Structural Change, Inequality, and Capital Flows*

Florian Trouvain
University of Michigan

April 2022

Abstract

High saving rates in fast growing "miracle" economies and the associated capital outflows have long been a puzzle in the international economics literature. I provide evidence that the demand for safe assets is systematically higher for urban (non-agricultural) relative to rural (agricultural) households suggesting a strong precautionary savings motive in urban areas. I combine this with the insight that miracle economies display fast structural change out of traditional farming. The interplay of structural change and rising demand of safe assets of urban households can account for the puzzling capital outflows during the growth miracle. I then develop a tractable model of miracle growth and human capital risk that rationalizes these findings. The key ingredients of the model are structural transformation away from traditional agricultural production, a heterogeneous income growth experience of households in the urban sector, and initial uncertainty about a household's success in the urban economy. The model characterizes in closed form the trade-off between consumption smoothing and precautionary savings, and offers a simple sufficient statistic to sign the direction of capital flows along the transition path.

Keywords: Capital Flows, Structural Change, Inequality, Human Capital Risk, Urban Rural differences

*Many thanks to my advisors John Leahy, Linda Tesar, and Dimitriy Stolyarov. I thank Paco Buera, Michael Koch, Rui Zhang, Rudi Bachmann, Hoyt Bleakley, Matthew Shapiro, Brit Sharoni, Wenjian Xu, Hongwei Xu, Pablo Ottonello, Steve Lugauer, Chris Carroll, Damiano Sandri, Andrei Levchenko, Sebastian Sotelo, Rachel Ngai, Karol Mazur, and John Laitner for their insights, as well as comments from the audience of the SED 2021 Annual meeting, the European Winter Meeting of the Econometric Society 2021, the Indian Winter Meeting of the Econometric Society 2021, as well as the CEPR-STEG Theme 3 workshop 2021. Funding from the Michigan Ross China Research Initiative is gratefully acknowledged.
Email: ftrouv@umich.edu.

1 Introduction

In a seminal article 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 et al., 2005; Gourinchas and Jeanne, 2013; Prasad et al., 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.[1] 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 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,

[1] See the recent paper by Coeurdacier et al. (2019) for a quantitative model that features consumption smoothing along the transition path.

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 et al., 2000).[2] 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.[3] In the model at hand the importance of the 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 et al., 2008; Mendoza et al., 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 reviews relevant literature. Section 3 provides a set of stylized facts relating to miracle growth, structural change, and urban-rural differences. Section 4 develops a simple model that connects those facts

[2]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 (Chamon et al., 2013).

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

and studies the tradeoff between catch-up growth and risk. Section 5 provides a simple quantification of the model. Section 6 concludes.

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.[4] It is important to note, however, that the residual component of household income fluctuations a la Blundell et al. (2008) is usually not sufficient to generate capital outflows during a growth miracle. Coeurdacier et al. (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,[5] 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 et al. (2014).

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. Lucas (2004) models the connection between development, urban-rural migration, and human capital accumulation.

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

[5]Santaeulalia-Llopis and Zheng (2018) use the method of Blundell et al. (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.

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 et al., 2014; Morten, 2019; Lagakos et al., 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 et al., 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.[6] Differences in total wealth accumulation between urban and rural households in Subsaharian 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 et al., 2008; Mendoza et al., 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 et al. (2007, 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.[7] Another central paper in this field is Song et al. (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 not involved in entrepreneurial activity.[8] 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, 2018) argue how the one-child policy in China can lead to savings pressure

[6]There is a current debate to what extent the urban rural wage gap reflects selection (Lagakos and Waugh, 2013; Young, 2013; Gollin et al., 2014; Hicks et al., 2017; Lagakos 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.

[7]Caballero et al. (2008) and Mendoza et al. (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.

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

and capital outflows. Wei and Zhang (2011) posit that the high Chinese saving rates are driven by the gap in the sex ratio, and Curtis et al. (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.

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.[9]

3.1 Macro Facts

Figure 1 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. The countries that drive this negative correlation are also known as "miracle economies", a term coined by 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.[10] As mentioned in the introduction, fast

[9]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 et al. (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 out what concepts are used where in the empirical work.

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

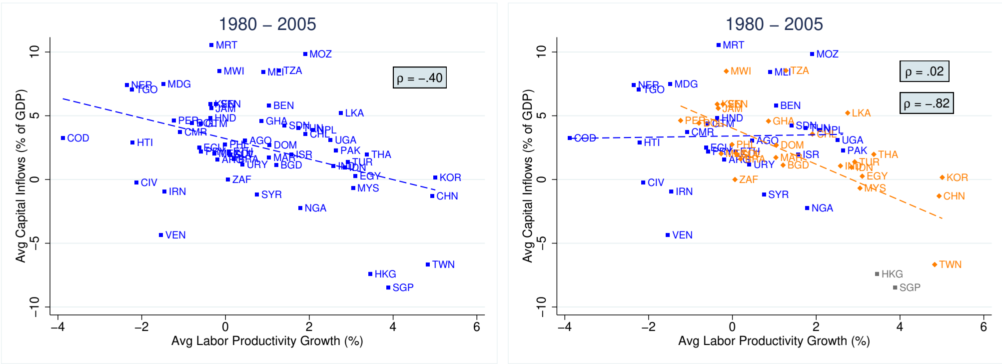

Figure 1: Relationship between average capital inflows and average labor productivity growth for a cross section of emerging markets following Gourinchas and Jeanne (2013). Orange diamonds represent emerging markets with above average decline of the agricultural employment share. Current-account-to-GDP 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.

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 2 highlights 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).[11]

Figure 3 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 15, 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 et al., 2005).[12] Based 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.[13] Household

[11] The recent handbook chapter by Herrendorf et al. (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.

[12] 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.

