastern Europe and Eurasia. This choice is informed by the fact that the rise in the German trade-to-GDP ratio is largely attributable to the rise of China and the fall of the Iron Curtain (Dauth, Findeisen, and Suedekum 2014). The relevant 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 D. 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 S. O. Becker et al. (2019) in mapping BACI and OECD industries to the German industry classification, see their paper for details.

A.8 Additional Information on Aggregate Wages and Employment

A.8.1 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.

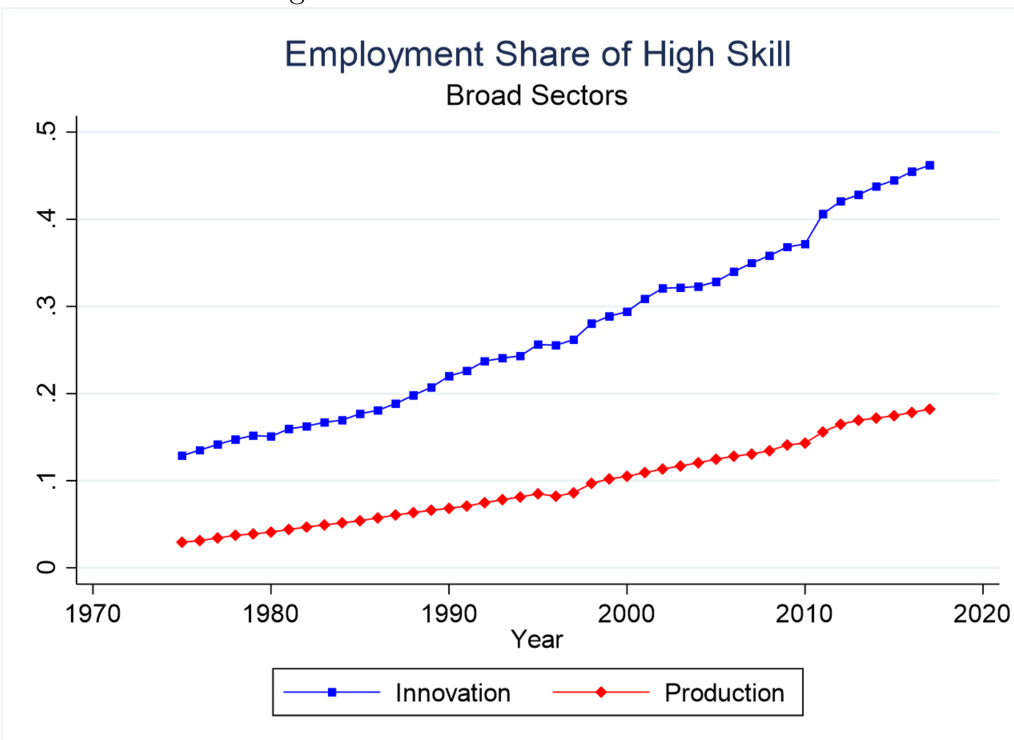
Figure A7: Sectoral Classification Innovation vs. Production

Industry Digit	variable_n	sector code	label	baseline	baseline plus Finance and IT
5-Steller	(w93_5)	65110	Zentralbanken	0	1
5-Steller	(w93_5)	65121	Kreditbanken einschliesslich Zw	0	1
5-Steller	(w93_5)	65124	Genossenschaftliche Zentralban	0	1
5-Steller	(w93_5)	65126	Realkreditinstitute	0	1
5-Steller	(w93_5)	65127	Kreditinstitute mit Sonderaufg	0	1
5-Steller	(w93_5)	65210	Institutionen fuer Finanzierung	0	1
5-Steller	(w93_5)	65220	Spezialkreditinstitute	0	1
5-Steller	(w93_5)	65231	Kapitalanlagegesellschaften	0	1
5-Steller	(w93_5)	65233	Sonstige Finanzierungsinstitut	0	1
5-Steller	(w93_5)	67110	Effekten- und Warenterminboerse	0	1
5-Steller	(w93_5)	67120	Effektenver und	0	1
5-Steller	(w93_5)	67130	Sonstige mit dem Kreditgewerbe	0	0
5-Steller	(w93_5)	72201	Softwareberatung	0	1
5-Steller	(w93_5)	72202	Softwareentwicklung	0	1
5-Steller	(w93_5)	72301	Datenerfassungsdienste	0	1
5-Steller	(w93_5)	72302	Datenverarbeitungs- und Tabell	0	1
5-Steller	(w93_5)	72303	Bereitstellungsdienste fuer Tei	0	1
5-Steller	(w93_5)	72304	Sonstige Datenverarbeitungsdie	0	1
5-Steller	(w93_5)	72400	Datenbanken	0	1
5-Steller	(w93_5)	72500	Instandhaltung und Reparatur v	0	0
5-Steller	(w93_5)	72601	Informationsvermittlung	0	1
5-Steller	(w93_5)	72602	Mit der Datenverarbeitung verb	0	1
5-Steller	(w93_5)	73101	Forschung und Entwicklung im	1	1
5-Steller	(w93_5)	73102	Forschung und Entwicklung im	1	1
5-Steller	(w93_5)	73103	Forschung und Entwicklung im	1	1
5-Steller	(w93_5)	73104	Forschung und Entwicklung im	1	1
5-Steller	(w93_5)	73105	Forschung und Entwicklung im	1	1
5-Steller	(w93_5)	73201	Forschung und Entwicklung im	1	1
5-Steller	(w93_5)	73202	Forschung und Entwicklung im	1	1
5-Steller	(w93_5)	74111	Rechtsanwaltskanzleien mit Not	0	0
5-Steller	(w93_5)	74112	Rechtsanwaltskanzleien ohne No	0	0
5-Steller	(w93_5)	74113	Notariat	0	0
5-Steller	(w93_5)	74114	Patentanwaltskanzleien	1	1
5-Steller	(w93_5)	74115	Sonstige Rechtsberatung	0	0
5-Steller	(w93_5)	74121	Praxen von Wirtschaftspruefern,	0	0
5-Steller	(w93_5)	74122	Praxen von vereidigten Buchprue	0	0
5-Steller	(w93_5)	74131	Marktforschung	1	1
5-Steller	(w93_5)	74132	Meinungsforschung	0	0
5-Steller	(w93_5)	74141	Unternehmensberatung	1	1
5-Steller	(w93_5)	74142	Public-Relations-Beratung	1	1
5-Steller	(w93_5)	74151	Beteiligungsgesellschaften mit	0	1
5-Steller	(w93_5)	74152	Sonstige Beteiligungsgesellsch	0	1
5-Steller	(w93_5)	74153	Geschlossene Immobilienfonds m	0	0
5-Steller	(w93_5)	74154	Geschlossene Immobilienfonds m	0	0
5-Steller	(w93_5)	74155	Komplementaergesellschaften	0	1
5-Steller	(w93_5)	74156	Verwaltung und Fuehrung von	1	1
5-Steller	(w03_5)	65110	Zentralbanken	0	1
5-Steller	(w03_5)	65121	Kreditbanken einschliesslich Zwei	0	1
5-Steller	(w03_5)			0	
5-Steller	(w03_5)	65124	Genossenschaftliche Zentralbanker	0	1
5-Steller	(w03_5)	65126	Realkreditinstitute	0	1
5-Steller	(w03_5)	65127	Kreditinstitute mit Sonderaufgabe	0	1
5-Steller	(w03_5)	65210	Institutionen fuer Finanzierungsl	0	1
5-Steller	(w03_5)	65220	Spezialkreditinstitute	0	1
5-Steller	(w03_5)	65231	Kapitalanlagegesellschaften	0	1
5-Steller	(w03_5)	65233	Sonstige Finanzierungsinstitution	0	1
5-Steller	(w03_5)	67110	Effekten- und Warenboersen	0	1
5-Steller	(w03_5)	67120	Effektenvermittlung und	0	1
5-Steller	(w03_5)			0	
5-Steller	(w03_5)	72210	Verlegen von Software	0	1
5-Steller	(w03_5)	72221	Softwareberatung	0	1
5-Steller	(w03_5)	72222	Entwicklung und Programmierung	0	1
5-Steller	(w03_5)	72223	Sonstige Softwareentwicklung	0	1
5-Steller	(w03_5)	72301	Datenerfassungsdienste	0	1
5-Steller	(w03_5)	72303	Bereitstellungsdienste fuer Teiln	0	1
5-Steller	(w03_5)	72305	Sonstige Datenverarbeitungsdienst	0	1
5-Steller	(w03_5)	72400	Datenbanken	0	1
5-Steller	(w03_5)	72601	Informationsvermittlung	0	1
5-Steller	(w03_5)	72602	Mit der Datenverarbeitung verbun	0	1
5-Steller	(w03_5)	73101	Forschung und Entwicklung im	1	1
5-Steller	(w03_5)	73102	Forschung und Entwicklung im	1	1
5-Steller	(w03_5)	73103	Forschung und Entwicklung im	1	1
5-Steller	(w03_5)	73104	Forschung und Entwicklung im	1	1
5-Steller	(w03_5)	73105	Forschung und Entwicklung im	1	1
5-Steller	(w03_5)	73201	Forschung und Entwicklung im	1	1
5-Steller	(w03_5)	73202	Forschung und Entwicklung im	1	1
5-Steller	(w03_5)	74111	Rechtsanwaltskanzleien mit Notari	0	0
5-Steller	(w03_5)	74112	Rechtsanwaltskanzleien ohne Nota	0	0
5-Steller	(w03_5)	74113	Notariate	0	0
5-Steller	(w03_5)	74114	Patentanwaltskanzleien	1	1
5-Steller	(w03_5)	74115	Sonstige Rechtsberatung	0	0
5-Steller	(w03_5)	74121	Praxen von Wirtschaftsprueferinne	0	0
5-Steller	(w03_5)	74122	Praxen von vereidigten Buchpruefe	0	0
5-Steller	(w03_5)	74123	Praxen von Steuerberaterinnen un	0	0
5-Steller	(w03_5)	74124	Praxen von Steuerbevollmaechtigte	0	0
5-Steller	(w03_5)	74125	Buchfuehrung (ohne Datenverarbe	0	0

Industry Di	variable_n	sector cod	label	baseline	baseline plus 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 Ho	1	1
5-Steller	(w03_5)	74152	Managementtaetigkeiten von sor	1	1
5-Steller	(w03_5)	74153	Geschlossene Immobilienfonds m	0	0
5-Steller	(w03_5)	74154	Geschlossene Immobilienfonds m	0	0
5-Steller	(w03_5)	74155	Komplementaergesellschaften	0	1
5-Steller	(w03_5)	74156	Verwaltung und Fuehrung von	1	1
5-Steller	(w08_5)	62011	Entwicklung und Programmierung	0	1
5-Steller	(w08_5)	62019	Sonstige Softwareentwicklung	0	1
5-Steller	(w08_5)	62020	Erbringung von Beratungsleistung	0	1
5-Steller	(w08_5)	62030	Betrieb von Datenverarbeitungse	0	1
5-Steller	(w08_5)	62090	Erbringung von sonstigen Dienstl	0	1
5-Steller	(w08_5)	63110	Datenverarbeitung, Hosting und	0	1
5-Steller	(w08_5)	63120	Webportale	0	0
5-Steller	(w08_5)	63910	Korrespondenz- und Nachrichtenl	0	0
5-Steller	(w08_5)	63990	Erbringung von sonstigen Informa	0	0
5-Steller	(w08_5)	64110	Zentralbanken	0	1
5-Steller	(w08_5)	64191	Kreditbanken einschliesslich Zwei	0	0
5-Steller	(w08_5)	64192	Kreditinstitute des Sparkassensek	0	0
5-Steller	(w08_5)	64193	Kreditinstitute des Genossenscha	0	0
5-Steller	(w08_5)	64194	Realkreditinstitute	0	1
5-Steller	(w08_5)	64195	Kreditinstitute mit Sonderaufgab	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 No	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 Wirtschaftsprueferinn	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 Ho	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

In the following plot A8 I show how the skill share in the research sector diverges from the skill share in the production sector.

Figure A8: Skill-Share across Sectors

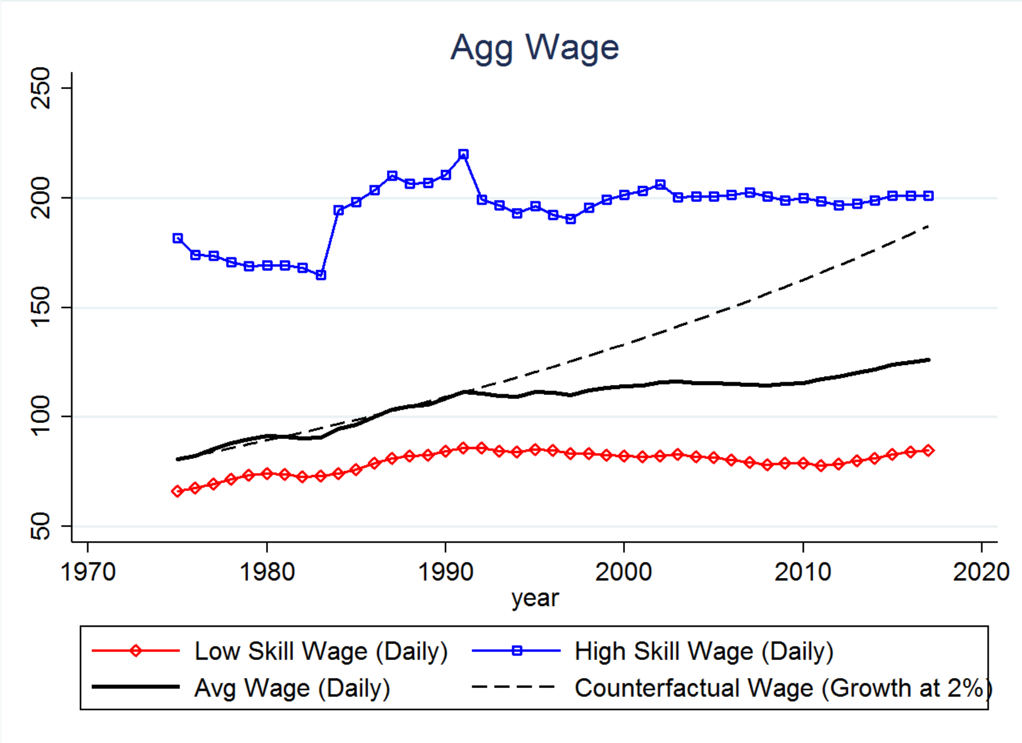


IAB BHP data. Measure divides full time skilled labor in each sector-group by total full time employment. Note the divergence that sets in since the 1990s.

A.8.2 Wage Stagnation in Germany

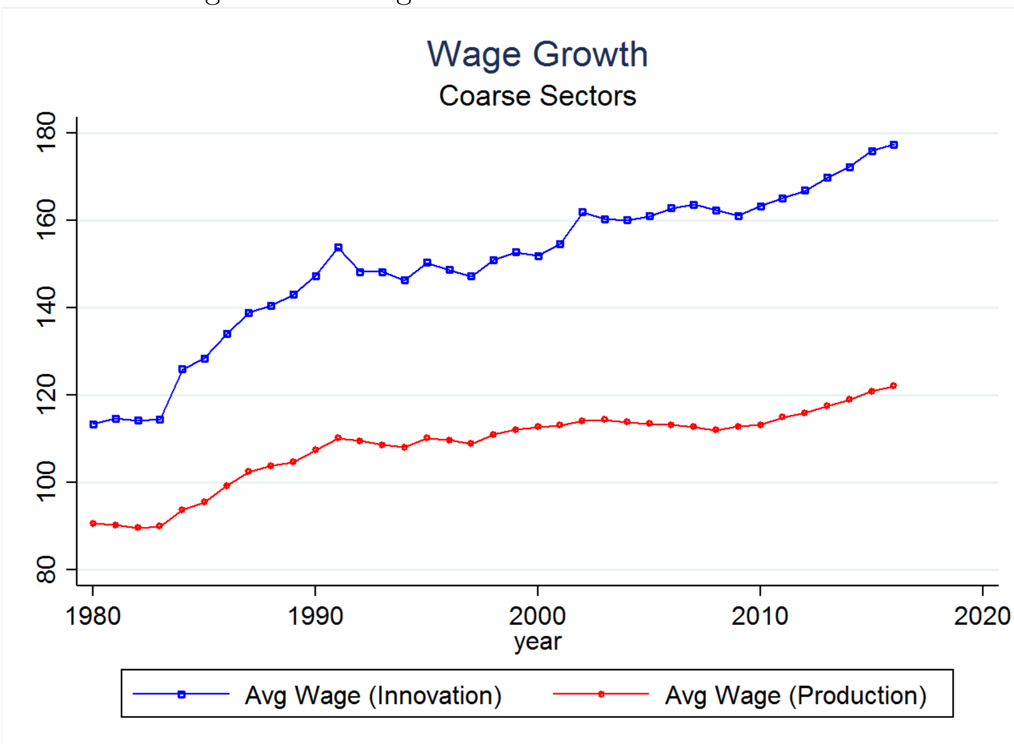
I plot average daily wages using the BHP data from the IAB over a long horizon. I plot the aggregate average wage, i.e. total labor income divided by total employment. As observed in a number of studies(Card, Heining, and Kline 2013; Doepke and Gaetani 2020) the skill premium does not respond as strong 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.

Figure A9: Wage Stagnation in Germany over the long run



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 A10, 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.

Figure A10: 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.

A.8.3 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 A1. Both a shift-share based measure of exporting following D. 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.[186]

Table A1: Barro Coefficient with Controls

	Controls in base period	$\hat{\beta}^{1986-1994}_{Barro}$		$\hat{\beta}^{1994-2006}_{Barro}$		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).

[186] D. 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.

Appendices to Chapter II

B.1 Simple Infinite Horizon Economy with Poisson Arrival of Type

Deriving equation (2.21**):**

I derive the relationship for a general utility function u

$$\begin{aligned}
V_{t_m} &= \max \mathbb{E}_{\hat{t}} \mathbb{E}_{\varphi} \left[\int_{t_m}^{\infty} exp\left(-\rho[s-t_m]\right) u\left(c_s\right) ds | T = \hat{t} \right] \\
&= \max \mathbb{E}_{\hat{t}} \left[\int_{t_m}^{\hat{t}} exp\left(-\rho[s-t_m]\right) u\left(c_s\right) ds + exp\left(-\rho[\hat{t}-t_m]\right) \mathbb{E}_{\varphi} V\left(\varphi, a_{\hat{t}}\right) \right] \\
&= \max \mathbb{E}_{T} \left[\int_{t_m}^{T} exp\left(-\rho[s-t_m]\right) u\left(c_s\right) ds + exp\left(-\rho[T-t_m]\right) \mathbb{E}_{\varphi} V\left(\varphi, a_T\right) \right] \\
&= \max \int_{t_m}^{\infty} \lambda exp\left(-\lambda[t-t_m]\right) \left[\int_{t_m}^{t} exp\left(-\rho[s-t_m]\right) u\left(c_s\right) ds + exp\left(-\rho[t-t_m]\right) \mathbb{E}_{\varphi} V\left(\varphi, a_t\right) \right] dt,
\end{aligned}$$

where the first line conditions on the arrival time. The second line splits up the integral, conditional on the arrival time, into the time before and after the agent learned about their type. Note that the second line also implicitly reflects the independence of the Poisson arrival process, and the agent's type. This allows me to simply compute the expectation over the type-space.