[13] The measurement of global capital flows is challenging (Coppola et al., 2020), but the qualitative finding

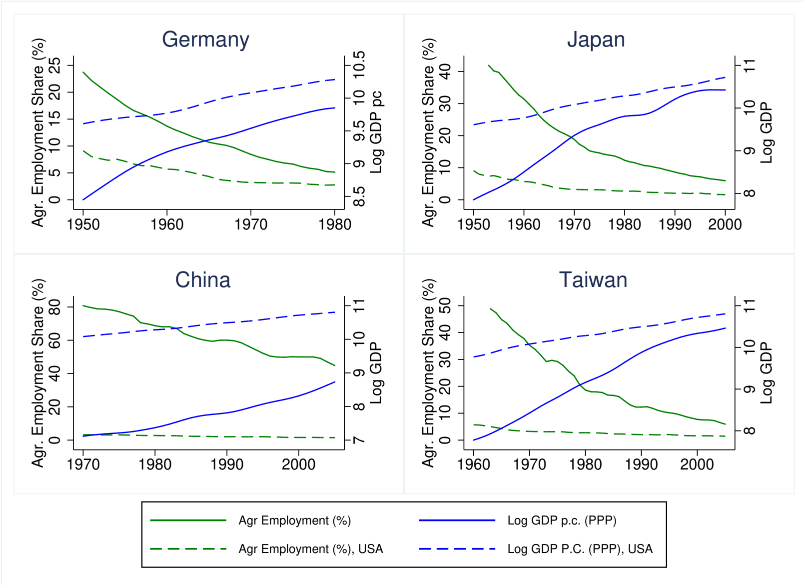

Figure 2: Relationship between agricultural employment share and convergence in GDP for Germany, Japan, China, and Taiwan. GDP series in purchasing power parity taken from the Penn World Tables 9.1. GDP is smoothed using an hp-filter with smoothing parameter of 8.5.

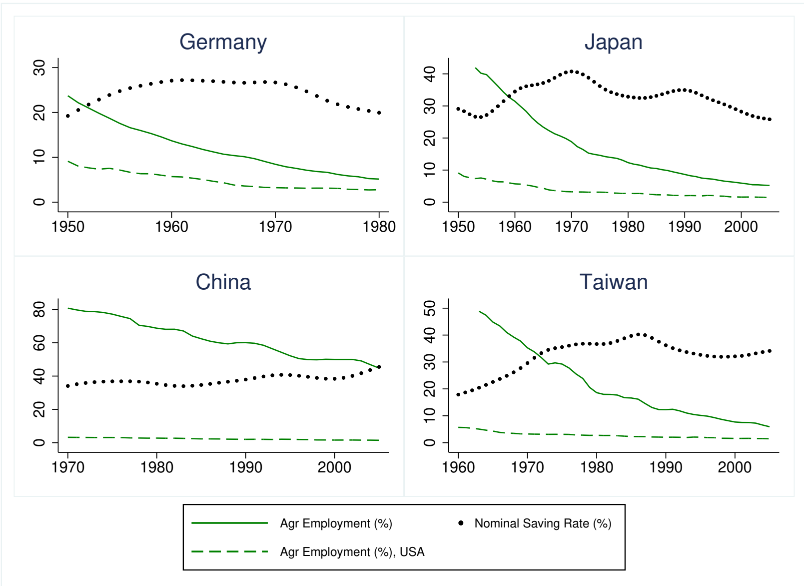

Figure 3: Relationship between agricultural employment share and aggregate nominal saving rate for Germany, Japan, China, and Taiwan. Saving rate is smoothed using an hp-filter with smoothing parameter of 8.5.

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.[14]

that growth miracles are associated with capital outflows is robust (Gourinchas and Jeanne, 2013).

[14]To corroborate the relationship between structural change and savings pressure, I offer additional evidence from cross country regressions in the appendix 8.7.

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 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.[15] 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.

3.2.1 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.[16] 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 zero assets, then it immediately follows that the asset-to-income ratio reads

[15] 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.

[16] 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.

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

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.[17]

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

$$\frac{a^i}{y^i} = \alpha + \beta D_i + \Gamma' X_i + \epsilon_i \tag{3.1}$$

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 between 23 and 60 years old, in line with previous work (He et al., 2018; Storesletten et al., 2004). Additional details on sample selection is provided in the appendix in section 8.4. 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 8.4. My preferred variable to understand urban rural differences is $urban\_cfps$ which is a community based measure that groups villages into

[17] 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 et al. (2003), Carroll and Samwick (1997), Fuchs-Schündeln and Schündeln (2005) or He et al. (2018).

urban and rural areas provided by the CFPS. This measure is different from the Census Bureau's definition. In the appendix in subsection 8.4 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 et al., 2008; Mendoza et al., 2009) and the theoretical model in the next section. Table 5 and table 6 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 4 plots the regression coefficient for rural ($\hat{\alpha}$) and urban ($\hat{\alpha}+\hat{\beta}$) households, based on equation 3.1, 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.

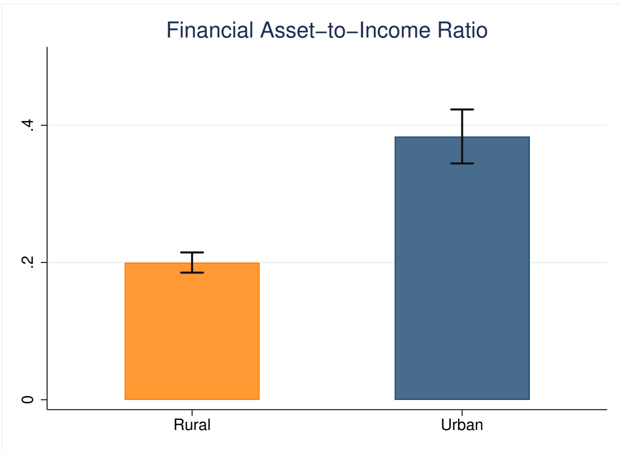

Figure 4: Cross sectional Urban-rural median differences in financial asset-to-income ratios for CFPS 2012 with 95% confidence intervals. See table 21 for additional information.

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 Zhang, 2011). In table 21 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,[18] 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.

[18]This includes the sex of the household head, the share of household members over 60, as well as whether

Table 1: Median regression with urban-rural dummy for CFPS 2012

	(1) fin_asset_income	(2) fin_asset_income	(3) fin_asset_income	(4) fin_asset_income	(5) 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 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.

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[19]

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

3.3 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.[20] The harder problem, however, is to write down a model with forward-looking agents that allow

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.

[19]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.

[20]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.

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 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 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 has been documented by a number of papers in the literature (Townsend, 1994; Santaeulalia-Llopis and Zheng, 2018; De Magalhães and Santaeulàlia-Llopis, 2018a).[21] 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 et al. (2008) is not going to be powerful enough.[22] 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.[23]

Figure 5 plots the rise in wage-inequality for the case of urban China.[24] The key question is whether rising inequality represents ex-ante risk during a growth miracle, which leads to

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

[22] 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 et al. (2019).

[23] The work by Buera and Shin (2017) or Song et al. (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.

[24] 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.

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 5 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 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. While the cross-sectional plot isn't really informative, I use the panel dataset constructed

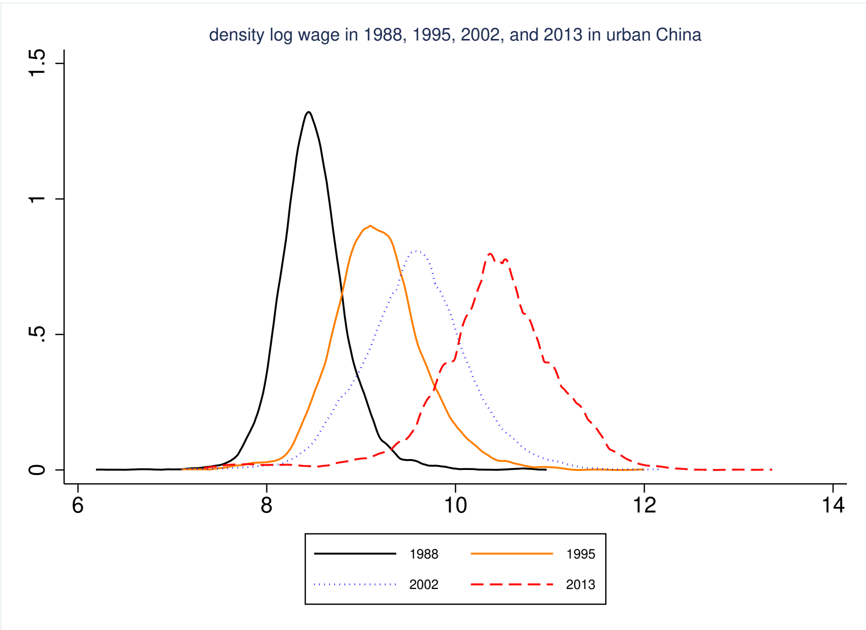

Figure 5: Data based on Chinese Household Income Project (CHIP), see Shi (2009). Density plot of log of real wage income for fulltime male household heads between the age of 23 and 60 in urban China in 1988, 1995, 2002, and 2013.

by Santaeulalia-Llopis and Zheng (2018) in figure 6 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.[25] The same holds true for household consumption including expenditures on food, utilities, health, and semidurable supplies. It seems indeed possible that

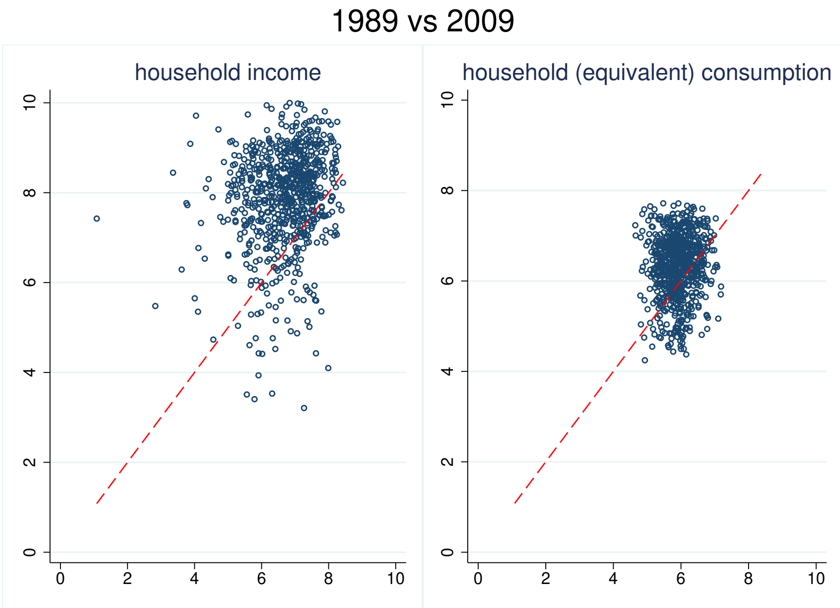

Figure 6: Log of household income and household food consumption in equivalent units based on the China Health and Nutrition Survey (CHNS). Data directly taken from Santaeulalia-Llopis and Zheng (2018), see their paper for details.

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

The second key ingredient in the model is that growth itself is unevenly distributed. In subsection 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.[26] 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.[27] 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.

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

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

[26] Piketty et al. (2019) document how top income inequality has shot up during the Chinese growth miracle, at a rate that is unprecedented in modern history.