The next step is to change the order of integration. In order to do so, we need to keep track of the boundaries of integration. In this problem, we have $t > s > t_m$. Then, changing the order of integration means that we first integrate over t. In that case, the lower boundary is s, and there is no upper bound on the values that t can take. This gives the following solution

$$
\begin{aligned}
V_{t_m} &= \max \int_{t_m}^{\infty} \lambda exp\left(-\lambda[t-t_m]\right)\left[\int_{t_m}^{t} exp\left(-\rho[s-t_m]\right)u\left(c_s\right)ds + exp\left(-\rho[t-t_m]\right)\mathbb{E}_{\varphi}V\left(\varphi, a_t\right)\right]dt \\
&= \max \int_{t_m}^{\infty} exp\left(-\rho[s-t_m]\right)u\left(c_s\right)\left[\int_{s}^{\infty} \lambda exp\left(-\lambda[t-t_m]\right)dt\right]ds \\
&+ \int_{t_m}^{\infty} \lambda exp\left(-\lambda[t-t_m]\right)exp\left(-\rho[t-t_m]\right)\mathbb{E}_{\varphi}V\left(\varphi, a_t\right)dt \\
&= \max \int_{t_m}^{\infty} exp\left(-\left(\lambda+\rho\right)[s-t_m]\right)\left[u\left(c_s\right)+\lambda\mathbb{E}_{\varphi}\left[V\left(\varphi, a_s\right)\right]\right]ds.
\end{aligned}
$$

Proposition 1 − 4

To prove and understand the propositions I first derive the dynamics for households in the high growth regime. Next, I will show how this fits into the whole equilibrium and in particular into the migration decision. The dynamics are very similar to the standard neoclassical growth model, and a phase diagram analysis allows for a general characterization.

First, let's focus on the households in the high-growth regime. Recall the household Euler equation as well as the budget constraint that determine the dynamics in the high growth regime:

$$
\begin{aligned}
\frac{\dot{c}_s}{c_s} &= \frac{\lambda}{\eta}\left\{\mathbb{E}_{\varphi}\left[\left(\frac{c_s}{\varphi y_s + \left[g^*\left(\eta-1\right)+\rho\right]a_s}\right)^{\eta} - 1\right]\right\} + \frac{r^*-\rho}{\eta} \\
\dot{a}_s &= r^* a_s + y_s - c_s
\end{aligned}
$$

The first challenge is to obtain a system of differential equations that leads to a steady state. In the main part of the paper I call this "pseudo" steady state. The reason is that households won't reside in this equilibrium forever, but are pulled out randomly according to the Poisson arrival process of their type. Nonetheless, I can use the steady state analysis in combination with a phase diagram to understand the equilibrium dynamics of households in the high-growth regime. The only difference is that for the actual solution of the model, I would need to send households onto the convergence process toward the pseudo steady state and then pull them out randomly consistent with the Poisson process.

Since I have growth in my model, $\dot{c}_s = 0$ is not going to be a solution. In order to obtain a stationary system, I define a new system where consumption and assets are normalized by income. This choices is motivated by Carroll (1997) who shows that asset-to-income ratios are stationary in a particular type of precautionary savings models.[187] His insight generalizes to the framework at hand as well. Let $X_s := \frac{c_s}{y_s}$ denote the consumption-to-income ratio, and let $Z_s := \frac{a_s}{y_s}$ denote the asset-to-income ratio. This leads to the following differential equations

$$\frac{\dot{X}_s}{X_s} = -\left(g - g^*\right) + \frac{\lambda}{\eta}\left\{\mathbb{E}_\varphi\left[\left(\frac{X_s}{\left[g^*\left(\eta - 1\right) + \rho\right] Z_s + \varphi}\right)^\eta\right] - 1\right\} \tag{B1}$$

$$\frac{\dot{Z}_s}{Z_s} = -\left(g - r^*\right) + \frac{1}{Z_s}\left(1 - X_s\right). \tag{B2}$$

The loci for equation $(B1)$ and $(B2)$ are given by

$$X_{**} = \left\{\frac{\left(g - g^*\right)}{\lambda}\eta + 1\right\}^{\frac{1}{\eta}}\left\{\mathbb{E}_\varphi\left[\left(\frac{1}{\varphi + \left[g^*\left(\eta - 1\right) + \rho\right] Z_{**}}\right)^\eta\right]\right\}^{-\frac{1}{\eta}} \tag{B3}$$

$$Z_{**} = \frac{1 - X_{**}}{g - r^*} \tag{B4}$$

where $**$ denotes the steady state values. Now there are two equilibria that could emerge: a situation with precautionary savings and capital outflows, or an equilibrium with consumption smoothing and capital inflows. I am going to focus on the equilibrium with precautionary savings, which means $Z_{**} > 0$.

In that case is easy to show that $(B3)$ is increasing in Z_{**} and $(B4)$ is strictly decreasing in X_{**} where I assume $g > r^*$, a mild assumption in the "miracle growth" context of this paper where $g = 7\%$. This leads to a unique steady state solution (if it exists). For existence, we need the intercept of $(B3)$ to be below the intercept of $(B4)$ which is satisfied as long as $\left(\frac{g-g^*}{\lambda}\eta + 1\right) < \mathbb{E}_\varphi\left[\left(\frac{1}{\varphi}\right)^\eta\right]$. I assume this inequality is satisfied in order for the economy to exhibit capital outflows.

A short comment is in order. Whether this inequality holds or not depends on the consumption smoothing force embodied in the term $\frac{g-g^*}{\lambda}\eta + 1$ on the one hand, and the precautionary motive reflected in $\mathbb{E}_\varphi\left[\left(\frac{1}{\varphi}\right)^\eta\right]$ on the other. First, for the standard case of a log-normal type distribution, $\log(\varphi) \sim N(-\frac{\sigma^2}{2}, \sigma^2)$, it is well known that $\mathbb{E}_\varphi\left[\left(\frac{1}{\varphi}\right)^\eta\right] = \exp\left(\frac{\sigma^2}{2}\eta\left[1 + \eta\right]\right)$. For $\eta = 2$, any σ^2 above .37 will suffice. For $\eta = 4$, any σ^2 above 0.16 is sufficient to dominate the consumption smoothing motive.

However, the argument extends beyond the log-normal case. What is needed is the notion

[187] His results operate in a discrete time framework where shocks to permanent income are modeled as random walk with the error following a log normal distribution

of a mean-preserving spread that can be applied in the context at hand. Note that Mas-Colell, Whinston, Green, et al. (1995) define a mean-preserving spread of a random variable X, based on the work of Diamond and Stiglitz (1974), in the following way: $X' = X + e$ s.t. $\mathbb{E}[e] = 0$. This does not work in the context at hand because I need to ensure that the type φ is always greater zero. So if one wanted to define a mean preserving spread, one would have to do something like $\varphi' = \varphi + e$ but this implies that $e \geq -\varphi$ which automatically induces statistical dependence among the random variables. When dealing with random variables that have to satisfy $\mathbb{E}[\varphi] = 1$, I propose the following version of a mean preserving spread $\varphi' = \varphi\epsilon$ with $\mathbb{E}[\epsilon\varphi] = \mathbb{E}[\epsilon]\mathbb{E}[\varphi] = 1$. Note that I have put zero distributional assumptions on neither φ nor ϵ other than that they have to be unity in expectation. Now define $\varphi_k = \Pi_{j=1}^{k}\epsilon_j$, where the ϵ's are iid draws.

What I want to show is that for a sufficient amount of uncertainty about a household's type there exist a solution that sustains capital outflows, beyond log-normality. To see that this is the case, consider $\mathbb{E}[\varphi_k^{-\eta}]$. A higher k here is a mean-preserving spread of the type distribution. Now note that

$$
\begin{aligned}
\mathbb{P}\left(\varphi_K > y\right) &= \mathbb{P}\left(\Pi_{j=1}^{K}\epsilon_j < y\right) \\
&= \mathbb{P}\left(\sum_{j=1}^{K}\log(\epsilon_j) < \log(y)\right) \\
&= \mathbb{P}\left(\frac{\sum_{j=1}^{K}\log(\epsilon_j)}{K} < \frac{\log(y)}{K}\right) \\
&= \mathbb{P}\left(\hat{\mathbb{E}}_K\log(\epsilon_j) < \frac{\log(y)}{K}\right) \\
&= \mathbb{P}\left(\hat{\mathbb{E}}_K\log(\epsilon_j) < \frac{\log(y)}{K}\right)
\end{aligned}
$$

where $y < 1$, and $\hat{\mathbb{E}}$ denotes the sample average. Note that because of Jensen's inequality it must be that $\mathbb{E}[\log(\epsilon)] < \log(\mathbb{E}[\epsilon]) = 0$. Without loss of generality, assume $\mathbb{E}[\log(\epsilon)] = \frac{\log(y)}{M}$

for some $M > 0$. Now subtract the expectation of $\log(\epsilon)$

$$\begin{aligned}
\mathbb{P}\left(\varphi_K < y\right) &= \mathbb{P}\left(\hat{\mathbb{E}}_K \log(\epsilon_j) < \frac{\log(y)}{K}\right) \\
&= \mathbb{P}\left(\hat{\mathbb{E}}_K \log(\epsilon_j) - \mathbb{E}[\log(\epsilon)] < \frac{\log(y)}{K} - \mathbb{E}[\log(\epsilon)]\right) \\
&= \mathbb{P}\left(\hat{\mathbb{E}}_K \log(\epsilon_j) - \mathbb{E}[\log(\epsilon)] < -\log(y)\left(\frac{1}{M} - \frac{1}{K}\right)\right) \\
&> \mathbb{P}\left(\left|\hat{\mathbb{E}}_K \log(\epsilon_j) - \mathbb{E}[\log(\epsilon)]\right| < -\log(y)\left(\frac{1}{M} - \frac{1}{K}\right)\right)
\end{aligned}$$

Where the last inequality follows from the fact that $\{X : X < B\} = \{X : X \leq -B\} \cup \{X : -B < X < B\} \supset \{X : B < X < -B\} = \{X : |X| < B\}$.

$$\begin{aligned}
\mathbb{P}\left(\varphi_K < y\right) &> \mathbb{P}\left(\left|\hat{\mathbb{E}}_K \log(\epsilon_j) - \mathbb{E}[\log(\epsilon)]\right| < -\log(y)\left(\frac{1}{M} - \frac{1}{K}\right)\right) \\
&= 1 - \mathbb{P}\left(\left|\hat{\mathbb{E}}_K \log(\epsilon_j) - \mathbb{E}[\log(\epsilon)]\right| \geq -\log(y)\left(\frac{1}{M} - \frac{1}{K}\right)\right) \\
&\geq 1 - \frac{\mathbb{E}(\epsilon)}{K\left[-\log(y)\left(\frac{1}{M} - \frac{1}{K}\right)\right]^2}
\end{aligned}$$

where the last inequality follows from Chebyshev's inequality. Put differently, for large enough K the probability of φ being very small converges to one. It then follows that for large enough K, we have

$$\mathbb{E}\left[\left(\frac{1}{\varphi}\right)^{\eta}\right] > \mathbb{P}(\varphi < x)\left(\frac{1}{x}\right)^{\eta}.$$

But I can make $\mathbb{P}(\varphi < x)\left(\frac{1}{x}\right)^{\eta}$ arbitrarily large as I am increasing K. Intuitively, the mean-preserving spread that I introduced shifts more and more mass on a very small x, but a small x lets the expression explode. Given the results so far, it is easy to show that for any y and for any number $L \in \{1, 2, ...\}$, I can find a K such that $\mathbb{E}\left[\left(\frac{1}{\varphi_K}\right)^{\eta}\right] > L$. This concludes the treatment of the general case. Note that nowhere did I assume anything about the precise shape and support of the distribution. The purpose of this derivation was to show that the results do not rely on the particularities of the log normal distribution, or on the continuity of the type space. After obtaining this general result, I will focus on the case of log-normally distributed types since this the main exercise in the paper.

Next, I need to sign the derivative of the differential equations to draw the phase diagram, evaluated at the locus

$$\frac{d\frac{\dot{X}_s}{X_s}}{dZ}|_{X_{**}} < 0 \tag{B5}$$

$$\frac{d\frac{\dot{Z}_s}{Z_s}}{dX}|_{Z_{**}} < 0. \tag{B6}$$

Figure B1 displays the phase diagram. The dashed line represents the stable arm which is the unique trajectory of the system. Consumption is a control variable and jumps up when the household enters the high-growth regime so as to end up on the stable arm. From then on, the consumption-to-income ratio and the asset-to-income ratio increases till the steady state is reached. Consequently, consumption and assets grow at a rate higher than income, a standard result of models of precautionary savings.

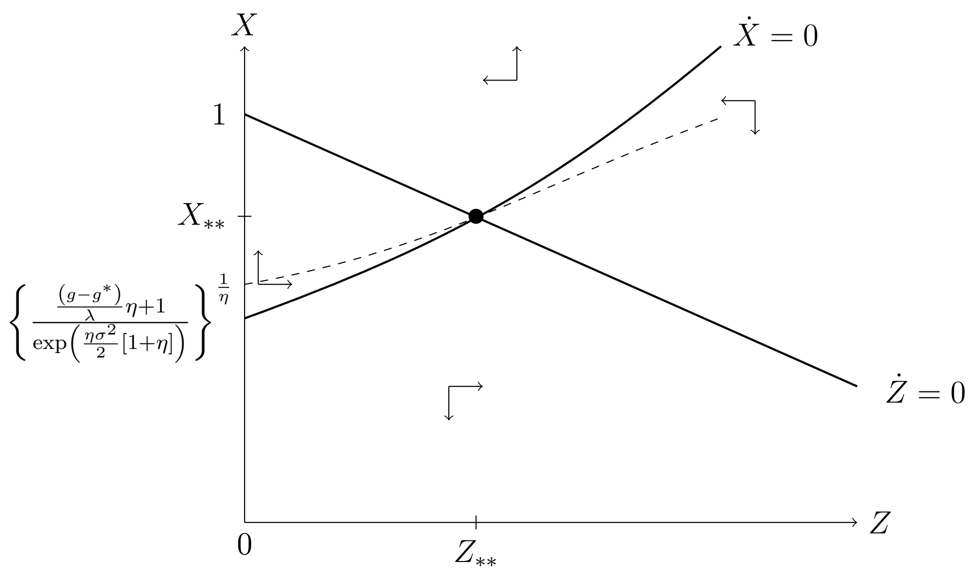


Figure B1: Pseudo steady state analysis of asset-to-income and consumption-to-income ratio in high-growth regime

Law of motion out of agriculture in general case

Next, I show that I can pin down the dynamics of the agricultural share even though I cannot solve for the transitional dynamics in the high growth regime in closed form. Once I establish this, we can conclude that an equilibrium exists, that is well behaved, in which the interaction of structural change, inequality, and growth generates capital outflows despite convergence growth.

From before, everything is captured in the asset-to-income and consumption-to-income ratio. Those dynamics are always the same, hence consumption and assets for households entering in the high growth regime at a later point in time are scaled up by a factor $\exp(g^* t)$ but otherwise identical.

To see this, note that the pseudo steady state analysis for an individual household always start at t^i_m, that is when the household leaves the country side. To see this, note that $\left\{\frac{c_t}{y_t}, \frac{a_t}{y_t}\right\}$ for households in the high-growth regime is only determined by the time spend in the high-growth regime $t - t_m$. This follows from the fact that the two (normalized) first order conditions do not depend on any variable that is a function of time t, the only thing one needs to keep track of here is $t - t_m = s$. It follows that the asset-to-income and consumption-to-income ratio are independent of calendar time, and only depend on the time spend in the high-growth regime. This property allows me to rewrite the consumption and asset profile of any agent in the high-growth regime as follows as follows:

$$\begin{aligned}
c\left(t_m, t\right) = & \qquad c_0\left(t - t_m\right) exp(g^* t_m) \\
a\left(t_m, t\right) = & \qquad a_0\left(t - t_m\right) exp\left(g^* t_m\right) \\
y\left(t_m, t\right) = & \qquad exp\left(g\left[t - t_m\right]\right) exp\left(g^* t_m\right) \\
= & \qquad y_0\left(t - t_m\right) exp\left(g^* t_m\right)
\end{aligned}$$

Proof:

$$X\left(s\right) = \frac{c\left(t, t_m\right)}{y\left(t, t_m\right)} = \frac{c\left(t+k, t_m+k\right)}{y\left(t+k, t_m+k\right)} \qquad \forall k \in \left[-t_m, \infty\right)$$

with a slight abuse of notation where $t_m + k$ represent the consumption profile of an agent that entered the city k units of time later. Since the income process is exogenous, and every agent starts at the level $y\left(t_m\right) = A_t$, we can rewrite the equality

$$\frac{c\left(t, t_m\right)}{y\left(t - t_m, 0\right) A_{t_m}} = \frac{c\left(t - t_m, 0\right)}{y\left(t - t_m, 0\right)}$$

This implies that $c\left(t, t_m\right) = A_{t_m} c\left(t - t_m, 0\right)$ or short $c_0\left(t - t_m\right) \exp(g^* t_m)$. The case for assets is analogous.

Of course, actual consumption and asset profiles need to be rescaled by income to obtain observed household asset holdings and consumption. The feature of the model that consumption and assets, after accounting for the time spend in the high-growth regime, can simply be scaled up by A_{t_m} is key for tractability. As I show next, this allows me to derive the law of motion of workers out of the urban sector in closed form.

When considering the indifference condition that household on the country side consider before moving to the city, we can simplify this problem as follows. The value function reads

$$V_{t_m} = \max_{c_s} \mathbb{E}_{\varphi,T} \int_{t_m}^{\infty} exp\left(-\rho[s-t_m]\right) \frac{c(\varphi,s,T)^{1-\eta}}{1-\eta} ds. \tag{B7}$$

There is no simple solution for this expression in closed form. But, we can use the fact that all choices simply scale in income, a simplification that obtains from the CRRA utility function, in combination with the multiplicative type shock. Then we rescale consumption by $\exp\left(g^* t_m\right)$. This allows us to rewrite the expression as follows

$$V_{t_m} = \exp\left(-g^*[\eta-1]t_m\right) \max_{c_s} \mathbb{E}_{\gamma,T} \int_{t_m}^{\infty} exp\left(-\rho[s-t_m]\right) \frac{c_0(\varphi,s-t_m,T-t_m)^{1-\eta}}{1-\eta} ds \tag{B8}$$

$$= \exp\left(-g^*[\eta-1]t_m\right) V_0 \tag{B9}$$

This delivers the important result that

$$\dot{V_{t_m}} = -\left(g^*[\eta-1]\right) V_{t_m} \tag{B10}$$

Two comments are in order. First, B10 only makes sense when V_0 is well defined. A sufficient condition for this to be the case is $\eta > 1 - \frac{\lambda+\rho}{g}$ and we need a finite expectation with respect to the type draw as well. Second, everything works out with log utility as well.[188] For finite utility in the low-growth regime, as well as in the industrialized world, we also need $\rho > [1-\eta]g^*$ which is assumed throughout the paper.

To finally pin down the law of motion out of the urban sector, we need to plug B10 into the indifference condition (2.13) that leaves a rural hosuehold indifferent between migrating and staying on the countryside. Conveniently, the time derivative is independent of V_0. Intuitively, migrants always face the same type of convergence process, the only difference is

[188] Notes are available upon request.

that the overall wage rate in the urban sector keeps growing.

$$\begin{aligned}
\frac{(w_t^r)^{1-\eta}}{1-\eta} &= \tau^{\eta-1}\left(\rho V_t - \dot{V}_t\right) \\
\frac{(w_t^r)^{1-\eta}}{1-\eta} &= \tau^{\eta-1}\left(\rho V_t - g^*\left[1-\eta\right] V_t\right) \\
\frac{(w_t^r)^{1-\eta}}{1-\eta} &= \tau^{\eta-1}\exp\left(g^*\left[1-\eta\right] t\right) V_0\left(\rho - g^*\left[1-\eta\right]\right) \\
w_t^r &= \exp\left(g^* t\right)\frac{\left[\left(1-\eta\right) V_0\left(\rho + g^*\left[\eta-1\right]\right)\right]^{\frac{1}{1-\eta}}}{\tau}.
\end{aligned}$$

Now, I only need to make use of the fact that the compensation on the country side is given by $w_t^r = (L_t^r)^{-[1-\alpha]}$. Using this yields the share of households on the country side

$$L_t^r = \{\tau\}^{\frac{1}{(1-\alpha)}}\exp\left(-\frac{g^*}{1-\alpha}t\right)\{V_0\left[1-\eta\right]\left[\rho + g^*\left[\eta-1\right]\right]\}^{\frac{1}{(1-\alpha)(\eta-1)}}$$

And even though there is no closed form solution for V_0, to the extent that we observe the initial share of workers on the country side L_0^r, we obtain a structural relationship based on observables

$$L_t^r = L_0^r \exp\left(-\frac{g^*}{1-\alpha}t\right).$$

I proceed by deriving the law of motion of households in the high and low growth regime, respectively. The convenient stochastic process delivers simple solutions. The labor resource constraint together with the Poisson process of drawing your type allows me to characterize the change in the different types L_t^u, $M_{t,0}$, $M_{t,1}$ as laws of motion. First, note that from (2.17) we get

$$\frac{dL_t^u}{dt} = -\frac{dL_t^r}{dt}. \tag{B11}$$

Moreover, the agents in the city that have drawn their type $M_{t,1}$ and the ones that haven't $M_{t,0}$ add up to L_t^u and thus

$$\begin{aligned}
\frac{dL_t^u}{dt} &= \frac{dM_{t,0}}{dt} + \frac{dM_{t,1}}{dt} \\
&= \frac{dM_{t,0}}{dt} + \lambda M_{t,0},
\end{aligned}$$

where the second line follows from the Poisson process, and the fact that there is a continuum of agents $M_{0,t}$. Rearranging yields the change in the fraction of agents that grow at the high rate g_h,

$$\frac{dL_t^u}{dt} - \lambda M_{t,0} = \frac{dM_{t,0}}{dt}. \tag{B12}$$

Whether this term is positive or negative depends on the relative strength of migration (inflow) and Poisson arrival process (outflow), and in the limit, $M_{\infty,1} = 1$. The change in the mass of agents that reside in the low-growth regime in the city is simply given by

$$dM_{t,1} = \lambda dt M_{t,0}. \tag{B13}$$

using the Poisson arrival.[189]

The fraction of agents in the city at time zero is $1 - L_0^r$. I assume that they all have to still draw their type and start growing at the high growth rate.[190] This would be agents in the set $M_{0,0}$. Using (B12) together with (2.28) yields

$$\dot{M}_{t,0} + \lambda M_{t,0} = \frac{g^*}{1-\alpha} L_0^r \exp\left(-\frac{g^*}{1-\alpha}t\right). \tag{B14}$$

Using $\exp(\lambda t)$ as integrating factor, this differential equation can be solved and yields

$$M_{t,0} = M_{0,0} exp(-\lambda t) + \frac{g^* L_0^r}{\lambda(1-\alpha) - g^*}\left[\exp(-\frac{g^*}{1-\alpha}t) - \exp(-\lambda t)\right]. \tag{B15}$$

Similarly, the mass of households $M_{t,1}$ can be obtained

$$\begin{aligned}
M_{t,1} &= \lambda \int_0^t M_{s,0} ds + M_{0,1} \\
&= M_{0,0}\left[1 - \exp(-\lambda t)\right] && \text{(B16)} \\
&+ \left(\frac{g^* L_0^r}{\lambda(1-\alpha) - g^*}\right)\left\{\frac{\lambda(1-\alpha)}{g^*}\left[1 - \exp(-\frac{g^*}{1-\alpha}t)\right] - \left[1 - \exp(-\lambda t)\right]\right\} && \text{(B17)} \\
&+ M_{0,1} && \text{(B18)}
\end{aligned}$$

The changing shares of agents that don't know their type will be key to generate hump-shaped aggregate saving rates. Of course, in order to aggregate things up onto the macro level and

[189]There is a law of large numbers operating in the background here.
[190]I can relax that assumption very easily and will do so in a later section.

get the aggregate savings in the economy right, we need to use appropriate weights that we attach to each household. Those weights will be based on the income of the household, which is why I study the dynamics of the income distribution next.

Proof of Proposition 2.4.2:

Using the phase diagram, proof of proposition one becomes straightforward. First, note that by Jensen's inequality consumption growth of the equilibrium with degenerate inequality distribution is a lower bound for consumption growth in the model with inequality

$$\mathbb{E}_{\varphi}\left[\left(\frac{1}{\rho a_s + \varphi y_s}\right)^{\eta}\right] > \left(\frac{1}{\rho a_s + \mathbb{E}_{\varphi}\varphi y_s}\right)^{\eta} = \left(\frac{1}{\rho a_s + y_s}\right)$$

The first part of the proposition, however, claims that consumption growth is strictly higher for agents in the high-growth regime relative to the industrialized world. To see this, consider the same phase diagram as before, except now $\sigma = 0$ which in turn implies that the X and Z loci intersect somewhere where $Z_{**} < 0$ and $X_{**} > 1$. Also note that along the transition path, $X_s > 1$ and $Z_s < 0$.

Next, by contradiction, suppose that

$$\frac{\dot{c}}{c} = \frac{\lambda}{\eta}\left[\left(\frac{c_s}{y_s + [\rho + [\eta - 1]g^*]a_s}\right)^{\eta} - 1\right] + g^* < g^*$$

This implies that $X_s < 1 + [\rho + [\eta - 1]g^*]Z_s$. But since $Z_s < 0$ and $X_s > 1$, together with $\rho + [\eta - 1]g^* > 0$, we arrived at a contradiction.

Proof of Proposition 2.4.2:

As argued before, the phase diagram analysis shows that the steady state as well as the transition path display $X_s > 1$. Clearly, if the consumption-to-income ratio is greater unity at all times in the high-growth regime there must be capital inflows to finance the gap between output and consumption.

Proof of Proposition 2.4.2:

Note that the proof for proposition 2.4.2 follows simply from the fact that proposition 2.4.2 and 2.4.2 hold for any positive and finite λ.

Migration Decision:

Given that agents are on the conjectured equilibrium, I can solve the arbitrage condition that pins down the flow agents into the city. Recall

$$log(w_t^r) \overset{!}{=} \rho V_t - \dot{V}_t.$$

Given the conjectured equilibrium, I obtain a closed form solution for the value function as follows.

$$
\begin{aligned}
V_{t,0} =& \int_t^{\infty} exp(-(\lambda+\rho)[s-t]) \left\{ log(y_s) + \lambda \mathbb{E}_{\varphi} V_{s,1}(\varphi, a_s) \right\} ds \\
=& \int_t^{\infty} exp(-(\lambda+\rho)[s-t]) \left\{ [log(y_t) + g[s-t]] + \lambda \left[\frac{log(y_s)}{\rho} + \mathbb{E}_{\varphi} \left[\frac{log(\varphi)}{\rho} \right] + \frac{g^*}{\rho^2} \right] \right\} ds \\
=& \int_t^{\infty} exp(-(\lambda+\rho)[s-t]) \left\{ [log(y_t) + g[s-t]] + \lambda \left[\frac{log(y_s)}{\rho} + \mathbb{E}_{\varphi} \left[\frac{log(\varphi)}{\rho} \right] + \frac{g^*}{\rho^2} \right] \right\} ds \\
=& \int_t^{\infty} exp(-(\lambda+\rho)[s-t]) \left\{ [log(y_t) + g[s-t]] + \lambda \left[\frac{log(y_s)}{\rho} - \frac{\sigma^2}{2\rho} + \frac{g^*}{\rho^2} \right] \right\} ds \\
=& \int_t^{\infty} exp(-(\lambda+\rho)[s-t]) \left\{ [log(y_t) + g[s-t]] (1 + \frac{\lambda}{\rho}) + \lambda \left[\frac{g^*}{\rho^2} - \frac{\sigma^2}{2\rho} \right] \right\} ds \\
=& \left[\frac{log(y_t)}{\lambda+\rho} + \frac{g}{(\lambda+\rho)^2} \right] (1 + \frac{\lambda}{\rho}) + \frac{\lambda}{\lambda+\rho} \left[\frac{g^*}{\rho^2} - \frac{\sigma^2}{2\rho} \right] \\
=& \left[\frac{log(y_t)}{\lambda+\rho} \right] (1 + \frac{\lambda}{\rho}) + \frac{1}{\lambda+\rho} \frac{1}{\rho^2} [\lambda g^* + \rho g] - \frac{\lambda}{\lambda+\rho} \frac{\sigma^2}{2\rho} \\
=& \frac{log(y_t)}{\rho} + \frac{1}{\lambda+\rho} \frac{1}{\rho^2} [\lambda g^* + \rho g] - \frac{\lambda}{\lambda+\rho} \frac{\sigma^2}{2\rho}.
\end{aligned}
$$

Now I can differentiate this expression with respect to t, and plug it back into the arbitrage condition that keeps agents on the country side indifferent between staying and moving, hence

$$
\begin{aligned}
log(w_t^r) &= \rho \left(\frac{log(y_t)}{\rho} + \frac{1}{\lambda+\rho} \frac{1}{\rho^2} [\lambda g^* + \rho g] - \frac{\lambda}{\lambda+\rho} \frac{\sigma^2}{2\rho} \right) - \dot{V}_t \\
&= log(y_t) + \frac{1}{\lambda+\rho} \frac{1}{\rho} [\lambda g^* + \rho g] - \frac{\lambda}{\lambda+\rho} \frac{\sigma^2}{2} - \dot{V}_t \\
&= log(y_t) + \frac{1}{\lambda+\rho} \frac{1}{\rho} [\lambda g^* + \rho g] - \frac{\lambda}{\lambda+\rho} \frac{\sigma^2}{2} - \frac{g^*}{\rho} \\
&= log(A_t) + \frac{1}{\lambda+\rho} (g - g^*) - \frac{\lambda}{\lambda+\rho} \frac{\sigma^2}{2}.
\end{aligned}
$$

B.1.1 agricultural productivity gap/urban rural wage gap

There is a link between the model and the literature on the urban rural wage gap (Harris and Todaro 1970; Young 2013; Lagakos and Waugh 2013; Hicks et al. 2017) and the agricultural productivity gap (Caselli 2005; Restuccia, D. T. Yang, and X. Zhu 2008; Gollin, Lagakos, and Waugh 2013). In the model, ex post inequality in the city can drive a wedge between urban and rural wages, and reduce the share of people working in urban areas relative to a world with complete markets. To see this, I solve a version of the model in log utility where the precautionary savings and the consumption smoothing motive exactly offset each other. Then, I get a closed form solution of the value function, and can pin down the rural wage as a function of the state of the technology in the city, as well as convergence growth and inequality, σ^2. This relationship is captured in equation (B19)

$$\log(w_t^r) = \log(A_t) + \frac{1}{\lambda+\rho}\left(g - g^*\right) - \frac{\lambda}{\lambda+\rho}\frac{\sigma^2}{2} - \log\left(\tau\right). \tag{B19}$$

In my model there is no selection on skill (ex ante everyone is the same!) but an econometrician that would compare labor productivity approximated by average log wages in a cross section of workers may conclude that there is an agricultural productivity gap, while the actual reason is a compensating risk differential. Wage growth in the rural region will be identical to wage growth in the city, and since there is potential for convergence growth in the city in contrast to the country side, rural workers are compensated for that by the term $\frac{1}{\lambda+\rho}\left(g - g^*\right)$. Note that without the wedge, assuming that convergence growth and precautionary savings cancel, it can be shown that the urban wage at time t_m would be smaller than the rural wage at t_m. This happens because rural households need to be compensated for the lack of high-growth opportunity in equilibrium. Given log utility, this force dominates the risk adjustment, that pushed down the rural wage. This is why we need $\tau > 1$, i.e. there needs to be an additional wedge to migration.

B.2 Derivation of income distribution

From the main text we obtain the following equation

$$\log(y(t, t_{m_i}, T_i, \varphi)) = 1(T_i \geq t \geq t_{m_i})[g - g^*][t - t_{m_i}] + 1(T_i < t)\{[g - g^*][T_i - t_{m_i}] + \log\varphi\} \tag{B20}$$

This makes clear that the income of each agent is pinned down by the quadruple $\{t, t_{m_i}, T_i, \varphi\}$. In order to compute the density of income, we need to keep track of how much time each household spent in the high growth regime, $T_i - t_{m_i}$. It turns out to be convenient to split the households into two groups, the ones that are in the high growth regime, relative to the ones that are in the absorbing state of low growth. In all this, keep in mind that there is a mass point of agents that "entered" the city at time zero. Some of those are people who have already been there before before the "beginning of time". Some jump over at time zero to ensure the migration arbitrage condition holds.

I start by computing the conditional density in the high-growth regime, which is slightly easier. Also note that I compute the conditional densities, relative to the whole population. When we take a final step to map those densities into the variance of the log of income into the city, we need to make sure to normalize appropriately so that the conditional probabilities over the city dwellers add up to unity. That means we need to normalized the densities by $M_0 + M_1$. At every point in time there is a cohort of migrants that enter at $t_{m_i} = t$. The size of the cohort is given by the flow of workers out of agricultural activity. Next, note that at time t there is only a fraction of the cohort left because of the Poisson arrival of drawing your type. As a consequence, the CDF reads

$$F(z|t) = \frac{M_{0,0}exp(-\lambda t) + \int_0^z \frac{g^*}{1-\alpha}\Lambda exp(-\frac{g^* s}{1-\alpha})exp(-\lambda(t-s))ds}{M_{0,t}} D(z \in (0,t])$$

Proof:

$$\begin{aligned}
P\left(t_{m_i} \leq z | T_i > t\right) =& P\left(t_{m_i} < z | T_i > t\right) && \text{(B21)}\\
=& \frac{P\left(t_{m_i} < z \cap T_i > t\right)}{P\left(T_i > t\right)} && \text{(B22)}\\
=& \frac{\mathbb{E}_{t_{m_i}}\left[P\left(T_i > t | t_{m_i}\right)\right] D\left(t_{m_i} < z\right)}{P\left(T_i > t\right)} && \text{(B23)}\\
=& \frac{\mathbb{E}_{t_{m_i}}\left[P\left(T_i - t_{m_i} > t - t_{m_i} | t_{m_i}\right)\right] D\left(t_{m_i} < z\right)}{P\left(T_i > t\right)} && \text{(B24)}\\
=& \frac{\int_0^z exp(-\lambda(t-s))dF\left(s\right)}{M_{0,t}} && \text{(B25)}\\
=& \frac{M_{0,0}\exp\left(-\lambda t\right) + \int_0^z exp(-\lambda(t-s))\frac{g^*}{1-\alpha}\Lambda exp(-\frac{g^* s}{1-\alpha})ds}{M_{0,t}} && \text{(B26)}
\end{aligned}$$

where D is an indicator function, and $f(s)ds = \frac{g^*}{1-\alpha}\Lambda \exp(-\frac{g^* t}{1-\alpha})$ is the size of the cohort entering the city at s, and $\Lambda = L_0$ keeps track of the initial share of agricultural workers. Thus, we have derived the distribution of t_{m_i} for households in the high growth regime. Now we can simply use this distribution to compute conditional moments – the reason is that given t, t_{m_i} is the only variable that impacts relative household income and inequality WITHIN the group of high-growth households. When we do that, the only thing to keep in mind is that there is a mass point at zero, i.e. $P\left(t_{m_i} = 0 | T_i > t\right) = \frac{M_{0,t}\exp(-\lambda t)}{M_{1,t}}$. As mentioned in the main text, an implicit assumption is that a law of large numbers operates within each cohort. A non-trivial assumption that is usually taken for granted in applied economic models (Arkolakis 2010; Luttmer 2007). Next, I show how to derive the distribution of the agents in the low-growth regime. This is harder because income depends on two random variables (ignoring the type draw here because it is easy to handle), namely the time of migration t_{m_i} and the time of leaving the high growth regime T_i.