[27] 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 Chamon et al. (2013) that are much closer to estimates in the US.

displayed in section 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

$$\max_{[H_t^u]} A_t H_t^u - w_t H_t^u, \tag{4.1}$$

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,

$$Y_t^r = A_0^r \left(L_t^r\right)^\alpha. \tag{4.2}$$

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.[28] 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 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^*.

$$A_t^u = A_0^u \exp(g^* t) \tag{4.3}$$

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 (Lucas, 2018).[29] 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

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

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

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^i_m 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 $t < T_i$

$$\begin{aligned} y^i_t =& w_t h^i_t \\ y^i_t =& w_t \exp((g - g^*)[t - t^i_m]). \end{aligned} \tag{4.4}$$

Then, using (4.1), the log of income for household i equals

$$\log(y^i_t) = \begin{cases} \log(A_t) + [t - t^i_m](g - g^*) & \text{if } t \leq T^i \\ \log(A_t) + [T^i - t^i_m](g - g^*) + \log(\varphi^i) & \text{if } t > T^i \end{cases} \tag{4.5}$$

which, in terms of growth rates reads

$$\frac{d\log(y^i_t)}{dt} = \begin{cases} g_h & \text{if } t < T^i \\ g^* & \text{if } t > T^i \end{cases} \tag{4.6}$$

for agents in the city with

$$F(T^i - t^i_m) \sim exponential(\lambda). \tag{4.7}$$

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

$$w^r_t = c^r_t \quad \forall t, \forall i \in S_r. \tag{4.8}$$

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

$$\dot{a}_t = \begin{cases} [r^* a_t + y_t^0(t_m) - c_t]\, dt & \text{if } t < T \\ [r^* a_t + \varphi^i y_t^1(t_m, T) - c_t]\, dt & \text{if } t \geq T. \end{cases} \tag{4.9}$$

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 7, 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).[30] 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.

[30] See Herrendorf et al. (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).

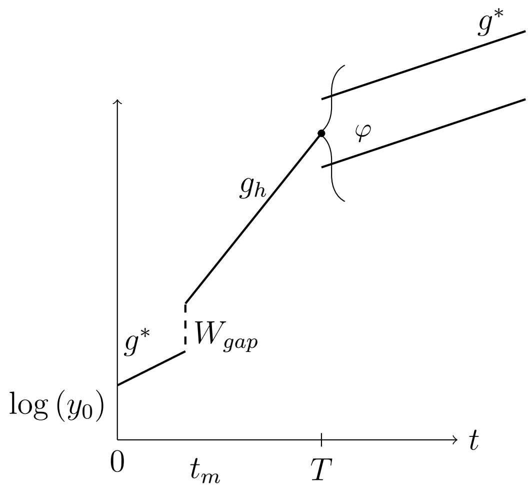

Figure 7: Income process from household starting in the agricultural sector

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 but in principal all household variables should be indexed by i

$$\max_{[c_t, a_t, t_m]} \mathbb{E}_{\varphi,T}\left[\int_0^\infty \exp(-\rho t)\frac{(c_t)^{1-\eta}}{1-\eta}dt\right]. \tag{4.10}$$

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 et al. (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”.[31] 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 2 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

[31] 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.

status as an illegal immigrant in the United States (Piketty et al., 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.[32]

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 at every point in time. To avoid counterfactual implications for the urban-rural wage gap W_{gap}, I introduce a migration cost.[33] The cost is paid in utility terms and proportional to the utility associated with moving to the city.[34] 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

$$\Delta \frac{(w_t^r)^{1-\eta}}{1-\eta} + \tau^{\eta-1}(1-\Delta\rho)V_{t+\Delta} = \tau^{\eta-1}V_t. \tag{4.11}$$

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

$$\frac{(w_t^r)^{1-\eta}}{1-\eta} = \tau^{\eta-1}\left(\rho V_t - \dot{V}_t\right). \tag{4.12}$$

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 (4.11) 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

[32] The hukou system is complex and has seen multiple reforms since 1980. In the appendix in section 8.4 I discuss the hukou system in a little more depth.

[33] 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).

[34] 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.

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

- households maximize utility given (4.10),(4.12), 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 (4.1),(4.2)
- 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 (7.13), (4.16)
- labor markets (4.15), (4.16) clear, and goods markets are consistent with asset markets (4.17)
- the no-Ponzi-scheme condition (4.13) is satisfied, i.e.

$$\lim_{t\to\infty} \exp(-r^* t) a_t^i \geq 0, \forall i \in [0,1]. \tag{4.13}$$

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

$$Y_t^r = (L_t^r)^\alpha \tag{4.14}$$

$$\begin{aligned} Y_t^u = {} & A_t \int_{i\in M_0} \int_{t_{m,i}}^{t} \exp((g-g^*)[s - t_{m,i}]) ds di \\ & + A_t \int_{i\in M_1} \varphi_i \int_{t_{m,i}}^{T_i} \exp((g-g^*)[s - t_{m,i}]) ds di. \end{aligned} \tag{4.15}$$

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

$$L_t^u + L_t^r = 1. \tag{4.16}$$

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

$$\int_{i\in S_u} \dot{a}_i di = Y_t^u + r^* A_t^b - C_t^u, \tag{4.17}$$

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

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

$$\frac{\dot{c}_s}{c_s} = \frac{1}{\eta}\left(r^* - \rho\right). \tag{4.18}$$

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

$$\begin{aligned} V\left(a_T; \varphi, T\right) &= \int_T^\infty \exp\left(-\rho\left(s-T\right)\right) \frac{c_s^{1-\eta}}{1-\eta} ds \\ &= \int_T^\infty \exp\left(-\rho\left(s-T\right)\right) \frac{\left(y_s + \left[g^*\left(\eta-1\right)+\rho\right] a_s\right)^{1-\eta}}{1-\eta} ds \\ &= \frac{\left(y_T + \left[g^*\left(\eta-1\right)+\rho\right] a_T\right)^{1-\eta}}{1-\eta} \left\{\frac{1}{g^*\left(\eta-1\right)+\rho}\right\}. \end{aligned} \tag{4.19}$$

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 4.20

$$V_{t_m} = \max_{c_s} \int_{t_m}^\infty exp\left(-\left(\lambda+\rho\right)\left[s - t_m\right]\right) \left[\frac{c_s^{1-\eta}}{1-\eta} + \lambda \mathbb{E}_\varphi\left[V\left(\varphi y_s, a_s\right)\right]\right] ds, \tag{4.20}$$