At every point in time t, there is a distribution over the time of migration from zero to t of households in the high growth regime. A random fraction λ is drawn from this distribution at every instant. Again, using some law of large numbers in the background, we can conclude that the fraction of people that are pushed into the set $M_{1,t}$, which migrated at time t_m before some threshold k reads

$$\begin{aligned}
P(i \in \dot{M}_{1,t} : t_m(i) \leq k|t) =& P(i \in M_{0,t} : t_m(i) < k|t) \\
=& F\left(k|t\right) \\
=& \frac{M_{0,0}exp(-\lambda t) + \int_0^k \frac{g^*}{1-\alpha} \Lambda exp(-\frac{g^* s}{1-\alpha}) exp(-\lambda(t-s))ds}{M_{0,t}}
\end{aligned}$$

$$\frac{\lambda \int_0^t M_{0,x} F(k|x) dx}{\lambda \int_0^t M_{0,x} dx} = \frac{\lambda \int_0^t M_{0,x} \frac{M_{0,0}exp(-\lambda x) + \int_0^z \frac{g^*}{1-\alpha} \Lambda exp(-\frac{g^* s}{1-\alpha}) exp(-\lambda(x-s))ds}{M_{0,x}} dx}{M_{1,t}}$$

We proceed as follows:

$$\begin{aligned}
\frac{P\left(t_{m_i} < k \cap T_i < t\right)}{P\left(T_i < t\right)} =& \frac{\mathbb{E}_{T_i} P\left(t_{m_i} < k | T_i = z\right) D\left(T_i < t\right)}{M_{1,t}} \\
=& \frac{\int_0^t F\left(k|x\right) M_{0,x} \lambda dx}{M_{1,t}} \\
=& \frac{\int_0^t F\left(k|x\right) M_{0,x} \lambda dx}{M_{1,t}}
\end{aligned}$$

This expression essentially asks: how many people entered before time k, and how many of them are left, since they are drawn out at rate λ. Importantly, and something that I messed up initially, we need to distinguish between two cases. There are outflows of M_0 before time k, and there are outflows of M_0 after time k. Any outflow before time k means that all the agents that flew into the M_1 pool left the country side at t_m below k. Therefore, $F\left(k|x\right) = 1\ \forall x < k$. For the outflows that happen after k, there is a distribution of types, some of which entered early and some of which entered late, in particular after t_m. This leads to the following expression

$$
\begin{aligned}
P\left(t_{m_i} \leq k|i \in M_{1,t}\right) =& \frac{\lambda \int_0^k M_{0,x} dx}{\lambda \int_0^t M_{0,x} dx} \\
&+ \frac{\lambda \int_k^t M_{0,x} \frac{M_{0,0}\exp(-\lambda x) + \int_0^k \frac{g^*}{1-\alpha}\Lambda \exp(-\frac{g^* s}{1-\alpha})\exp(-\lambda(x-s))ds}{M_{0,x}} dx}{\lambda \int_0^t M_{0,x} dx} \\
=& \frac{M_{1,k}}{M_{1,t}} \\
&+ \frac{\lambda \int_k^t M_{0,0}\exp(-\lambda x)dx + \lambda \int_k^t \int_0^k \frac{g^*}{1-\alpha}\Lambda \exp(-\frac{g^* s}{1-\alpha})\exp(-\lambda(x-s))dsdx}{M_{1,t}} \\
=& \frac{M_{1,k}}{M_{1,t}} \\
&+ \frac{M_{0,0}\exp(-\lambda t)\left(exp(\lambda\left[t-k\right])-1\right)}{M_{1,t}} \\
&+ \frac{\lambda \frac{g^*}{1-\alpha}\Lambda \int_k^t \exp\left(-\lambda x\right) \int_0^k \exp\left(\left[\frac{(1-\alpha)\lambda - g^*}{1-\alpha}\right] s\right) dsdx}{M_{1,t}} \\
=& \frac{M_{1,k} + M_{0,0}\exp(-\lambda t)\left(exp(\lambda\left[t-k\right])-1\right)}{M_1, t} \\
&+ \frac{\lambda \frac{g^*}{1-\alpha}\Lambda \int_k^t \exp\left(-\lambda x\right) \frac{1-\alpha}{(1-\alpha)\lambda - g^*}\left[\exp\left(\left[\frac{(1-\alpha)\lambda - g^*}{1-\alpha}\right] k\right) - 1\right] dx}{M_{1,t}} \\
=& \frac{M_{1,k}}{M_{1,t}} \\
&+ \frac{\left(\exp\left(-k\lambda\right) - \exp\left(-\lambda t\right)\right)\left\{M_{0,0} + \frac{\Lambda g^*}{g^* - (1-\alpha)\lambda}\left[1 - \exp\left(\left[-\frac{g^* - (1-\alpha)\lambda}{1-\alpha}\right] k\right)\right]\right\}}{M_{1,t}}
\end{aligned}
$$

It is easy to verify that this is a well defined CDF. Moreover, note that again we have a mass point at zero, which is an implication of the initial share of people in the city at time zero. To compute the expected log of income, as well as the variance, however, we also need to characterize the distribution of T_i. Thankfully, given that we know the marginal density of t_{m_i} of low-growth households, the only thing we need to know is the conditional density $f_{T_i|t_{m_i}}(x)$.

This is simply a truncated exponential distribution, and can be derived from the initial assumption that $T - t_m$ is exponential distributed, i.e. the time spent in the high growth regime follows an exponential distribution because of the memoryless Poisson process. But we need to incorporate the information that the distribution is truncated at $t - t_m$.

Moreover, there might be some confusion as to why the distribution is not uniform, as is usually the case with Poisson processes. Note, however, that here drawing your type is an absorbing state. This makes a big difference to the classic Poisson process where over an interval of time, multiple arrivals happen. In that world, conditioning on exactly one arrival over a time interval, would indeed yield a uniform distribution. But not here because agents can at most get one arrival. Fun fact: The first order approximation at zero is the same, which is intuitive because the standard process and my absorbing state agree at that point.

$$P\left(T_i - t_{m_i} \leq x | T_i < t, t_{m_i}\right) = \frac{1 - \exp\left(-\lambda\left[x\right]\right)}{1 - \exp\left(-\lambda\left[t - t_{m_i}\right]\right)} D\left(x \in [0, t - t_{m_i}]\right)$$

micro moments

Armed with the CDF, it is straightforward to compute the conditional first and second moment of the log of income in the cross section of households, mimicking the plots in 2.3. The conditional variance and mean for households in the high-growth regime are given by the following integral

$$\begin{aligned} V_0 =& [g - g^*]^2 \mathbb{E}\left[[t - t_{m_i}]^2 | T_i \geq t\right] \\ E_0 =& [g - g^*] \mathbb{E}\left[[t - t_{m_i}] | T_i \geq t\right] \end{aligned}$$

There is no conceptual difficulty to solve for the moments using pen and paper – integration is straight forward. I recommend, however, to use software because the the expressions do not simplify nicely and no additional insight is gained in spite of much pain. Especially for the moments conditional on households being in the absorbing state.

We also need the moments for households in the low growth regime, those are slightly more complicated because we had to handle both t_{m_i} as well as T_i, both of which are modeled as random variables.

$$
\begin{aligned}
m_k &= \int (\log y)^k \, dF\left(y|i \in M_{1,t}\right) \\
&= [g-g^*]^k \, \mathbb{E}\left[\left(T_i - t_{m_i}\right)^k |i \in M_{1,t}\right] \\
&= [g-g^*]^k \int \mathbb{E}\left[\left(T_i - t_{m_i}\right)^k |t_{m_i} = z\right] dF_{t_{m_i}}\left(z|i \in M_{1,t}\right) \\
&= [g-g^*]^k \int_0^t \int \left(T_i - z\right)^k \frac{\lambda \exp\left(-\lambda\left[x\right]\right)}{1-\exp\left(-\lambda\left[t-t_{m_i}\right]\right)} dx \, dF_{t_{m_i}}\left(z|i \in M_{1,t}\right) \\
&= [g-g^*]^k \int_0^t \int_0^{t-z} \left(x\right)^k \frac{\lambda \exp\left(-\lambda\left[x\right]\right)}{1-\exp\left(-\lambda\left[t-z\right]\right)} dx \, dF_{t_{m_i}}\left(z|i \in M_{1,t}\right) \\
&= [g-g^*]^k \int_0^t \int_0^{t-z} \left(x\right)^k \frac{\lambda \exp\left(-\lambda\left[x\right]\right)}{1-\exp\left(-\lambda\left[t-z\right]\right)} dx \, dF_{t_{m_i}}\left(z|i \in M_{1,t}\right)
\end{aligned}
$$

Now we have all the pieces together to compute the conditional expectations. In case you are wondering where $\log A_t$ shows up, I am dropping it because it is an inessential constant that cancels in case of the variance of the log of income. Note that all urban workers enjoy growth in A equally. I do add it back in when computing mean income.

time-dependent distribution of the log of income

Lastly, based on the previous insights we can also derive the entire income distribution at every point in time in closed form. While stationary distributions are often tractable, there are few applications that allow for a characterization of the transitional income distribution which is made possible here by imposing strong assumptions on the income process. As before, I focus on the distribution of the relative log of income, i.e. I drop $\log A_t$.

$$
\begin{aligned}
P\left(\log y_{i,t} \leq k\right) &= \frac{M_{0,t}}{M_{0,t}+M_{1,t}} P\left(\log y \leq k|i \in M_{0,t}\right) + \frac{M_{1,t}}{M_{0,t}+M_{1,t}} P\left(\log y \leq k|i \in M_{1,t}\right) \\
&= \frac{M_{0,t}}{M_{0,t}+M_{1,t}} \left\{P_0\left(t - t_m \leq \frac{k}{g-g^*}\right) D\left(0 \leq k \leq t\left(g-g^*\right)\right) + D\left(k > t\left(g-g^*\right)\right)\right\} \\
&+ \frac{M_{1,t}}{M_{0,t}+M_{1,t}} \mathbb{E}_{t_m} \mathbb{E}_{\log\varphi} P_1\left(0 < T_i - t_m < \frac{k-\log\varphi}{g-g^*}|t_m, \varphi\right) \\
&+ \frac{M_{1,t}}{M_{0,t}+M_{1,t}} P\left(\log\varphi \leq k\right) \frac{M_{1,0}}{M_{1,t}}
\end{aligned}
$$

This first step is easy, where I simply condition on being in the high or low growth regime as before. The second line obtains because the probability of the log of income to be smaller than some threshold k is zero when k is negative, and unity when k exceeds the maximum log income that a household in the high-growth regime could have obtained, namely $t(g-g^*)$, which is the log of income of a household that has been growing high since time zero. Note that P_0 denotes the conditional probability, and D is an indicator function.

Note that the fourth line arises because there is a mass point at time zero of households that do not not experience fast income growth, i.e. $P(t_{m_i} - T^i = 0) = M_{1,0}$. I suppose that they have been hit by the same inequality shock, though, so that at time zero there is a non-degenerate distribution. Next, I focus on the third line, which represents the probability of a low-growth household to have log income below some threshold k. Here, the analysis is complicated by the type draw. For example, households that did not spend much time in the high growth regime might still have a large income if they received a very high type draw φ.

$$
\begin{aligned}
&\frac{M_{1,t}}{M_{0,t}+M_{1,t}}\mathbb{E}_{t_m}\mathbb{E}_{\log\varphi}P_1\left(T_i - t_m < \frac{k-\log\varphi}{g-g^*}|t_m,\varphi\right)\\
=&\frac{M_{1,t}}{M_{0,t}+M_{1,t}}\mathbb{E}_{t_m}\mathbb{E}_{\log\varphi}P_1\left(T_i - t_m \le \frac{k-\log\varphi}{g-g^*}\right)D\left(0\le k-\log\varphi \le (g-g^*)(t-t_m)\right)\\
+&\frac{M_{1,t}}{M_{0,t}+M_{1,t}}\mathbb{E}_{t_m}\mathbb{E}_{\log\varphi}P_1\left(T_i - t_m \le \frac{k-\log\varphi}{g-g^*}\right)D\left(\frac{k-\log\varphi}{g-g^*}<0\right)\\
+&\frac{M_{1,t}}{M_{0,t}+M_{1,t}}\mathbb{E}_{t_m}\mathbb{E}_{\log\varphi}P_1\left(T_i - t_m \le \frac{k-\log\varphi}{g-g^*}\right)D\left(\frac{k-\log\varphi}{g-g^*}>t-t_m\right)\\
=&\frac{M_{1,t}}{M_{0,t}+M_{1,t}}E_{t_m}E_{\log\varphi}P_1\left(T_i - t_m \le \frac{k-\log\varphi}{g-g^*}\right)D\left(0\le k-\log\varphi \le t-t_m\right)\\
+&\frac{M_{1,t}}{M_{0,t}+M_{1,t}}E_{t_m}E_{\log\varphi}*0*D\left(\frac{k-\log\varphi}{g-g^*}<0\right)\\
+&\frac{M_{1,t}}{M_{0,t}+M_{1,t}}E_{t_m}E_{\log\varphi}*1*D\left(\frac{k-\log\varphi}{g-g^*}>t-t_m\right)
\end{aligned}
$$

where again I use the fact that the probability of $T_i - t_m < 0$ is zero, and the probability of $T_i - t_m < x$ is one when $x > t - t_m$. The first statement says that the income accrued during the high growth phase is non negative. The second says that time spent in the high growth regime is bounded from above by $t - t_m$ for each agent in the absorbing state. In turn, the probability that a household spent less time in the high growth regime than $t - t_m$ is one.

In the next step, I use the normality of $\log\varphi$. Moreover, integrating against $\log\varphi$ requires to get the right boundaries. For $T_i - t_m$ to be non-negative, $\log\varphi$ can be no larger than k.

Similar reasoning leads to the lower bound $k-(t-t_m)(g-g^*)$. A larger $\log\varphi$ implies that the household in the high growth regime is below that threshold with probability one. Let Φ denote the CDF of the standard normal distribution.

$$
\begin{aligned}
=&\frac{M_{1,t}}{M_{0,t}+M_{1,t}}E_{t_m}\int_{[k-(t-t_m)(g-g^*)]}^{k} P_1\left(T_i-t_m\leq\frac{k-\log\varphi}{g-g^*}\right)dF\left(\log\varphi\right)\\
&+\frac{M_{1,t}}{M_{0,t}+M_{1,t}}E_{t_m}\Phi\left(\frac{k-\left(g-g^*\right)\left(t-t_m\right)}{\sigma}+\frac{\sigma}{2}\right)\\
=&\frac{M_{1,t}}{M_{0,t}+M_{1,t}}E_{t_m}\int_{[k-(t-t_m)(g-g^*)]}^{k}\frac{1-\exp\left(-\lambda\left[\frac{k-\log\varphi}{g-g^*}\right]\right)}{1-\exp\left(-\lambda\left[t-t_m\right]\right)}dF\left(\log\varphi\right)\\
&+\frac{M_{1,t}}{M_{0,t}+M_{1,t}}E_{t_m}\Phi\left(\frac{k-\left(g-g^*\right)\left(t-t_m\right)}{\sigma}+\frac{\sigma}{2}\right)\\
=&\frac{M_{1,t}}{M_{0,t}+M_{1,t}}E_{t_m}\frac{1}{1-\exp\left(-\lambda\left[t-t_m\right]\right)}\\
&\times\int_{\{k-(t-t_m)(g-g^*)\}}^{k}1-\exp\left(-\lambda\left[\frac{k}{g-g^*}\right]\right)\exp\left(+\lambda\left[\frac{\log\varphi}{g-g^*}\right]\right)dF\left(\log\varphi\right)\\
&+\frac{M_{1,t}}{M_{0,t}+M_{1,t}}E_{t_m}\Phi\left(\frac{k-\left(g-g^*\right)\left(t-t_m\right)}{\sigma}+\frac{\sigma}{2}\right)
\end{aligned}
$$

which can be rewritten as

$$
\begin{aligned}
&=\frac{M_{1,t}}{M_{0,t}+M_{1,t}}E_{t_m}\frac{1}{1-\exp(-\lambda[t-t_m])}\left[\Phi\left(\frac{k}{\sigma}+\frac{\sigma}{2}\right)-\Phi\left(\frac{k-(g-g^*)(t-t_m)}{\sigma}+\frac{\sigma}{2}\right)\right]\\
&-\frac{M_{1,t}}{M_{0,t}+M_{1,t}}E_{t_m}\frac{\exp\left(-\lambda\left[\frac{k}{g-g^*}\right]\right)}{1-\exp(-\lambda[t-t_m])}\int_{\{k-(t-t_m)(g-g^*)\}}^{k}\exp\left(\frac{\lambda}{g-g^*}\log\varphi\right)dF\left(\log\varphi\right)\\
&+\frac{M_{1,t}}{M_{0,t}+M_{1,t}}E_{t_m}\Phi\left(\frac{k-(g-g^*)(t-t_m)}{\sigma}+\frac{\sigma}{2}\right)
\end{aligned}
$$

In order to proceed, we need to know what the moment generating function of a double truncated normal distribution looks like. If you stare at the penultimate line long enough, you will note that this will help us pin down the value of the integral. Wikipedia knows the answer ($https://en.wikipedia.org/wiki/Truncated\_normal\_distribution$).

$$\mathbb{E}[x^t|a\leq x\leq b]=\exp\left(\mu t+\frac{\sigma^2t^2}{2}\right)\left[\frac{\Phi\left(\beta-\sigma t\right)-\Phi\left(\alpha-\sigma t\right)}{\Phi\left(\beta\right)-\Phi\left(\alpha\right)}\right]$$

with $\alpha = \frac{a-\mu}{\sigma}$ and $\beta = \frac{b-\mu}{\sigma}$. Using this result yields

$$
\begin{aligned}
&= \frac{M_{1,t}}{M_{0,t}+M_{1,t}} E_{t_m} \frac{1}{1-\exp\left(-\lambda\left[t-t_m\right]\right)} \left[\Phi\left(\frac{k}{\sigma}+\frac{\sigma}{2}\right) - \Phi\left(\frac{k-\left(g-g^*\right)\left(t-t_m\right)}{\sigma}+\frac{\sigma}{2}\right)\right] \\
&- \frac{M_{1,t}}{M_{0,t}+M_{1,t}} E_{t_m} \exp\left(-\frac{\lambda}{g-g^*}\left(k-\frac{\sigma^2}{2}\left(\frac{\lambda}{g-g^*}-1\right)\right)\right) \\
&\times \left[\Phi\left(\frac{k}{\sigma}+\frac{\sigma}{2}-\sigma\frac{\lambda}{g-g^*}\right) - \Phi\left(\frac{k-(t-t_m)(g-g^*)}{\sigma}+\frac{\sigma}{2}-\sigma\frac{\lambda}{g-g^*}\right)\right] \\
&+ \frac{M_{1,t}}{M_{0,t}+M_{1,t}} E_{t_m} \Phi\left(\frac{k-\left(g-g^*\right)\left(t-t_m\right)}{\sigma}+\frac{\sigma}{2}\right)
\end{aligned}
$$

Finally, make sure to compute the expectation against the appropriate density of $f_{t_{m_i}|i\in M_{1,t}}(t_m)$, so we get

$$
\begin{aligned}
&= \frac{M_{1,t}}{M_{0,t}+M_{1,t}} \int_0^t \frac{1}{1-\exp\left(-\lambda\left[t-t_m\right]\right)} \left[\Phi\left(\frac{k}{\sigma}+\frac{\sigma}{2}\right) - \Phi\left(\frac{k}{\sigma}+\frac{\sigma}{2}-\frac{\left(g-g^*\right)\left(t-t_m\right)}{\sigma}\right)\right] dF_1(t_m) \\
&- \frac{M_{1,t}}{M_{0,t}+M_{1,t}} \int_0^t \exp\left(-\frac{\lambda}{g-g^*}\left(k-\frac{\sigma^2}{2}\left(\frac{\lambda-\left(g-g^*\right)}{g-g^*}\right)\right)\right) \\
&\times \left[\Phi\left(\frac{k}{\sigma}+\frac{\sigma}{2}-\sigma\frac{\lambda}{g-g^*}\right) - \Phi\left(\frac{k}{\sigma}+\frac{\sigma}{2}-\sigma\frac{\lambda}{g-g^*}-\frac{(t-t_m)(g-g^*)}{\sigma}\right)\right] dF_1(t_m) \\
&+ \frac{M_{1,t}}{M_{0,t}+M_{1,t}} \int_0^t \Phi\left(\frac{k-\left(g-g^*\right)\left(t-t_m\right)}{\sigma}+\frac{\sigma}{2}\right) dF_1(t_m)
\end{aligned}
$$

This completes the derivation, since we know the density of $f(t_m)$ and can simply compute the integral. Putting all the pieces together then yields the CDF. Importantly, the CDF is a function of time. In the main part of the paper I show how this simple framework can deliver inequality dynamics that mimic the ones observed in the data.