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:

$$H = \exp\left(-\left(\lambda+\rho\right)s\right)\left\{\frac{c_s^{1-\eta}}{1-\eta} + \lambda\mathbb{E}_{\varphi}\left[V\left(\varphi y_s, a_s\right)\right]\right\} + q_s\left[r^* a_s + y_s - c_s\right]. \tag{4.21}$$

After plugging (4.19) into (4.21), the optimality conditions with respect to c_s and a_s read

$$dc: \quad q_s = \exp\left(-\left(\lambda+\rho\right)s\right)\left(\frac{1}{c_s}\right)^{\eta} \tag{4.22}$$

$$da: \quad -\dot{q} = \exp\left(-\left(\lambda+\rho\right)s\right)\lambda\mathbb{E}_{\varphi}\left[\left(\frac{1}{\left[g^*\left(\eta-1\right)+\rho\right]+\varphi y_s}\right)^{\eta}\right] + q_s\left[r^*\right]. \tag{4.23}$$

Taking logs of (4.22) and differentiating with respect to s, plugging in (4.23), 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

$$\frac{\dot{c}_s}{c_s} = \frac{\lambda}{\eta}\left\{\mathbb{E}_{\varphi}\left[\left(\frac{\left[g^*\left(\eta-1\right)+\rho\right]a_s + \varphi y_s}{c_s}\right)^{-\eta}\right] - 1\right\} + g^*. \tag{4.24}$$

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 (4.24) 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 (4.24).

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 (4.24) 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.[35]

Proposition 1. *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 $\frac{a_t}{y_t}\left[\left(\eta-1\right)g^*+\rho\right] > -\underline{\varphi}$.*

Proof. By contradiction, suppose an agent borrows above the borrowing level. Then there is a range of values $\varphi \in [\underline{\varphi}, -\frac{a_t}{y_t}\left[\left(\eta-1\right)g^*+\rho\right]]$ where the agent would have to consume weakly below zero. Furthermore, assuming that $\int_{\underline{\varphi}}^{-\frac{a_t}{y_t}\left[\left(\eta-1\right)g^*+\rho\right]} dF(\varphi) = \epsilon > 0$, then such

[35] 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.

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 1 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.[36]

Proposition 2. *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^*.*

Proof. 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.[37] 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 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.[38]

Proposition 3. *A model without income inequality, i.e. $\varphi_i = 1, \forall i$, cannot generate capital outflows.*

Proof. See appendix. □

Proposition 3 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

[36]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 1. 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.

[37]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.

[38]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.

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.[39]

Proposition 4. *Without human capital risk, there are no capital outflows, even if expectations about convergence are downward biased, $\tilde{\lambda} > \lambda$.*

Proof. See appendix. □

This proposition makes clear that biased expectations per se are not a solution to the capital flow puzzle.[40] 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 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.

Proposition 5. *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*

$$\frac{g_h - g^*}{\lambda}\eta < \mathbb{E}_\varphi\left[\left(\frac{1}{\varphi}\right)^\eta\right] - 1 \tag{4.25}$$

[39]Recall that average time spend in the high growth regime is inversely related to the arrival rate.
[40]Thanks to Pablo Ottonello for drawing my attention to this point.

Proof. See appendix. □

Proposition 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 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 4.25 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.[41] 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 (4.24), similar 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.[42] 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 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

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

[42]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.

appendix:

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

ii) 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.

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 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.[43] 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 et al., 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 7.4 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

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

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 comovement 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 10 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 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 et al., 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.

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 (Chamon and Prasad, 2010; Chamon et al., 2013). Nonetheless, in Appendix 7.5, 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 Song et al. (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.

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 Poission process leads to Pareto-distributed income due to catch-up growth.[44] 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%![45]

For the coefficient of relative risk aversion (that simultaneously pins down the elasticity of intertemporal substitution) I pick $\eta = 2$. The results are sensitive to this number. Proposition 4.2 precisely shows how the parameters of the model, and in particular η, pin down the direction of capital flows.[46] I model the type distribution as a draw from a Log-Normal distribution, i.e.

$$log(\varphi) \sim N(-\frac{\sigma^2}{2}, \sigma^2). \tag{5.1}$$

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.[47] 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 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 (Krueger et al., 2016).[48]

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

[45] 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 roughly 206% instead of 250%.

[46] 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).

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

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

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.[49] 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.[50] 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.

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

$$L_t^r = L_0^r \exp\left(-\frac{g^*}{1-\alpha}t\right). \tag{5.2}$$

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

[49]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.

[50]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.

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 5.2. This law of motion of agricultural employment leads to the following estimating equation,

$$\log\left(L^r_{t,c}\right) = \beta_{0,c} + \beta_1 * t + \epsilon_{t,c}, \tag{5.3}$$

where I added a random error term. The estimating equation (5.3) 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 8.6.1 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 (Lucas, 2018), implies an estimate of $\hat{\alpha} \approx \frac{5}{9}$.[51] Figure 8 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 Tombe and Zhu (2019).[52]

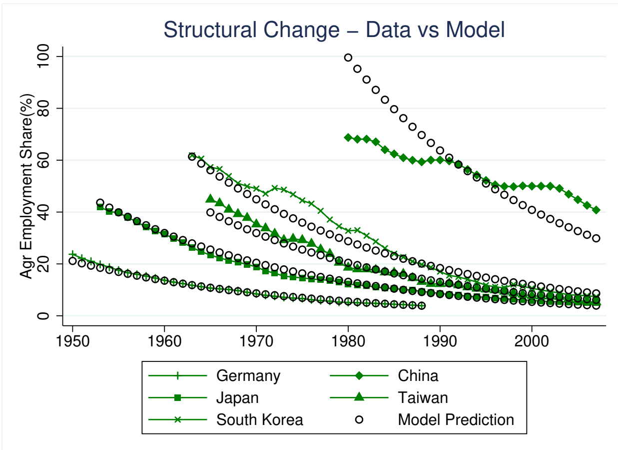

Figure 8: Prediction based on random effects model.

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

[51]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.

[52]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 7.1.1.

randomness washes out in the aggregate,[53]

$$g_{agg}\left(t\right)=\underbrace{\left(g-g^{*}\right)\frac{Y_{0,t}}{Y_{t}}+\left(\frac{g^{*}}{1-\alpha}\right)\frac{\hat{W}_{gap}L_{t}^{r}}{Y_{t}}}_{\text{catch-up growth}}+\underbrace{g^{*}}_{\text{long run growth}} \tag{5.4}$$

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 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 9 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 9 also shows the convergence in terms of log output per capita relative to the United States.[54]

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 7.2. 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 10 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

[53] Derivation:

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

[54] 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.

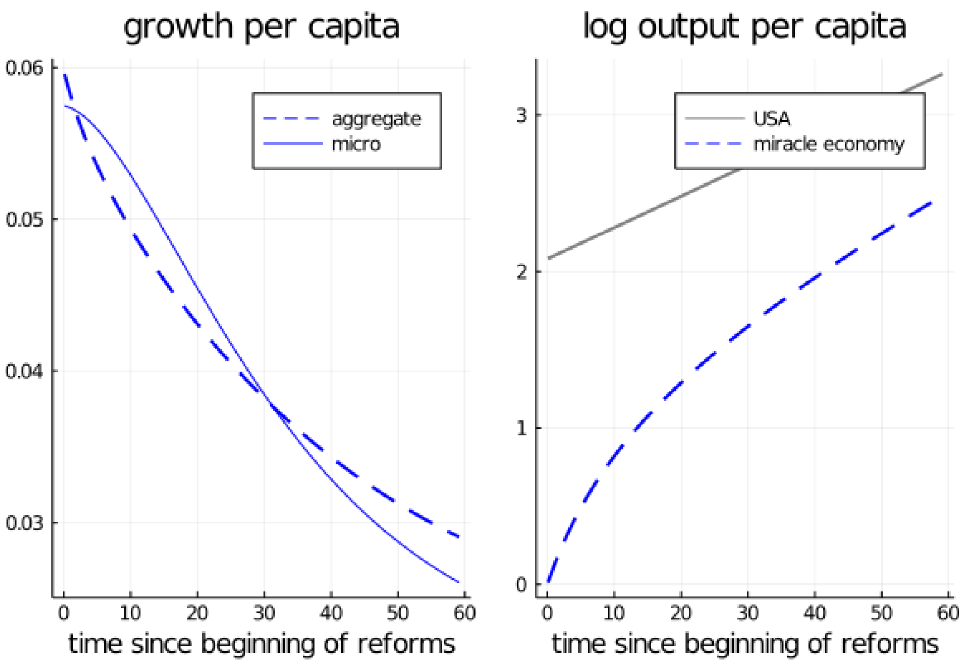

Figure 9: Per capita growth rates and log output computed using the values provided in table 5. Per capita GDP is assumed to be 8 times larger in the US at time zero.

for a relatively long time. The normalized stationary distribution of the log of income is

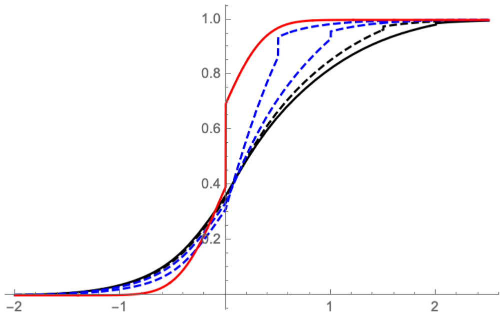

Figure 10: Simulated CDF of log of income, normalized by g^*t. The figure plots the distribution at time 0, 10, 20, 30, and 40 years in after the reforms started. Lines to the right are later periods compared to the left. The CDF displays jumps that stem from the initial share of households that start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

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).[55]

Armed with this CDF I can compute the log variance of income for non-agricultural

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

households displayed in table 2. The results are broadly inline with the fast rise of inequality in China.

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

Table 2: 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 3 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.[56]

5.4 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,

[56] 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.

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 4.24 and the budget constraint, I employ a simple shooting algorithm to solve the household problem.[57] Figure 14 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 11 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 11 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 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 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.

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 $y\left(t - t_m^i | T^i > t\right)$ and gives as output the asset-to-income ratio plotted in figure 11. Note that there is a one-to-one mapping between time spent in the high growth regime and income growth, and we

[57] 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.

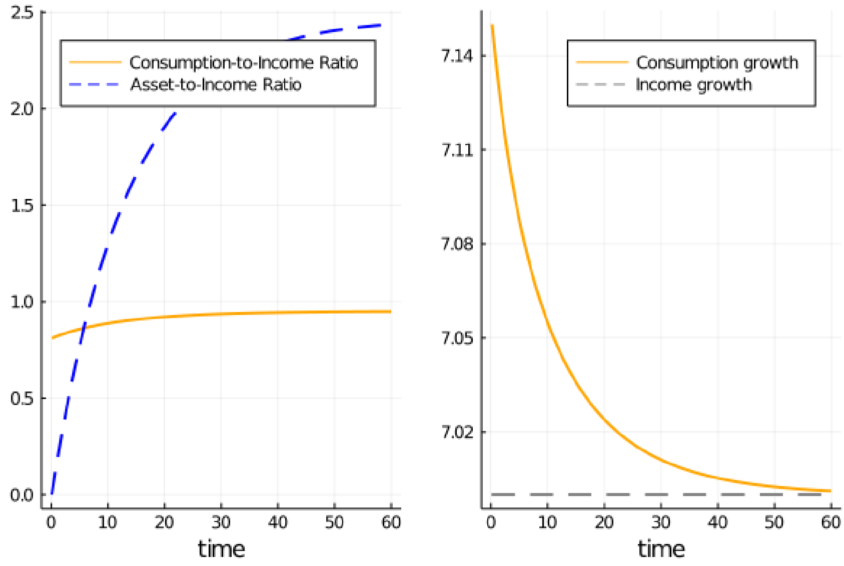

Figure 11: Transitional dynamics in high-growth regime using parameters from table 5. Time here is measured as time passed since the household switched sectors.