B.3 Derivation of f_0, f_1

The conditional density for household income in the high growth regime reads

$$
f_0\left(k\right) = \frac{1}{M_{0,t}} \frac{g^*}{\left(g-g^*\right)\left(1-\alpha\right)} L_t k^{-\left(\frac{\lambda}{g-g^*}-\frac{g^*}{(1-\alpha)(g-g^*)}\right)-1} \tag{B27}
$$

where the probability mass at $y = \exp\left((g-g^*)t\right)$ is equal to $\frac{M_{0,0}\exp(-\lambda t)}{M_{t,0}}$. That is to say, there is a positive mass of agents who start growing fast at time zero, and this mass point shrinks exponentially over time.

The conditional densities for household income based on convergence growth only (ignoring the type draw) in the low-growth regime reads

$$f_1\left(k\right) = \frac{1}{M_{1,t}}\frac{\lambda}{\left(g-g^*\right)}k^{-\frac{\lambda}{(g-g^*)}-1}\left\{1 - M_{1,0} - L_t^r\left[k^{\frac{g^*}{(g-g^*)(1-\alpha)}}\right]\right\} \tag{B28}$$

with a mass point at 1 with probability $M_{0,1}$.

Proof:

Let's start with f_0. Use a simple change of variable, and knowledge of the distribution of t_{m_i}.[191] Then,

$$\begin{aligned}
P\left(y_{0,t}^i \le k\right) &= P\left(\log y_{0,t}^i \le \log k\right)\\
&= P\left(\left(g-g^*\right)\left(t-t_{m_i}\right) \le \log k\right)\\
&= P\left(\left(g-g^*\right)\left(t-t_{m_i}\right) \le \log k\right)\\
&= P\left(t - t_{m_i} \le \frac{\log k}{g-g^*}\right) D\left(k \le \exp\left(\left(g-g^*\right)t\right)\right)\\
&= P\left(t - t_{m_i} \le \frac{\log k}{g-g^*}\right) D\left(0 \le \frac{\log k}{g-g^*} \le t\right)
\end{aligned}$$

where D is an indicator function that keeps track of the time-dependent support of the distribution. Keeping in mind the mass point at zero, we could obtain the continuous part of the density by differentiating the previous expression with respect to k. Note that the density of $x = t - t_m$ can be obtained using a straightforward change of variable and (B26).

[191] Note that I keep everything normalized by the constant growth rate g^* to study the stationary distributions.

Differentiating with respect to k, after using $f_X(x) = \frac{\frac{g^*}{1-\alpha} L_0^r exp\left(-\frac{g^*}{1-\alpha}t\right) exp\left(-\left(\lambda - \frac{g^*}{1-\alpha}\right)x\right)}{M_{0,t}}$ yields

$$\begin{aligned}
f_0(k) &= \frac{1}{k(g-g^*)} f_x\left(\frac{\log k}{g-g^*}\right) \\
&= \frac{g^*}{k(g-g^*)(1-\alpha)} \frac{L_0^r \exp\left(-\frac{g^*}{1-\alpha}t\right) \exp\left(-\left(\frac{\lambda}{g-g^*} - \frac{g^*}{(1-\alpha)(g-g^*)}\right)\log k\right)}{M_{0,t}} \\
&= \frac{1}{M_{0,t}} \frac{g^*}{(g-g^*)(1-\alpha)} L_t k^{-\left(\frac{\lambda}{g-g^*} - \frac{g^*}{(1-\alpha)(g-g^*)}\right)-1}.
\end{aligned}$$

The mass point at $x = 0$ follows by noting that there is an initial mass of households $M_{0,0}$ in the high growth regime, and these households are pulled out randomly by the Poisson process. Hence that share declines at an exponential rate λ.

Now let's focus on f_1. Now we again can use previous results about the log of income distribution to compute the output

$$\begin{aligned}
P(y_t \le k | T_i < t) &= P(\log y_t \le \log k | T_i < t) \\
&= P((g-g^*)(T_i - t_{m_i}) \le \log k | T_i < t) \\
&= \mathbb{E}_{t_m} P\left((T_i - t_{m_i}) \le \frac{\log k}{(g-g^*)} | T_i < t, t_m\right)
\end{aligned}$$

Using the same trick as before, I condition on t_m to then integrate over it. In doing so, I simplify the problem because I know the marginal distribution of t_m, and I also know that T_i given t_{m_i} is a truncated exponential distribution, by the Poisson arrival of learning your type. First, recall the conditional density for t_m is

$$f_{t_m|T<t}(k) = \frac{1}{M_{1,t}} \left(\frac{g^* L_0^r}{1-\alpha}\right) \left[\exp\left(-\frac{g^*}{1-\alpha}k\right) - \exp\left(\frac{\lambda(1-\alpha) - g^*}{1-\alpha}k\right) \exp(-\lambda t)\right].$$

Then, we can go ahead and compute the density. First, we use the law of iterated expectations to split the expression into two pieces, where D is again an indicator function.

$$
\begin{aligned}
P\left(y_t \leq k|T_i < t\right) &= \mathbb{E}_{t_m} P\left(T_i - t_{m_i} \leq \frac{\log k}{(g-g^*)}|T_i < t, t_m\right) \\
&= \mathbb{E}_{t_m}\left\{\frac{1-\exp\left(-\lambda\left[\frac{\log k}{(g-g^*)}\right]\right)}{1-\exp\left(-\lambda\left[t-t_{m_i}\right]\right)} D\left(\frac{\log k}{(g-g^*)} \leq t - t_{m_i}\right) + D\left(\frac{\log k}{(g-g^*)} > t - t_{m_i}\right)\right\} \\
&= \mathbb{E}_{t_m}\left\{\frac{1-\exp\left(-\lambda\left[\frac{\log k}{(g-g^*)}\right]\right)}{1-\exp\left(-\lambda\left[t-t_{m_i}\right]\right)} D\left(\frac{\log k}{(g-g^*)} \leq t - t_{m_i}\right)\right\} \\
&+ \int_0^t \left\{D\left(\frac{\log k}{(g-g^*)} > t - t_{m_i}\right)\right\} dF\left(t_m | i \in M_{1,t}\right).
\end{aligned}
$$

Note that we need to account for the mass point at zero again that comes from the share of agents in the city that already know their type, and we also need to keep track of the mass point of agents that start growing at the high rate,

$$
\begin{aligned}
&= \int_0^{t-\frac{\log k}{(g-g^*)}} \left\{\frac{1-\exp\left(-\lambda\left[\frac{\log k}{(g-g^*)}\right]\right)}{1-\exp\left(-\lambda\left[t-t_{m_i}\right]\right)}\right\} dF_1\left(t_m\right) + F\left(t_m = 0\right)\frac{1-\exp\left(-\lambda\left[\frac{\log k}{(g-g^*)}\right]\right)}{1-\exp\left(-\lambda\left[t\right]\right)} \\
&+ \int_{t-\frac{\log k}{(g-g^*)}}^t 1 * dF\left(t_m | i \in M_{1,t}\right) + \frac{M_{1,0}}{M_{1,t}} \\
&= \int_0^{t-\frac{\log k}{(g-g^*)}} \left\{\frac{1-\exp\left(-\lambda\left[\frac{\log k}{(g-g^*)}\right]\right)}{1-\exp\left(-\lambda\left[t-t_{m_i}\right]\right)}\right\} dF_1\left(t_m\right) + F\left(t_m = 0\right)\frac{1-\exp\left(-\lambda\left[\frac{\log k}{(g-g^*)}\right]\right)}{1-\exp\left(-\lambda\left[t\right]\right)} \\
&+ F_1\left(t\right) - F_1\left(t - \frac{\log k}{g-g^*}\right) + \frac{M_{1,0}}{M_{1,t}} \\
&= \int_0^{t-\frac{\log k}{(g-g^*)}} \left\{\frac{1-\exp\left(-\lambda\left[\frac{\log k}{(g-g^*)}\right]\right)}{1-\exp\left(-\lambda\left[t-t_{m_i}\right]\right)}\right\} dF_1\left(t_m\right) + F\left(t_m = 0\right)\frac{1-\exp\left(-\lambda\left[\frac{\log k}{(g-g^*)}\right]\right)}{1-\exp\left(-\lambda\left[t\right]\right)} \\
&+ F_1\left(t\right) - F_1\left(t - \frac{\log k}{g-g^*}\right) + \frac{M_{1,0}}{M_{1,t}}.
\end{aligned}
$$

Now we use the density $f_1(t_m|T^i < t)$, whic

$$= \frac{1}{M_{1,t}} \int_0^{t-\frac{\log k}{(g-g^*)}}$$

$$\times \left\{ \frac{1-\exp\left(-\lambda\left[\frac{\log k}{(g-g^*)}\right]\right)}{1-\exp\left(-\lambda\left[t-t_m\right]\right)} \right\} \left(\frac{g^* L_0^r}{1-\alpha}\right) \left[\exp\left(-\frac{g^*}{1-\alpha}t_m\right) - \exp\left(\frac{\lambda\left(1-\alpha\right)-g^*}{1-\alpha}t_m\right)\exp\left(-\lambda t\right)\right] dt_m$$

$$+ \frac{\left(1-\exp\left(-\lambda t\right)\right) M_{0,0}}{M_{1,t}} \frac{1-\exp\left(-\lambda\left[\frac{\log k}{(g-g^*)}\right]\right)}{1-\exp\left(-\lambda t\right)}$$

$$+ F_1\left(t\right) - F_1\left(t - \frac{\log k}{g-g^*}\right) + \frac{M_{1,0}}{M_{1,t}}$$

$$= \frac{1}{M_{1,t}} \int_0^{t-\frac{\log k}{(g-g^*)}} \left\{ \frac{1-\exp\left(-\lambda\left[\frac{\log k}{(g-g^*)}\right]\right)}{1-\exp\left(-\lambda\left[t-t_m\right]\right)} \right\} \left(\frac{g^* L_0^r}{1-\alpha}\right) \exp\left(-\frac{g^*}{1-\alpha}t_m\right) \left[1-\exp\left(-\lambda\left[t-t_m\right]\right)\right] dt_m$$

$$+ \frac{M_{0,0}}{M_{1,t}} \left(1-\exp\left(-\frac{\lambda}{(g-g^*)}\log k\right)\right)$$

$$+ F_1\left(t\right) - F_1\left(t - \frac{\log k}{g-g^*}\right) + \frac{M_{1,0}}{M_{1,t}}$$

which simplifies to

$$\begin{aligned} &= \tfrac{1}{M_{1,t}} \int_0^{t-\frac{\log k}{(g-g^*)}} \left\{1-\exp\left(-\lambda\left[\tfrac{\log k}{(g-g^*)}\right]\right)\right\} \left(\tfrac{g^* L_0^r}{1-\alpha}\right) \exp\left(-\tfrac{g^*}{1-\alpha}t_m\right) dt_m \\ &\quad + \tfrac{M_{0,0}}{M_{1,t}} \left(1-\exp\left(-\tfrac{\lambda}{(g-g^*)}\log k\right)\right) \\ &\quad + F_1\left(t\right) - F_1\left(t-\tfrac{\log k}{g-g^*}\right) + \tfrac{M_{1,0}}{M_{1,t}} \\ &= \tfrac{L_0^r}{M_{1,t}} \left\{1-\exp\left(-\lambda\left[\tfrac{\log k}{(g-g^*)}\right]\right)\right\} \left\{1-\exp\left(-\tfrac{g^*}{1-\alpha}\left(t-\tfrac{\log k}{(g-g^*)}\right)\right)\right\} \\ &\quad + \tfrac{M_{0,0}}{M_{1,t}} \left(1-\exp\left(-\tfrac{\lambda}{(g-g^*)}\log k\right)\right) \\ &\quad + F_1\left(t\right) - F_1\left(t-\tfrac{\log k}{g-g^*}\right) + \tfrac{M_{1,0}}{M_{1,t}} \end{aligned}$$

which reads

$$
\begin{aligned}
&= \tfrac{L_0^r}{M_{1,t}} \left\{1-\exp\left(-\tfrac{\lambda}{(g-g^*)}\log k\right)\right\}\left\{1-\exp\left(-\tfrac{g^*}{1-\alpha}t\right)\exp\left(\tfrac{g^*}{(g-g^*)(1-\alpha)}\log k\right)\right\} \\
&\quad + \tfrac{M_{0,0}}{M_{1,t}}\left(1-\exp\left(-\tfrac{\lambda}{(g-g^*)}\log k\right)\right) \\
&\quad + F_1(t) - F_1\left(t-\tfrac{\log k}{g-g^*}\right)+\tfrac{M_{1,0}}{M_{1,t}} \\
&= \tfrac{L_0^r}{M_{1,t}}\left\{\exp\left(-\tfrac{g^*}{1-\alpha}t\right)\exp\left(\tfrac{g^*-\lambda(1-\alpha)}{(g-g^*)(1-\alpha)}\log k\right)-\exp\left(-\tfrac{g^*}{1-\alpha}t\right)\exp\left(\tfrac{g^*}{(g-g^*)(1-\alpha)}\log k\right)\right\} \\
&\quad + \tfrac{L_0^r+M_{0,0}}{M_{1,t}}\left[1-\exp\left(-\tfrac{\lambda}{(g-g^*)}\log k\right)\right] \\
&\quad + F_1(t) - F_1\left(t-\tfrac{\log k}{g-g^*}\right)+\tfrac{M_{1,0}}{M_{1,t}}.
\end{aligned}
$$

Using the definition of L_t^r as well as the normalization $M_{0,0}+M_{1,0}+L_0^r=1$, we get

$$
\begin{aligned}
&= \frac{L_t^r}{M_{1,t}}\left\{\exp\left(-\frac{\lambda(1-\alpha)-g^*}{(g-g^*)(1-\alpha)}\log k\right)-\exp\left(\frac{g^*}{(g-g^*)(1-\alpha)}\log k\right)\right\} \\
&+ \frac{1-M_{1,0}}{M_{1,t}}\left[1-\exp\left(-\frac{\lambda}{(g-g^*)}\log k\right)\right] \\
&+ F_1(t)-F_1\left(t-\frac{\log k}{g-g^*}\right)+\frac{M_{1,0}}{M_{1,t}} \\
&= \frac{1}{M_{1,t}}\left\{(1-M_{1,0})\left[1-k^{-\frac{\lambda}{(g-g^*)}}\right]+L_t^r\left[k^{-\frac{\lambda(1-\alpha)-g^*}{(g-g^*)(1-\alpha)}}-k^{\frac{g^*}{(g-g^*)(1-\alpha)}}\right]+M_{1,0}\right\} \\
&+ F_1(t)-F_1\left(t-\frac{\log k}{g-g^*}\right)
\end{aligned}
$$

Now we can differentiate this expression with respect to k to obtain

$$
\begin{aligned}
f\left(y|y \in Y_1\right) =& \frac{1}{M_{1,t}} \left\{ \left(1 - M_{1,0}\right) \left[\frac{\lambda}{(g-g^*)} k^{-\frac{\lambda}{(g-g^*)}-1} \right] \right\} \\
& - \frac{1}{M_{1,t}} \left\{ \frac{L_t^r}{(g-g^*)(1-\alpha)} \left[\left(\lambda(1-\alpha) - g^*\right) k^{-\frac{\lambda(1-\alpha)-g^*}{(g-g^*)(1-\alpha)}-1} + g^* k^{\frac{g^*}{(g-g^*)(1-\alpha)}-1} \right] \right\} \\
& + \frac{1}{M_{1,t}} \frac{g^* L_0^r}{k(g-g^*)(1-\alpha)} \\
& \times \left[\exp\left(-\frac{g^*}{1-\alpha} \left(t - \frac{\log k}{g-g^*} \right) \right) - \exp\left(\frac{\lambda(1-\alpha) - g^*}{1-\alpha} \left(t - \frac{\log k}{g-g^*} \right) \right) \exp(-\lambda t) \right] \\
=& \frac{1}{M_{1,t}} \left(1 - M_{1,0}\right) \left[\frac{\lambda}{(g-g^*)} k^{-\frac{\lambda}{(g-g^*)}-1} \right] \\
& - \frac{1}{M_{1,t}} \frac{L_t^r}{(g-g^*)(1-\alpha)} \left[\left(\lambda(1-\alpha) - g^*\right) k^{-\frac{\lambda(1-\alpha)-g^*}{(g-g^*)(1-\alpha)}-1} + g^* k^{\frac{g^*}{(g-g^*)(1-\alpha)}-1} \right] \\
& + \frac{1}{M_{1,t}} \frac{g^* L_t^r}{(g-g^*)(1-\alpha)} \left[k^{\frac{g^*}{(g-g^*)(1-\alpha)}-1} - k^{-\frac{\lambda(1-\alpha)-g^*}{(g-g^*)(1-\alpha)}-1} \right] \\
=& \frac{1}{M_{1,t}} \left(1 - M_{1,0}\right) \left[\frac{\lambda}{(g-g^*)} k^{-\frac{\lambda}{(g-g^*)}-1} \right] \\
& - \frac{1}{M_{1,t}} \frac{L_t^r}{(g-g^*)(1-\alpha)} \left[\left(\lambda(1-\alpha) - g^* + g^*\right) k^{-\frac{\lambda(1-\alpha)-g^*}{(g-g^*)(1-\alpha)}-1} + \left(g^* - g^*\right) k^{\frac{g^*}{(g-g^*)(1-\alpha)}-1} \right]
\end{aligned}
$$

hence we obtain

$$
f\left(y_0|y_0 \in Y_1\right) = \frac{1}{M_{1,t}} \frac{\lambda}{(g-g^*)} k^{-\frac{\lambda}{(g-g^*)}-1} \left\{ 1 - M_{1,0} - L_t^r \left[k^{\frac{g^*}{(g-g^*)(1-\alpha)}} \right] \right\}
$$

for the continuous part of the density.

B.4 A version with even catch-up growth

To see how heterogeneous and risky income growth is central, let's consider a model economy where convergence growth is deterministic and occurs up until time T, when the households draw their type φ as before. I hold the aggregate level of convergence fixed but distribute the growth it takes to get there evenly. Foreshadowing a calibration exercise in the next section, I require that urban per capita GDP grows by a factor of 3.5 relative to the rest of the world, with $g = .07$ and $g^* = .02$, and $M_{0,1} = .175$. Moreover, the interest rate is 5% and the discount factor ρ is .01. This means that $T = \frac{\log(2.06)}{g-g^*} \approx 22.4$. Noting that the household optimality condition leads to an equalization of (expected) marginal utilities, and in particular at time T with $\Delta \to 0$,

$$c_{T-\Delta}^{-\eta} = \mathbb{E}_\varphi \left[(\varphi y_T + (\rho + [\eta - 1] g^*) a_T)^{-\eta} \right],$$

I can ask what level of inequality, here in the form of the variance of the type draw σ^2, is needed to generate capital outflows along the transition path. In relation to figure **??**, this is like asking: what level of variance is needed to push the intercept of the consumption profile below y_0. This follows since the slope of the consumption profile is only pinned down by preference parameters and the interest rate in this version of the model[192]. Consequently, the optimality condition simplifies to

$$\exp\left(\eta \left[g - g^*\right] T\right) = \mathbb{E}_\varphi \left(\varphi^i + \exp\left(\left[r^* - g\right] T\right) \left[\frac{r^* - g^*}{g - r^*} \left[\exp\left(\left[g - r^*\right] T\right) - 1\right] - \left[1 - \exp\left(\left[g^* - r^*\right] T\right)\right] \right] \right)^{-\eta}$$

where I used the fact that income and consumption grow at rate g_h and g^*, as well as the budget constraint and the requirement that $y_0 = c_0$. Assuming that the type draw is log normal, the variance of the log of the type that is needed to solve this equation is a staggering 4.75, and completely out of range of any empirically sensible estimate of income dispersion.[193] This highlights the importance of the stochastic nature of the catch-up growth on the household level, not just for tractability but also to quantitatively account for household asset accumulation despite strong convergence growth.