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_t^b 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 in subsection 7.3.

$$\frac{A_t^b}{A_0^u \exp(g^* t)} = M_{t,0} \int_1^{\exp((g-g^*)t)} y_0 Z(y_0)\, dF_0(y_0) + M_{t,1} \int_1^{\exp((g-g^*)t)} y_0 Z(y_0)\, dF_1(y_0) \tag{5.5}$$

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

$$f_1(k) = \frac{1}{M_{1,t}} \frac{\lambda}{(g-g^*)} k^{-\frac{\lambda}{(g-g^*)}-1} \left\{ 1 - M_{1,0} - L_t^r \left[k^{\frac{g^*}{(g-g^*)(1-\alpha)}} \right] \right\} \tag{5.6}$$

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

$$f_0(k) = \frac{1}{M_{0,t}} \frac{g^*}{(g-g^*)(1-\alpha)} L_t k^{-\left(\frac{\lambda}{g-g^*} - \frac{g^*}{(1-\alpha)(g-g^*)}\right)-1} \tag{5.7}$$

where the probability mass at $y = \exp((g-g^*)t)$ 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 12. 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 13 shows the declining share of agriculture. The hump-shaped saving rate can only emerge as a compositional effect. That is, there needs to be

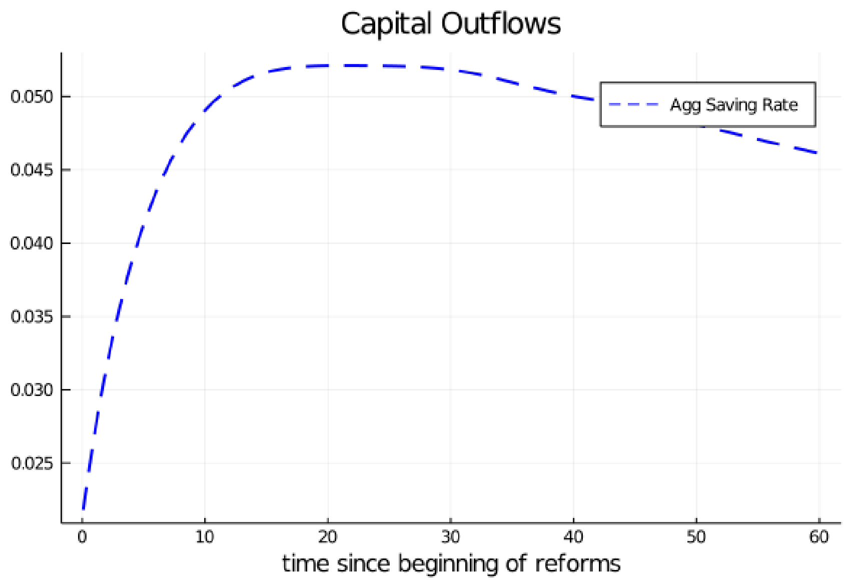

Figure 12: Aggregate saving rate over time. Since there is no capital in the model this coincides with the Current Account in the small open economy.

an increasing share of precautionary savers relative to total output for some time. This is precisely what happens as figure 13 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 12 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.

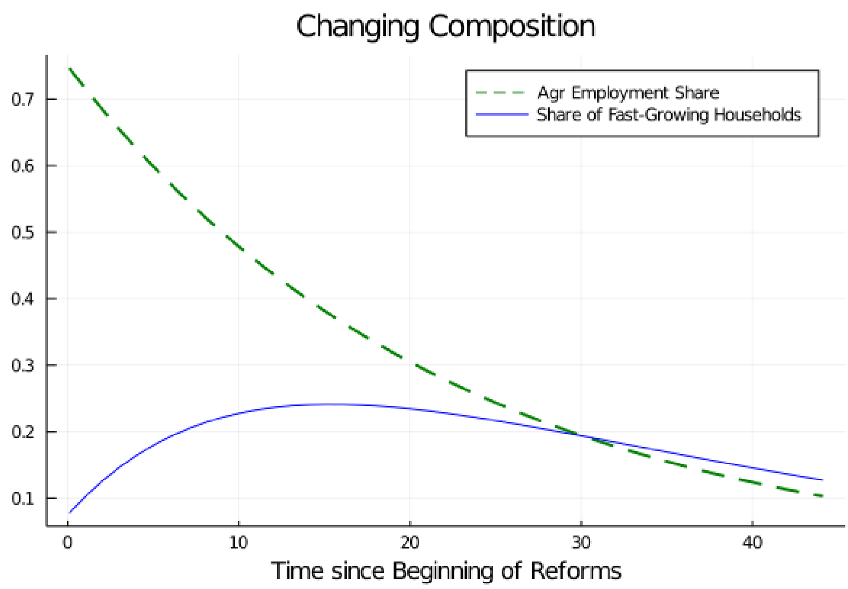

Figure 13: Changing share of population in agriculture, and in the high-growth regime.

5.5 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 11. 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.[58] 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. Table 17 in 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.[59] 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 predictions are born out by the data. In the appendix in table 18 the reader can verify that the median financial-assets-to-income ratios are systematically higher for households that always held an urban Hukou.[60]

In table 3 I report mean differences between households that always held an urban Hukou and switchers in terms of the growth rate of financial assets. I run a simple OLS regression

[58] 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 8.4 in the appendix.

[59] 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.

[60] 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

Table 3: Linear Regression for CFPS 2012 – 2016

	(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 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.

based on equation 3.1 but now with the growth of financial assets as the dependent variable.[61] 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,[62] 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 bufferstock savings.

6 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

around 1500 households which makes detecting differences harder compared to before.