There are two additional remarks worth pointing out. First, note the tension between

[192] This simplified model is inconsistent with the strong comovement of consumption and output in the data as mentioned before. Output growth and consumption growth track each other.

[193] In this simplified version of the model, the type draw is the only source of uncertainty. Inequality measured in terms of the log of income therefore would exceed standard measures of labor income inequality of .6 and household income inequality of around 1 (D. Krueger, Mitman, and Perri 2016).

inter-temporal and intra-temporal smoothing. In a world where every household converges to some average level, a large coefficient of relative risk aversion will induce a strong consumption smoothing motive. Since lifetime utility becomes increasingly defined by the lowest level of consumption as η increases, households want to smooth consumption and borrow against their future lifetime income. This happens in the case with deterministic convergence. In a world with risky growth, however, this logic is turned upside down. If households are very risk averse, they effectively attach more weight to the worst convergence growth path inducing them to built up savings.

B.5 A model version with capital

Here I introduce a simple version of the model with capital that leaves all qualitative conclusions unchanged. First, I reinterpret what used to be the labor supply in the rural economy l_r^i of household i as a composite intermediate input that uses both raw rural labor and capital as inputs in a Cobb-Douglas fashion with capital elasticity β. Hence, $x_r^i = k_i^\beta l_i^{1-\beta}$. Rural total output is now given by $A^r X^\alpha$. Suppose that rural households save and invest a fraction s of their income to buy more capital, as in Solow (1956). Suppose that as before, the composite factor x flows out of the rural economy at a rate such that there is constant income growth for rural households at rate g^*.

Now consider the capital accumulation of household i,

$$\begin{aligned} k_i &= sy_i \\ &= sA^r X^{\alpha-1} \end{aligned}$$

Now focus on the balanced growth path where capital is accumulated at rate g^*. The steady state level is of course endogenous and depends on the saving rate and productivity etc. As before, normalize l_i to unity. Suppose, for the sake of the argument, that the steady state capital–effective urban labor ratio is such that $\beta \left(\frac{k_t^i}{A_t^u}\right)^{\beta-1} = r^*$.

This choice is motivated by the desire to ensure that every household that leaves the rural sector brings a sufficient amount of capital with them so that the influx of workers into the urban sector does not raise the marginal product of capital. Note that if entering workers were to enter without any capital, we would have two offsetting forces on the direction of capital flows. On the one hand, entering households will increase the marginal product of capital – a simple labor supply shock that should lead to capital inflows. On the other hand, precautionary savings are accumulated, potentially inducing outflows. It is ex ante unclear which force dominates. In this modified version, however, every worker enters the city with a sufficient amount of capital to ensure that the marginal product of capital is left unchanged. As before, miracle growth increases the effective units of labor of each household. But as long as $k_{t_m}^i < a_{t_m}^{**}$, we know that the household will accumulate assets at a rate that is higher than their labor income growth. Simply put, the household problem has not changed, except now the household does not start with an asset-to-income ratio that is zero but with one that is large enough to match the labor supply they are contributing to the urban economy. Whether, in fact, the desired buffer-stock asset-to-income ratio is larger than the amount of capital already owned by the household depends on the parameters of the model, especially the rural saving rate s as well as the risk in the urban economy. If there is no risk in the form of the type draw, we know that this inequality is never going to hold. If there is a sufficient amount of risk, this may well be the case.

Lastly, the reader may worry that the type draw itself creates some complications by raising the marginal product of capital. Since production is constant returns to scale, it is easiest to consider one household first, and aggregate up in a last step

$$y_t^i = k_t^\beta \left(A_t \varphi h_t\right)^{1-\beta} \tag{B29}$$

The allocation of capital to each household unit is simply given by the first order condition with respect to k

$$A_t \varphi h_t \left(\frac{\beta}{r^*}\right)^{\frac{1}{1-\beta}} = k_t$$

Note that whenever a fraction of households draws their type, there is no jump in the aggregate demand for capital since the type draw is centered around one. Capital can be reshuffled within each cohort of agents drawing their type while leaving the overall demand for capital in the economy unchanged. This concludes the generalization of the model, showing that it is in principal able to accommodate the inclusion of capital as a factor of production.

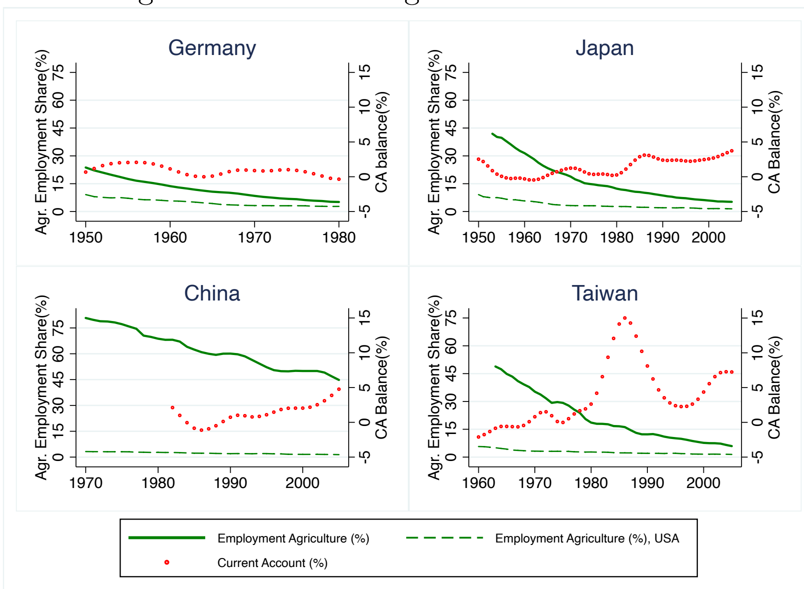
B.6 Aggregate Time Series

B.7 CHIP

When using the CHIP data, I make a choice to only focus on urban males aged 23 to 60. I also focus on full-time employees which leads me to drop workers that work for less than 6 hours a day, or workers that work less than 4 days a week. The reason for focusing on urban workers is that income is hard to observe on the country side, especially in 1988 because most people operate small scale farming units and do not earn a normal wage. Furthermore, the focus of my paper is on rural-urban structural change and inequality in urban areas seems like a better proxy for the kind of risk that an agricultural household is exposed to when entering a modern occupation.

It is also important to note that inequality in rural China is and was substantial (Piketty, L. Yang, and Zucman 2019). As mentioned in the introduction, my model does not necessar-

Figure B2: Str. Change & Current Account



Relationship between agricultural employment share and current account for Germany, Japan, China, and Taiwan. Current Account series is smoothed using an hp-filter with smoothing parameter of 8.5. Current account data is from the WDI, Taiwanese Statistical office, and the historical macro database from Jordà, Schularick, and Taylor (2017) (for Germany and Japan).

ily require inequality to be low on the country side. The key that makes inequality matter from an insurance and risk perspective is when it is combined with a learning-about-your-type mechanism. This seems more relevant in the city. This is why I plot the development of the average log wage and the variance of the log wage in 17 in urban areas, and omit the rural counterpart.

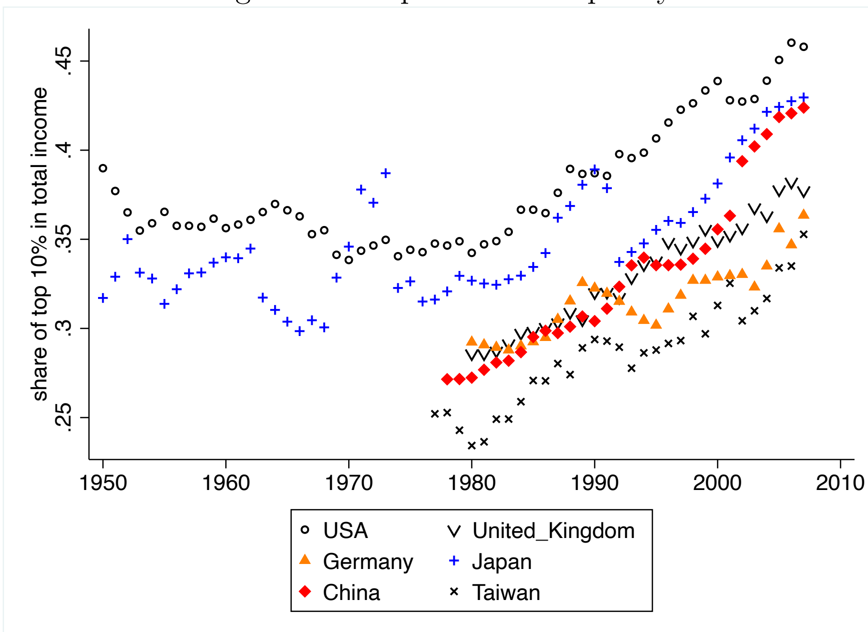
Table B1: Log Variance of Income

	no covariates				*age and age square netted out*			
	1988	1995	2002	2013	1988	1995	2002	2013
log mean income	8.49	9.18	9.57	10.41	8.54	9.19	9.55	10.46
log variance of income	0.13	0.24	0.32	0.46	0.12	0.22	0.32	0.45
Observations	7260	2853	3076	2927	7260	2853	3076	2927

Note: The table reports results from a simple linear regression of log income on a constant with and without a second order polynomial of age for male household heads of age 23 to 60. Mean income for the specification with age is projected for a household head of the age of 42, the sample average. Additional information on how the sample has been selected is in the appendix in section B.7 .

B.8 Aggregate Inequality

Figure B3: Top Income Inequality



Share of income of top 10% earners to total income from the World Inequality Database. Raw data.

B3 displays the raw data from the WID. I also added Germany and the UK. Germany is added for consistency with the other motivating figures. The issue with Germany is that the series is too short to reveal the trajectory of inequality measured in terms of the income share captured by the top 10%, and German unification also mechanically pushes down inequality since there was relatively little inequality in the former communist part of East Germany. See Card, Heining, and Kline (2013) for an analysis of the rise in West German wage inequality.

B.9 Data Appendix CFPS 2012

Sample Selection CFPS 2012

To run this regression, I restrict the sample to employed household heads that are between 23 and 60 years old, in line with previous work (He et al. 2018; Storesletten, Telmer, and Yaron 2004). This is done because students' or retirees' savings behavior is strongly related to life cycle patterns and not well captured by precautionary models. Moreover, I drop the households with the smallest 4% of income realizations for each group, for example in the urban-rural sample I drop the household below the 4th percentile of income per capita and consumption per capita within urban households, and I do the same within the sample of rural households. The CFPS does not define household heads, and I assume the highest earner is the household head. When running regressions for other waves or data sets I incorporate the same restrictions imposed on the sample here when possible. I do not use the sampling weights provided by the CFPS. The reason is twofold: first, I am not that interested in obtaining estimates that are representative of the Chinese economy as a whole. I aim to document urban-rural differences and for that I treat every observation with equal weight. Second, I am using the qreg2 command from Machado, P. Parente, and Santos Silva (2020) to obtain heteroscedastic-robust standard errors, which doesn't work with sampling weights. The results change very little, however, when using sampling weights from the CFPS and the standard qreg command.

Additional Information on Key Measurement Concepts

There are additional important remarks regarding measurement of important variables and concepts. I will discuss measurement of income, assets, consumption, and the definition of a family, as well as hukou status and urban vs. rural categories in turn. Unless otherwise indicated, the main source of information is from the data guide of the CFPS available under this link: https://opendata.pku.edu.cn/dataverse/CFPS?language=en.

Income

Income is measured annually, and adds up the different income source including transfer income, wages, rent and asset returns, bonuses, net profits etc. These variables are measured net-of-tax. Importantly, and adjustment has been made for families with agricultural production to take into account the fraction of production that is not being sold but consumed by the family directly.

Assets

This paragraph is copied from page 105 of the data guide of the CFPS:

"*In the CFPS 2010 and 2012 family questionnaires, the variable total_asset indicates the net family asset value, which was the difference between family total assets and total liabilities. Family assets include land, housing, financial assets, productive fixed assets and durable goods. Family liabilities include housing liabilities and non-housing liabilities. The value of land was estimated, for example, assuming that* 25% *of the family agricultural income comes from land and the return rate of land is* 8%*, and we could estimate the value of land (Mckinley, 1993). The housing property includes current residence and other housing. When calculating the value of house property, we counted a house with partial property rights as* *full property rights since we were not informed of the proportion of the property rights and a household has perpetuity. Financial assets include deposits, stocks, funds, bonds, financial derivatives, other financial products and borrowings. The data in 2010 did not contain the value of bonds, financial derivatives and other financial products. Productive fixed assets include productive firm assets, agricultural machinery and so on. Durable goods include automobiles, televisions, computers, refrigerators and other common household appliances. Housing liabilities is the number reported when answering the question about "housing debt with interest". Non-housing liabilities counts debt from education or medical care.*"

Essentially, the total-asset variable should represent the net asset position of the household, with the caveat that some financial products are missing for 2010. To the extent that financial products are more likely to be used by urban households, this should bias my results down, strengthening the empirical results.

Consumption

Measurement of consumption is a non-trivial challenge. First, note that durable goods are captured in the composite measure of consumption (*Expense*), except for cars. This is important because durable goods consume is more volatile than aspects of consumption that do not represent a long stream of services like durable goods. Second, an important issue in survey-based data is recall, especially when dealing with a yearly measure of consumption. There were 3 types of recall rates, last year, last month, and last week, and applied in the questionnaire when most appropriate. All answers have then been aggregated back up onto the year level. Accounting for different recall rates is important as shown in Deaton (2003).

Concept of a Household/Family

Every household in the CPFS has at least one Chinese national, and the family is defined as interdependent economic unit. Household members are defined as financially dependent immediate relatives, or non-immediate relatives who lived with the household for more than three consecutive months and are financially related to the sampled household. That includes families with household members who migrated for work to another city. It does not include family members that got married and started their own family. In the panel dimension, households are "split up" to keep track of those changes and follow the different household units over time. Ideally, the household unit therefore also incorporates migrant workers. A feature of the sample that has been exploited by Xu and Xie (2015).

Hukou Status

Here I discuss a few key aspect of the hukou system, based on Y. Song (2014). The hukous status is also known as household registration and separates rural vs. urban or agricultural vs. non-agricultural hukou, where the two terms are used interchangeably. In general, the hukou system is complex because local governments have much leeway in determining hukou policies within their jurisdiction. All Chinese national's personal hukou is characterized by two classifications: hukou type and hukou location. At birth, a child inherits both type and location from their parents. The hukou type refers to the urban vs. rural hukou distinction. The hukou location is passed on at birth, and a person can be distinguished by whether she has a local or a non-local hukou with respect to an administrative unit. The local hukou registration defines the eligibility for public services provided by local governments, i.e. the benefits are different for local vs non-local and urban vs rural hukou type.

Before 1980, households with rural hukou were not allowed to leave the country side and were mostly restricted to agricultural production. Only under special circumstances could households change their hukou. The main reasons that allowed households to change their hukou was recruitment by state-owned enterprises, college education, and joining the army.

From 1980 onward, many local governments have eased the restrictions associated with the hukou system and made internal migration relatively easier. As a result, many households migrated to the city but kept their rural hukou. While restrictions have been eased incrementally, Y. Song (2014) and the literature cited therein makes a convincing case that gaining access by changing type and location of hukou status is common only among a small minority of successful and rich individuals in the booming centers of China such as Shanghai or Bejing. Recently, a common way to obtain valuable urban local hukou status in Shanghai or Bejing is by simple money transfers and property purchases. On the other hand, some provinces have lifted to hukou restrictions altogether. Additionally, a rural hukou no longer means that households are bound to agricultural production. Some might have turned into successful entrepreneurs with potentially high asset-to-income ratios. Lastly, while I observe the hukou location, I do not know it relative to the current location of work. Yet, it seems to make little sense to only change the hukou type but not the location.[194]

Given all these caveats and measurement challenges, it is surprising that I can detect robust differences in asset-to-income ratios based on the change in the hukou type (agr vs. non-agr) of households.

Urban vs Rural

In general, the distinction between urban and rural areas is not sharp, and depends on context and country. Standard criterion are population density, living conditions and amenities, as well as industrial composition.

[194] Does someone know more about this than I do?

Qin and Y. Zhang (2014) highlight some of the difficulties with using urban rural definitions categories on the Census Bureaus definition. The definition has been shifting over time, and while in the 50s and 60s the hukous status was a good indicator of urban vs. rural household, this has been less true over the last decades as many families have moved to urban areas without urban hukou. Overall, the definition urban-rural is not comparable over time, so the best I can do is to state the definition for the 2010 Census.

The 2010 census bases the rural-urban distinction on the community level, which is the lowest level of administrative unit in China Gan et al. (2019). Urban areas are defines an area "of continuous built-up with urban facilities" (Qin and Y. Zhang 2014, p. 500). Gan et al. (2019) summarize the new 2010 definition as follows:

"*This new standard is solely based on land contiguity by actual construction, which refers to public facilities, residential facilities etc., either completed or under construction. For example, in districts, if a community is contiguous to the district-level government by actual construction, it is classified as urban; otherwise it should be regarded as rural. Industrial parks, economic development zones, colleges or farms that are not contiguous to the area where the local government is located but with population more than 3,000 are also categorized as urban. As a result, this rural and urban division does not directly take population densities, economic activities, or residential infrastructures into major consideration. Hence, there is a possibility that a community is officially reclassified from rural to urban only because its attribute of land contiguity has been changed. It is worth pointing out that the standard does not alter the administrative division, affiliation status or land planning; instead, it is mainly for statistical accounting use.*"(p. 7)

Importantly, urban areas are not identical to what would be categorized as a city. All cities are urban areas, but not all urban areas are cities. A "city" is an administrative unit and usually larger than the average urbanized area.

Qin and Y. Zhang (2014) both argue that much of the growth in urban areas is driven by internal migration and reclassification. Gan et al. (2019) argue that this reclassification often times does not reflect the living conditions of communities appropriately. In particular, some "urban" households resemble standard rural households in consumption, housing, and access to facilities.

Noting that measurement of urban vs rural households is difficult makes the strong correlations that I have found in terms of asset-to-income ratios even more remarkable. One would think that the measurement error biases my results toward zero.

Descriptive Statistics of CFPS 2012

Table ?? and B2 display mean and standard deviations of the raw sample, i.e. before I restrict on age etc.

Table B2: Descriptive Statistic CFPS 2012

	rural household		urban household	
	mean	sd	mean	sd
household consumption expenditure per capita	8868.731	10542.57	17610.29	27079.3
household consumption expenditure, equivalent scale (OECD)	11860.51	13547.23	22907.85	34092.44
household income per capita	9492.302	11995.93	17810.05	33908.04
household income, equivalent scale (OECD)	12719.38	15293.52	22945.31	41534.85
age household head	42.47808	16.32895	46.16107	16.43448
number of years of education of household head	6.331158	4.595607	8.871506	4.780266
number of household members	4.087349	1.827104	3.473356	1.564111
number of kids in household	1.225067	1.257898	.9471144	1.10556
share of people in household in 60s or older	.1829483	.3023401	.2063425	.3363822
Net family assets(yuan)	186761.9	386411.4	518963.2	1041060
Observations	5976		4973	

Note: Mean and standard deviation for the raw sample, i.e. before selection. The data is from the CFPS wave 2012. The urban rural definition follows the CFPS community definition (*urbancomm*) which is more closely tied to the level of development of a village.

Table B3: Descriptive Statistic CFPS 2012

	No		Yes	
	mean	sd	mean	sd
household consumption expenditure per capita	18487.24	27900.23	8483.957	9225.321
household consumption expenditure, equivalent scale (OECD)	23868.54	35053.16	11505.4	12287.46
household income per capita	18073.89	33746.14	9454.978	12640.82
household income, equivalent scale (OECD)	23133.06	41336.33	12795.27	16179.06
age household head	47.60268	17.19248	41.2295	15.30528
number of years of education of household head	8.70587	4.980868	6.54345	4.508898
number of household members	3.253098	1.542216	4.246161	1.760075
number of kids in household	.8909202	1.081448	1.266253	1.25989
share of people in household in 60s or older	.2369411	.364973	.1569864	.2694647
Net family assets(yuan)	511603.4	1054988	203641.7	410487.4
Observations	4923		6122	

Note: Mean and standard deviation for the raw sample, i.e. before selection. The data is from the CFPS wave 2012.

B.10 Additional robustness for standard and quantile regressions

Table B4 shows the result from the CHIP for 1995. The asset-to-income ratio is systematically higher in non-agricultural activity. Those differences are not simply driven by income, age, and even survive controlling for education. These results arise during a time that is before China's pro-market reforms in state owned companies (He et al. 2018) or before China joins the WTO. The results are in line with De Magalhães and Santaeulàlia-Llopis (2018b) who show systematic urban rural differences. From a macro point of view I consider column one as most informative, as income and education rise endogenously as households move into the city, as least over the very long run. I use the variable agr_fam, which is a dummy that takes on the value of one if the household earned some income in agriculture, to characterize "agricultural households". Obviously, this is not a perfect match since many agricultural families also earn some income from non-agricultural activity.

It is also worth pointing out that urban-rural differences in the financial-asset-to-income ratios are robust to dropping stocks from the financial assets. Results are available upon request.

Table B4: Median regression using dummy for agricultural occupation for CHIP 1995

	(1) fin_asset_income	(2) fin_asset_income	(3) fin_asset_income	(4) fin_asset_income	(5) fin_asset_income
agr	-0.0872***	-0.0623***	-0.0473***	-0.0217	-0.0439***
	(0.0116)	(0.0139)	(0.0131)	(0.0145)	(0.0128)
_cons	0.421***	0.118	0.217*	0.0706	0.330***
	(0.00978)	(0.115)	(0.120)	(0.120)	(0.108)
income	No	Yes	Yes	Yes	Yes
age & demographics	No	Yes	Yes	Yes	Yes
education	No	No	No	Yes	Yes
province fe	No	No	No	No	Yes
N	10553	10553	10553	10486	10486

Note: The dependent variable is the household financial-asset-to-income-ratio as defined before. This contains bank deposits, and financial assets, but excludes company assets, housing, land, and other fixed (productive) assets. This data set for 1995 is from the CHIP. Standard errors in parentheses. *, **,*** denote statistical significance at 1, 5, and 10 percent level.

Table B5 offers results for asset-to-consumption ratio, which might be a better proxy for permanent income.

Table B7 reports the results in 2012 based on the CFPS. Much has changed in China but the estimates are relatively stable, with a lack of statistical significance in the second to last column. The results for the asset-to-consumption ratio are again reported in table B8 in the appendix turn out similar and significant for every specification. It also contains a specification with an urban-rural dummy instead of an agricultural occupation dummy in table B9, which is very similar to table B8. Since in the CHIP it is not always clear with "urban" refers to an urban areas as defined by the Chinese Census Bureau, or a non-agricultural hukou, I prefer to use occupational dummies that partition households into

Table B5: Median regression using dummy for agricultural occupation for CHIP 1995

	(1)	(2)	(3)	(4)
	no_prod_assets_consum	no_prod_assets_consum	no_prod_assets_consum	no_prod_assets_consum
agr	-0.188***	-0.0593***	-0.0521***	-0.0353**
	(0.0159)	(0.0179)	(0.0184)	(0.0179)
income_pc		0.0000573***	0.0000544***	0.0000475***
		(0.00000454)	(0.00000519)	(0.00000354)
income_pc_sq		-7.75e-10***	-6.64e-10***	-5.21e-10***
		(1.71e-10)	(1.52e-10)	(4.40e-11)
age		0.00766	0.00601	0.00574
		(0.00643)	(0.00666)	(0.00642)
age_sq		-0.0000814	-0.0000425	-0.0000314
		(0.0000754)	(0.0000792)	(0.0000763)
sex_hhead			-0.00863	-0.0145
			(0.0214)	(0.0206)
share_kids			0.0656*	0.0568
			(0.0390)	(0.0377)
share_retired			-0.0992	-0.102
			(0.0841)	(0.0793)
boys			-0.00170	0.000650
			(0.0157)	(0.0151)
familysize			-0.0192***	-0.0186***
			(0.00544)	(0.00517)
educ_year_hhead				0.0106***
				(0.00224)
_cons	0.575***	0.139	0.209	0.123
	(0.0140)	(0.134)	(0.138)	(0.135)
N	10553	10553	10553	10486

Note: The dependent variable is the household nonproductive-asset-to-consumption-ratio as defined before. This contains bank deposits, and financial assets, but excludes company assets, housing, land, and other fixed (productive) assets. Standard errors in parentheses. *, **,*** denote statistical significance at 1, 5, and 10 percent level.

Table B6: Median regression with urban-rural dummy in CFPS 2012

	(1) total_asset_income	(2) total_asset_income	(3) total_asset_income	(4) total_asset_income	(5) total_asset_income
urban_cfps	1.727***	2.571***	2.055***	1.476***	1.513***
	(0.213)	(0.263)	(0.223)	(0.241)	(0.255)
_cons	4.987***	5.684***	8.170***	6.574***	9.162***
	(0.0906)	(0.199)	(1.320)	(1.313)	(1.698)
income	No	Yes	Yes	Yes	Yes
age & demographics	No	No	Yes	Yes	Yes
education	No	No	No	Yes	Yes
province fe	No	No	No	No	Yes
N	6978	6978	6978	6977	6977

Note: The dependent variable is the household total-asset-to-income-ratio. This contains bank deposits, other financial assets, house property, company assets minus any type of debt. Financial derivatives are missing for the year 2010. Income is total household income(net_faminc). Standard errors in parentheses. *, **,*** denote statistical significance at 1, 5, and 10 percent level.

Table B7: Median regression with agricultural dummy for CFPS 2012

	(1) fin_asset_income	(2) fin_asset_income	(3) fin_asset_income	(4) fin_asset_income	(5) fin_asset_income
fam_agr	-0.179***	-0.180***	-0.160***	-0.109***	-0.0774***
	(0.0179)	(0.0189)	(0.0186)	(0.0177)	(0.0188)
_cons	0.362***	0.364***	0.142	-0.0691	0.0826
	(0.0160)	(0.0188)	(0.0930)	(0.0885)	(0.196)
income	No	Yes	Yes	Yes	Yes
age & demographics	No	No	Yes	Yes	Yes
education	No	No	No	Yes	Yes
province fe	No	No	No	No	Yes
N	7135	7135	7135	7134	7134

Note: The dependent variable is the household financial-asset-to-income-ratio as defined before. This contains bank deposits, and financial assets, but excludes company assets, housing, land, and other fixed (productive) assets. This data set for 2012 is from the CFPS. Robust standard errors in parentheses. *, **,*** denote statistical significance at 1, 5, and 10 percent level.

agricultural and non-agricultural employment.

Table B8: Median regression with agricultural dummy for CFPS 2012

	(1)	(2)	(3)	(4)	(5)
	fin_asset_consum	fin_asset_consum	fin_asset_consum	fin_asset_consum	fin_asset_consum
fam_agr	-0.169***	-0.0846***	-0.0784***	-0.0387*	-0.0314*
	(0.0189)	(0.0205)	(0.0217)	(0.0220)	(0.0178)
_cons	0.355***	0.205***	0.0758	-0.0936	0.0411
	(0.0167)	(0.0262)	(0.117)	(0.113)	(0.204)
income	No	Yes	Yes	Yes	Yes
age & demographics	No	No	Yes	Yes	Yes
education	No	No	No	Yes	Yes
province fe	No	No	No	No	Yes
N	6299	6299	6299	6299	6299

Note: The dependent variable is the household nonproductive-asset-to-consumption-ratio as defined before. This contains bank deposits, and financial assets, but excludes company assets, housing, land, and other fixed (productive) assets. This data set for 2012 is from the CFPS. Robust standard errors in parentheses. *, **,*** denote statistical significance at 1, 5, and 10 percent level.

Table B9: Median regression using urban-rural dummy for CFPS 2012

	(1)	(2)	(3)	(4)	(5)
	fin_asset_consum	fin_asset_consum	fin_asset_consum	fin_asset_consum	fin_asset_consum
urban_cfps	0.141***	0.0571***	0.0438*	-0.0109	0.0165
	(0.0205)	(0.0220)	(0.0242)	(0.0220)	(0.0204)
_cons	0.204***	0.129***	-0.0383	-0.160	0.0237
	(0.00915)	(0.0181)	(0.113)	(0.107)	(0.202)
income	No	Yes	Yes	Yes	Yes
age & demographics	No	No	Yes	Yes	Yes
education	No	No	No	Yes	Yes
province fe	No	No	No	No	Yes
N	6299	6299	6299	6299	6299

Note: The dependent variable is the household nonproductive-asset-to-consumption-ratio as defined before. This contains bank deposits, and financial assets, but excludes company assets, housing, land, and other fixed (productive) assets. This data set for 2012 is from the CFPS. Robust standard errors in parentheses. *, **,*** denote statistical significance at 1, 5, and 10 percent level.

Table B10: Median regression using urban-rural dummy for CFPS 2012

	(1)	(2)	(3)	(4)	(5)
	fin_asset_income	fin_asset_income	fin_asset_income	fin_asset_income	fin_asset_income
urban_cfps	0.184***	0.183***	0.157***	0.0884***	0.0837***
	(0.0201)	(0.0214)	(0.0207)	(0.0225)	(0.0219)
_cons	0.200***	0.200***	0.0205	-0.134	0.00992
	(0.00753)	(0.00895)	(0.0935)	(0.0919)	(0.197)
income	No	Yes	Yes	Yes	Yes
age & demographics	No	No	Yes	Yes	Yes
education	No	No	No	Yes	Yes
province fe	No	No	No	No	Yes
N	7135	7135	7135	7134	7134

Note: The dependent variable is the household total-asset-to-income-ratio. This contains bank deposits, other financial assets, house property, company assets minus any type of debt. Financial derivatives are missing for the year 2010. Income is total household income(net_faminc). Robust standard errors in parentheses. *, **,*** denote statistical significance at 1, 5, and 10 percent level.

Table B11: Median regression with agricultural dummy for CFPS 2012

	(1)	(2)	(3)	(4)	(5)
	total_asset_income	total_asset_income	total_asset_income	total_asset_income	total_asset_income
fam_agr	-1.282***	-1.886***	-1.667***	-1.128***	-0.437**
	(0.184)	(0.247)	(0.219)	(0.235)	(0.222)
_cons	6.282***	7.441***	9.067***	6.850***	8.853***
	(0.158)	(0.343)	(1.295)	(1.281)	(2.043)
income	No	Yes	Yes	Yes	Yes
age & demographics	No	No	Yes	Yes	Yes
education	No	No	No	Yes	Yes
province fe	No	No	No	No	Yes
N	6978	6978	6978	6977	6977

Note: The dependent variable is the household total-asset-to-income-ratio. This contains bank deposits, other financial assets, house property, company assets minus any type of debt. Financial derivatives are missing for the year 2010. Income is total household income(net_faminc). Robust standard errors in parentheses. *, **,*** denote statistical significance at 1, 5, and 10 percent level.

Table B12: Median regression with agricultural dummy for CFPS 2012

	(1) total_asset_consum	(2) total_asset_consum	(3) total_asset_consum	(4) total_asset_consum	(5) total_asset_consum
fam_agr	-0.963***	-0.325	-0.375*	0.155	0.551***
	(0.196)	(0.210)	(0.209)	(0.213)	(0.206)
_cons	6.076***	4.790***	7.193***	5.254***	1.534
	(0.171)	(0.246)	(1.376)	(1.496)	(2.476)
income	No	Yes	Yes	Yes	Yes
age & demographics	No	No	Yes	Yes	Yes
education	No	No	No	Yes	Yes
province fe	No	No	No	No	Yes
N	6201	6201	6201	6201	6201

Note: The dependent variable is the household total-asset-to-consumption-ratio. This contains bank deposits, other financial assets, house property, company assets minus any type of debt. Financial derivatives are missing for the year 2010. Income is total household income(net_faminc). Robust standard errors in parentheses. *, **,*** denote statistical significance at 1, 5, and 10 percent level.

B.11 Hukou switchers – Additional Results

Here I report descriptive statistics and additional results for the Hukou switchers. Results for the total wealth-to-income ratio, which turn out to be higher for switchers are available upon request.[195]

Table B13: Descriptive Statistic for Hukou Status in CFPS 2012

	hukou_agr_agr		hukou_agr_urban		hukou_urban_urban	
	mean	sd	mean	sd	mean	sd
household consumption expenditure per capita	10082.23	15529.74	17799.61	24754.3	22689.54	31117.13
household consumption expenditure, equivalent scale (OECD)	13500.53	20156.03	23086.47	31865.94	28951.57	37742.95
household income per capita	9796.117	14051.43	18212.69	28312.34	23072.67	47103.14
household income, equivalent scale (OECD)	13112.05	17984.91	23409.2	35018.48	29086.18	56957.55
age household head	43.2861	15.71596	51.08042	16.90349	47.20337	15.98333
number of years of education of household head	6.109705	4.458168	9.115974	5.168185	11.00873	3.923733
number of household members	3.92855	1.739628	3.422867	1.61176	3.084841	1.349264
number of kids in household	1.206311	1.234826	.9195842	1.103108	.7504679	.9817902
share of people in household in 60s or older	.1777834	.3043956	.2771765	.377298	.2131505	.3415762
Net family assets(yuan)	222587.1	474903.3	565912.9	1289694	659155.2	1046383
Observations	6718		1828		1603	

Note: Summary statistics for CFPS 2012 based on Hukou status of the household.

[195] The higher total-asset-to-income ratio is coming from the relatively lower income of switchers, and is driven by housing wealth.

Table B14: Hukou Median Regression for CFPS 2012

	(1)	(2)	(3)	(4)	(5)
	fin_asset_income	fin_asset_income	fin_asset_income	fin_asset_income	fin_asset_income
hukou_switcher	-0.138***	-0.102**	-0.0489	-0.0243	-0.0317
	(0.0504)	(0.0476)	(0.0503)	(0.0452)	(0.0443)
_cons	0.500***	0.395***	-0.432	-0.602	-0.504
	(0.0383)	(0.0489)	(0.369)	(0.375)	(0.512)
income	No	Yes	Yes	Yes	Yes
age & demographics	No	No	Yes	Yes	Yes
education	No	No	No	Yes	Yes
province fe	No	No	No	No	Yes
N	1539	1539	1539	1539	1539

Note: The dependent variable is the financial asset-to-income ratio. Sample selection is the same as before. *, **,*** denote statistical significance at 1, 5, and 10 percent level.

Table B15: Regression for CFPS 2012 – 2016

	(1)	(2)	(3)	(4)	(5)
	g_total_asset	g_total_asset	g_total_asset	g_total_asset	g_total_asset
hukou_switcher	0.0205	0.0165	0.0205	0.0234	0.00900
	(0.0144)	(0.0144)	(0.0146)	(0.0148)	(0.0151)
_cons	0.115***	0.0975***	0.327**	0.282**	0.348**
	(0.0102)	(0.0104)	(0.138)	(0.139)	(0.147)
income growth	No	Yes	Yes	Yes	Yes
age & demographics	No	No	Yes	Yes	Yes
education	No	No	No	Yes	Yes
province fe	No	No	No	No	Yes
N	1115	1110	1110	1110	1110

Note: The dependent variable is growth in total household wealth. *, **,*** denote statistical significance at 1, 5, and 10 percent level. Rural households as well as the largest 1% of asset growth rates are dropped.

Table B16: Regression for CFPS 2012 – 2016

	(1)	(2)	(3)	(4)	(5)
	g_arc_fin_asset	g_arc_fin_asset	g_arc_fin_asset	g_arc_fin_asset	g_arc_fin_asset
hukou_switcher	0.108*	0.0952	0.0948	0.105*	0.0790
	(0.0604)	(0.0605)	(0.0618)	(0.0624)	(0.0663)
_cons	0.640***	0.576***	1.398***	1.235**	1.193**
	(0.0397)	(0.0445)	(0.502)	(0.517)	(0.555)
income growth	No	Yes	Yes	Yes	Yes
age & demographics	No	No	Yes	Yes	Yes
education	No	No	No	Yes	Yes
province fe	No	No	No	No	Yes
N	1115	1110	1110	1110	1110

Note: The dependent variable is growth in total household wealth. *, **,*** denote statistical significance at 1, 5, and 10 percent level. Rural households as well as the largest 1% of asset growth rates are dropped.

B.12 Structural Change Regression Results

Table B17: Agr. Share Regression Results

	random effects	fixed effects
$\hat{\beta}$	-0.0444	-0.0446
SE	(0.00346)	(0.000689)
R^2	.96	.98

Appendices to Chapter III

C.1 HJB & Komolgorov Forward Equation

Household Problem/HJB Equation

Considering the household problem, it is useful to separate the differential operator $\mathcal{A}$ into three pieces, one endogenous piece relating to household saving choices $\mathcal{A}_s$ ($\mathcal{A}_s v = v_W \dot{W}$), one exogenous piece relating to the drift-diffusion process $\mathcal{A}_h$, an infinitesimal generator that captures exit from the high growth regime $\mathcal{A}_{\delta_{\mathrm{L}}}$, and lastly a component capturing intergenerational human capital risk $\mathcal{A}_\lambda$ so that $\mathcal{A}v = \mathcal{A}_s v + \mathcal{A}_h v + \mathcal{A}_{\delta_{\mathrm{L}}} v + \mathcal{A}_\lambda v$. It is important to note that the infinitesimal operator associated with the household problem is not the adjoint of the infinitesimal operator of the KFE. While it remains true that the adjoint of the first two elements, $\mathcal{A}_s$ and $\mathcal{A}_h$ are indeed the appropriate operators for the KFE, the last piece $\mathcal{A}_\lambda$ is not. The reason is that $\mathcal{A}_\lambda$ involves computing expectations over the population, which, though related, is not identical to the income and type evolution of the offspring generation. I first focus on $\mathcal{A}$ before turning to $\mathcal{A}_{KFE}$.

The saving choices, and the drift-diffusion term are standard and identical to Achdou et al. (2022a). The term capturing intergenerational risk is new and merits further discussion. Assuming that the consumption-saving choice maximizes the household problem, one can write the HJB equation in stacked vector notation. Differential operators are approximated using a Markov transition matrix together with finite differences (a grid) and difference quotients and I denote the Kronecker product using $\bigotimes$ and $\iota_{I_a \times I_h \times I_j}$ is a vector of size $I_W \times I_h \times I_j$ with all entries equal to one, where $I_W \times I_h \times I_j$ is the number of grid points in the finite difference approximation. The HJB reads

$$
\begin{aligned}
\rho v\left(W,h,j\right) = \max u &+ \left[\mathcal{A}_s + \mathcal{A}_h + \mathcal{A}_{\delta_{\mathrm{L}}}\right] v \\
&+ \lambda\left(1-\beta\right)\hat{\delta}\left[\chi v\left(W,\underline{h},\mathrm{H}\right) + \left(1-\chi\right) v\left(W,\underline{h},\mathrm{L}\right) - v\left(W,h,j\right)\right] \\
&+ \lambda\left(1-\beta\right)\left(1-\hat{\delta}\right)\left[\sum p\left(h,j\right)\left(v\left(W,h,j\right) - v\left(W,h,j\right)\right)\right]
\end{aligned}
$$

where the last term depends itself on the stationary marginal income distribution, $\int p\left(W,h,j\right) dF\left(W\right) = p\left(h,j\right)$. This means that in order to solve this household problem one needs to first solve for the stationary marginal income distribution. I take $p(h,j)$ as given for now and show below how to solve for it. I next write the HJB in stacked matrix notation, which ultimately is the system of equations on which to iterate to find a solution

$$\begin{aligned}\rho v &= u + \left[\mathcal{A}_s + \mathcal{A}_h + \mathcal{A}_{\delta_{\mathrm{L}}}\right] v \\ &\quad + \lambda(1-\beta)\left(\left(\left[\iota_{I_h\times I_j}\hat{\delta} * [1-\chi\, 0 \ldots 0\, \chi\, 0 \ldots 0] + \left(1-\hat{\delta}\right)\iota_{I_h\times I_j} p'_{hj}\right] - Diagm(1)_{I_h\times I_j}\right) \bigotimes M_{I_W}\right) v\end{aligned}$$

where $Diagm(1)_{I_h\times I_j}$ is a diagonal matrix of ones of size $I_h \times I_j$ and $\bigotimes M_{I_W}$ is a Kronecker product with a matrix of ones of size I_W. Note that as in any well-defined continuous time Markov transition matrix, the rows sum up to zero.

KFE Equation

Now I derive the KFE, defined as $\mathcal{A}_{KFE} = \mathcal{A}_s^* + \mathcal{A}_h^* + \mathcal{A}_{\delta_{\mathrm{L}}}^* + \mathcal{A}_{\lambda,KFE}^*$. While the first three elements are the adjoint of the infinitesimal generator (aka transpose) from the household problem, the last one $\mathcal{A}_{\lambda,KFE}^*$ is not. The appropriate generator related to the intergenerational human capital risk is

$$\mathcal{A}_{\lambda,KFE} = \left(\lambda(1-\beta)\,\hat{\delta}\left(\left[\iota_{I_h\times I_j} * [1-\chi\, 0 \ldots 0\, \chi\, 0 \ldots 0]\right] - Diagm(1)_{I_h\times I_j}\right) \bigotimes M_{I_W}\right).$$

The KFE reads

$$\dot{p} = \left(\mathcal{A}_s^* + \mathcal{A}_h^* + \mathcal{A}_{\delta_{\mathrm{L}}}^*\right) p - \lambda(1-\beta)\left(\hat{\delta}\right) p + \mathbf{1}_{p(\underline{h},\mathrm{L})}\lambda(1-\beta)\,\hat{\delta}\,(1-\chi) + \mathbf{1}_{p(\underline{h},\mathrm{H})}\lambda(1-\beta)\,\hat{\delta}\,(\chi), \tag{C1}$$

where $\mathbf{1}_{\underline{p}}$ is an indicator function that represents the additional inflow at the worst income state $\underline{p}$. Since the income shocks are exogenous, on can solve for the distribution of income without solving the household consumption-saving problem, essentially because income dynamics are exogenous. I describe the dynamics of income in the main text, and have deferred the derivation of the boundary conditions to the next paragraph.

Note that I can derive the law of motion within each growth rate regime, where the unconditional probability that a household is in the high growth regime is defined as $p_{\mathrm{L}} := \int p(W,h,\mathrm{L})\, dW dh$,

$$\dot{p}_{\mathrm{L}} = \lambda(1-\beta)\,\hat{\delta}\left[(1-\chi)\, p_{\mathrm{H}} - (\chi)\, p_{\mathrm{L}}\right] + \delta_{\mathrm{L}} p_{\mathrm{H}}. \tag{C2}$$

The first element in (C2) on the right hand side accounts for the inflow of new cohorts where I add former high-growth households that become low growth households at rebirth with probability $1-\chi$, and I subtract $(\chi)\, p_{\mathrm{L}}$ households that leave and enter the high growth regime at rebirth. Moreover, I account for the fact that a fraction $\delta_{\mathrm{L}} p_{\mathrm{H}}$ flows into the low growth regime every instant. As an aside, in the steady state the ratio of high growth to low growth households equals $\frac{p_{\mathrm{H}}}{P_{\mathrm{L}}} = \frac{\lambda(1-\beta)\hat{\delta}\chi}{\lambda(1-\beta)\hat{\delta}(1-\chi)+\delta_{\mathrm{L}}}$, which can be derived by setting $\dot{p}_{\mathrm{L}} = 0$ and solving (C2).

Next, note that

$$\int \frac{\partial p\,(h,\mathrm{H})}{\partial t} dh = + \left[\mu_t^{\mathrm{H}} \underline{h} p\,(h,\mathrm{H})\right] - \frac{\partial}{\partial h}\left[\frac{\left(\sigma_t^{\mathrm{H}}\right)^2}{2} \underline{h}^2 p\,(\underline{h},\mathrm{H})\right] - \lambda\,(1-\beta_t)\,\hat{\delta} p_{\mathrm{H}} - \delta_{\mathrm{L}} p_{\mathrm{H}},$$

given regularity conditions, in particular that the first moment is well-defined so that $\lim_{h\to\infty} h^2 p\,(h,\mathrm{H}) = 0$. Now I can add up $\dot{p}_{\mathrm{H}} + \dot{p}_{\mathrm{L}}$, which has to equal zero.[196] Thus

$$\dot{p}_{\mathrm{L}} + \dot{p}_{\mathrm{H}} = \left[\mu_t^{\mathrm{H}} \underline{h} p\,(h,\mathrm{H})\right] - \frac{\partial}{\partial h}\left[\frac{\left(\sigma_t^{\mathrm{H}}\right)^2}{2} \underline{h}^2 p\,(\underline{h},\mathrm{H})\right] - \lambda\,(1-\beta_t)\,\hat{\delta} p_{\mathrm{H}} - \delta_{\mathrm{L}} p_{\mathrm{H}} + \lambda\,(1-\beta)\,\hat{\delta}\left[(1-\chi)\, p_{\mathrm{H}} - (\chi)\, p_{\mathrm{L}}\right.$$

which simplifies to

$$0 = -\left[\mu_t^{\mathrm{H}} \underline{h} p\,(h,\mathrm{H})\right] + \frac{\partial}{\partial h}\left[\frac{\left(\sigma_t^{\mathrm{H}}\right)^2}{2} \underline{h}^2 p\,(\underline{h},\mathrm{H})\right] + \lambda\,(1-\beta_t)\,\hat{\delta}\chi,$$

and is identical to the boundary condition in the main text. The case for the low growth regime is analogous.

C.2 Intergenerational Risk

C.2.1 Intergenerational elasticity

I want to prove that a regression of child on parent income, for the case with $\mu_j = \sigma_j = 0$, identifies the coefficient β, which thus represents an intergenerational elasticity. Consider the regression model

$$\log y_i' = \mu + \beta_{IGE} \log y_i + u_i,$$

[196] Since the probabilities have to add up to unity, I have $\dot{p}_{\mathrm{L}} + \dot{p}_{\mathrm{H}} = \frac{\partial}{\partial t}\left[p_{\mathrm{L}} + p_{\mathrm{H}}\right] = \frac{\partial}{\partial t}[1] = 0$.

where y' represents offspring income, u is an error term, and β_{IGE} is the regression coefficient known as intergenerational elasticity. The OLS formula implies

$$\beta_{IGE} = \frac{Cov\left(\log y_i', \log y_i\right)}{Var\left(\log y_i\right)}.$$

Using the stochastic process postulated in the paper, see equation (3.4), and using the law of total expectations and the fact that $\mathbb{E}\log y_i' = \mathbb{E}\log y_i = \mu$ in a stationary equilibrium, I can write

$$\begin{aligned}
\beta_{IGE} &= \frac{Cov\left(\log y_i', \log y_i\right)}{Var\left(\log y_i\right)} \\
&= \frac{\mathbb{E}\left[\left(\log y_i' - \mu_y\right)\left(\log y_i - \mu_y\right)\right]}{Var\left(\log y_i\right)} \\
&= \frac{\beta\mathbb{E}\left[\left(\log y_i' - \mu_y\right)\left(\log y_i - \mu_y\right)|X_i = 1\right] + (1-\beta)\,\mathbb{E}\left[\left(\log y_i' - \mu_y\right)\left(\log y_i - \mu_y\right)|X_i = 0\right]}{Var\left(\log y_i\right)} \\
&= \frac{\beta\mathbb{E}\left[\left(\log y_i - \mu_y\right)\left(\log y_i - \mu_y\right)|X_i = 1\right] + (1-\beta)\,\mathbb{E}\left[\left(\log u - \mu_y\right)\left(\log y_i - \mu_y\right)|X_i = 0\right]}{Var\left(\log y_i\right)} \\
&= \frac{\beta\mathbb{E}\left[\left(\log y_i - \mu_y\right)\left(\log y_i - \mu_y\right)|X_i = 1\right] + 0}{Var\left(\log y_i\right)} \\
&= \beta
\end{aligned}$$

where the penultimate line uses the fact that the new draw is independent of the parents' human capital for $X_i = 0$.

How does this carry over to the more general version in the paper where households are hit by persistent shocks permanently? Simulate and run regression. Even verify the first result with a simulation.

C.2.2 Intergenerational Risk and Non-Linear Income Distribution

One way to deal with the distributional consequences is to have a completely random draw from the distribution. A better way, especially from an empirical point of view, is to allow for some human capital destruction as the new generation takes over. That is, on average, children, at the beginning of their career, earn less than their parents. A simple way to introduce this aspect into the model is to let households draw from a tilted distribution of the form

$$P\left(y' < x|j\right) = \left[F_{y|j}\left(x\right)\right]^{\kappa}$$
$$\Leftrightarrow$$
$$\psi\left(x|j\right) = \kappa\left[F_{y|j}\left(x\right)\right]^{\kappa-1} f_{y|j}\left(x\right)$$

with $\kappa \leq 1$ and each distribution is conditional on the growth type, $j \in \{\mathrm{L}, \mathrm{H}\}$. I maintain that at birth, individuals enter the high growth regime with Bernoulli probability χ, exit from the high growth regime happens at rate δ_{L} and low growth is an absorbing state until the household is reborn. For $\kappa = 0$, all mass ends up on the worst income state within each type.[197] And the case of $\kappa = 1$ is identical with $\delta = 0$. Using this formulation, the KFE read as follows

$$0 = transpose\left(A - \lambda\left(1-\beta\right)I + \lambda\left(1-\beta\right)\left[\begin{array}{cc} \left(1-\chi\right)\psi_{\mathrm{L}} & \chi\psi_{\mathrm{H}} \\ \ldots & \ldots \\ \underbrace{\left(1-\chi\right)\psi_{\mathrm{L}}}_{1\times N/2} & \underbrace{\chi\psi_{\mathrm{H}}}_{1\times N/2} \end{array}\right]\right)p \tag{C3}$$

where $\psi_{\mathrm{L}} = \kappa\left(p_{\mathrm{L}}./\sum_i p_{\mathrm{L},i}\right)\left\{\underbrace{p'_{\mathrm{L}}\left[\begin{array}{ccc} 1 & \ldots & 1 \\ 0 & 1 & \ldots \\ 0 & 0 & 1 \end{array}\right]./\sum_i p_{\mathrm{L},i}}_{CDF}\right\}^{\kappa-1}$ is a vector that contains the conditional probability distribution and $\left(1-\chi\right)\psi_{\mathrm{L}}$ means that the scalar $\left(1-\chi\right)$ is multiplied which each entry of the vector. I make sure that $\sum \psi_{\mathrm{L},i} = 1$, i.e. I normalized by the sum if the conditional distribution does not add up to unity.[198] Note that$A$ contains the drift-diffusion terms as well as the exit from the high growth rate. The ranks of matrices are $R\left(A\right) = R\left(I\right) = N \times N, R\left(p\right) = N \times 1$ and note that UT is an upper triangular matrix of ones, with rank $N/2$. The notation ./ means a point-wise operation where every entry of the

[197]One natural extension is to make the probability χ a function of income, say starting a company requires starting capital so offspring from low-income families cannot pursue a high-growth strategy. I do not pursue this here but refer the reader to a large literature where productive investment and borrowing constraints interact to cause aggregate productivity losses, see Galor and Zeira (1993) for example.

[198]I have derived the shape of ψ using a continuous distribution, so the normalization helps overcome numerical discrepancy that arise as I move form a continuous to a discrete distribution.

vector or matrix is divided through by some scalar. Going from the overall distribution p to conditional distributions for each type requires normalizing by the total probability of being of said type.

This is a non-linear system of equations which cannot be solved by inverting the adjoint of the transition matrix. The reason is that the probabilities ψ themselves depend on the stationary distribution p. This fixed point problem, however, can be solved for instance through iteration. Specifically, I use the fact that

$$p_{n+1} = \left(I - \Delta \widetilde{A}_n\right)^{-1} p_n \tag{C4}$$

where $\widetilde{A}$ is the matrix in (C3) right before p. One can now simulate the stationary distribution simply by iterating on (C4), where it is clear that $\widetilde{A}$ is a function of p_n. The equation essentially simulates the transitions in time, and thus converges to a stationary distribution. Any initial guess with probabilities that sum up to unity works.

At this point, it is clear why I like to focus on the two special cases with $\kappa = 0$ and $\kappa = 1$ as in either case equation (C3) would become linear in the sense that the matrix would NOT be a function of p. The fact that sampling from the distribution renders the differential operator associated with the KFE nonlinear has been made in Gabaix et al. (2016). Inverting the matrix is much faster than iterating on (C4). To make matters worse, one would have to employ this iteration procedure at every step along the transition path. At this point, the non-linearity begins to impose a substantial computational burden as there are easily 1500 time steps, (150 years with .1 time increments) each of which requires finding a time-dependent solution to (C4).

C.3 Special Cases

Derivation of equation (3.20): Assume that the household is at an interior solution, and note that $c = y$ in virtue of the maximally tight borrowing constraint. Then

$$
\begin{aligned}
\frac{\dot{c}}{c} &= \frac{r-\rho}{\gamma} + \frac{\lambda(1-\beta)}{\gamma}\mathbb{E}_{c'}\left[\left(\frac{c'}{c}\right)^{-\gamma} - 1\right] + \frac{\phi_t}{\gamma}\left[\left(\frac{c_0}{c}\right)^{-\gamma} - 1\right] \\
\mu_{\mathrm{L}} &= \frac{r-\rho}{\gamma} + \frac{\lambda(1-\beta)}{\gamma}\left(\chi\left[\left(\frac{y_{\mathrm{H}}}{y_{\mathrm{L}}}\right)^{\gamma} - 1\right] + (1-\chi)(1-1)\right) + \frac{\phi_t}{\gamma}[1-1] \\
&\Leftrightarrow \\
r &= \rho + \lambda(1-\beta)\chi\left[\left(\frac{y_{\mathrm{H}}}{y_{\mathrm{L}}}\right)^{\gamma} - 1\right]
\end{aligned}
$$

where i used the fact that the household experiences no income growth, and that there is no effective disaster risk for poor households. They will be equally poor in the disaster scenario. Derivation of equation (3.21):

Assume that the household is at an interior solution, and note that $c = y$ in virtue of the maximally tight borrowing constraint. Then

$$
\begin{aligned}
\frac{\dot{c}}{c} &= \frac{r-\rho}{\gamma} + \frac{\lambda(1-\beta)}{\gamma}\mathbb{E}_{c'}\left[\left(\frac{c'}{c}\right)^{-\gamma} - 1\right] + \frac{\phi_t}{\gamma}\left[\left(\frac{c_0}{c}\right)^{-\gamma} - 1\right] \\
\mu_{\mathrm{H}} &= \frac{r-\rho}{\gamma} + \frac{\lambda(1-\beta)}{\gamma}\mathbb{E}_{c'}\left[\left(\frac{c'}{c}\right)^{-\gamma} - 1\right] + \frac{\phi_t}{\gamma}\left[\left(\frac{c_0}{c}\right)^{-\gamma} - 1\right] \\
&\Leftrightarrow \\
\gamma\mu_{\mathrm{H}} &= r - \rho + \lambda(1-\beta)\mathbb{E}_{c'}\left[\left(\frac{c'}{c}\right)^{-\gamma} - 1\right] + \phi_t\left[\left(\frac{c_0}{c}\right)^{-\gamma} - 1\right] \\
\gamma\mu_{\mathrm{H}} &= r - \rho + \lambda(1-\beta)\left[\chi(1-1) + (1-\chi)\left[\left(\frac{y_{\mathrm{H}}}{y_{\mathrm{L}}}\right)^{\gamma} - 1\right]\right] + \phi_t\left[\left(\frac{y_{\mathrm{H}}}{y_{\mathrm{L}}}\right)^{\gamma} - 1\right] \\
\gamma\mu_{\mathrm{H}} &= r - \rho + \lambda(1-\beta)(1-\chi)\left[\left(\frac{y_{\mathrm{H}}}{y_{\mathrm{L}}}\right)^{\gamma} - 1\right] + \phi_t\left[\left(\frac{y_{\mathrm{H}}}{y_{\mathrm{L}}}\right)^{\gamma} - 1\right] \\
\gamma\mu_{\mathrm{H}} &= r - \rho + \lambda(1-\beta)(1-\chi)[\Gamma^{\gamma} - 1] + \phi_t[\Gamma^{\gamma} - 1]
\end{aligned}
$$