[61]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.

[62]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.

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.

References

Aiyagari, S. R. (1994). Uninsured idiosyncratic risk and aggregate saving. *The Quarterly Journal of Economics 109*(3), 659–684.

Alcalá, F. and A. Ciccone (2004). Trade and productivity. *The Quarterly Journal of Economics 119*(2), 613–646.

Alfaro, L., S. Kalemli-Ozcan, and V. 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*, pp. 19–72. University of Chicago Press.

Alfaro, L., S. Kalemli-Ozcan, and V. Volosovych (2008). Why doesn't capital flow from rich to poor countries? an empirical investigation. *The review of economics and statistics 90*(2), 347–368.

Alvarez-Cuadrado, F. and M. Poschke (2011). Structural change out of agriculture: Labor push versus labor pull. *American Economic Journal: Macroeconomics 3*(3), 127–58.

Antras, P. (2001). Transitional dynamics of the savings rate in the neoclassical growth model. *V manuscript*.

Aoki, S. and M. Nirei (2017). Zipf's law, pareto's law, and the evolution of top incomes in the united states. *American Economic Journal: Macroeconomics 9*(3), 36–71.

Arkolakis, C. (2010). Market penetration costs and the new consumers margin in international trade. *Journal of political economy 118*(6), 1151–1199.

Attanasio, O. and M. 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.

Auclert, A., H. Malmberg, F. Martenet, and M. Rognlie (2020). Demographics, wealth, and global imbalances in the twenty-first century.

Barro, R. J. (2012). Convergence and modernization revisited. Technical report, National Bureau of Economic Research.

Bernanke, B. S. et al. (2005). The global saving glut and the us current account deficit. Technical report.

Bewley, T. (1977). The permanent income hypothesis: A theoretical formulation. *Journal of Economic Theory 16*(2), 252–292.

Blanchard, O. J. (1985). Debt, deficits, and finite horizons. *The Journal of Political Economy 93*(2), 223–247.

Blundell, R., L. Pistaferri, and I. Preston (2008). Consumption inequality and partial insurance. *American Economic Review 98*(5), 1887–1921.

Bryan, G., S. Chowdhury, and A. M. Mobarak (2014). Underinvestment in a profitable technology: The case of seasonal migration in bangladesh. *Econometrica 82*(5), 1671–1748.

Buera, F. J. and Y. Shin (2013). Financial frictions and the persistence of history: A quantitative exploration. *Journal of Political Economy 121*(2), 221–272.

Buera, F. J. and Y. Shin (2017). Productivity growth and capital flows: The dynamics of reforms. *American Economic Journal: Macroeconomics 9*(3), 147–85.

Caballero, R. J., E. Farhi, and P.-O. Gourinchas (2008). An equilibrium model of "global imbalances" and low interest rates. *American Economic Review 98*(1), 358–93.

Card, D., J. Heining, and P. Kline (2013). Workplace heterogeneity and the rise of west german wage inequality. *The Quarterly Journal of Economics 128*(3), 967–1015.

Carroll, C. D. (1994). How does future income affect current consumption? *The Quarterly Journal of Economics 109*(1), 111–147.

Carroll, C. D. (1997). Buffer-stock saving and the life cycle/permanent income hypothesis. *The Quarterly Journal of Economics 112*(1), 1–55.

Carroll, C. D., K. E. Dynan, and S. D. Krane (2003). Unemployment risk and precautionary wealth: Evidence from households' balance sheets. *Review of Economics and Statistics 85*(3), 586–604.

Carroll, C. D. and M. S. Kimball (1996). On the concavity of the consumption function. *Econometrica*, 981–992.

Carroll, C. D., J. Overland, and D. N. Weil (2000). Saving and growth with habit formation. *American Economic Review 90*(3), 341–355.

Carroll, C. D. and A. A. Samwick (1997). The nature of precautionary wealth. *Journal of Monetary Economics 1*(40), 41–71.

Caselli, F. (2005). Accounting for cross-country income differences. *Handbook of Economic Growth 1*, 679–741.

Chamon, M., K. Liu, and E. Prasad (2013). Income uncertainty and household savings in china. *Journal of Development Economics 105*, 164–177.

Chamon, M. D. and E. S. Prasad (2010). Why are saving rates of urban households in china rising? *American Economic Journal: Macroeconomics 2*(1), 93–130.

Chan, K. W. and W. Buckingham (2008). Is china abolishing the hukou system? *The China Quarterly 195*, 582–606.

Coeurdacier, N., H. Rey, and P. Winant (2019). Financial integration and growth in a risky world. *Journal of Monetary Economics*.

Coppola, A., M. Maggiori, B. Neiman, and J. Schreger (2020). Redrawing the map of global capital flows: The role of cross-border financing and tax havens. Technical report, National Bureau of Economic Research.

Curtis, C. C., S. Lugauer, and N. C. Mark (2015). Demographic patterns and household saving in china. *American Economic Journal: Macroeconomics 7*(2), 58–94.

De Magalhães, L. and R. 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. *Journal of Development Economics 134*, 350–371.

De Magalhães, L. and R. Santaeulàlia-Llopis (2018b). The consumption, income, and wealth of the poorest: An empirical analysis of economic inequality in rural and urban sub-saharan africa for macroeconomists. *Journal of Development Economics 134*, 350–371.

Deaton, A. (1991). Saving and liquidity constraints. *Econometrica 59*, 1221–1248.

Deaton, A. (2003). Adjusted indian poverty estimates for 1999-2000. *Economic and political Weekly*, 322–326.

Diamond, P. A. and J. E. Stiglitz (1974). Increases in risk and in risk aversion. *Journal of Economic Theory 8*(3), 337–360.

Ding, H. and H. He (2018). A tale of transition: An empirical analysis of economic inequality in urban china, 1986–2009. *Review of Economic Dynamics 29*, 106–137.

Epstein, L. G. and S. E. Zin (1991). Substitution, risk aversion, and the temporal behavior of consumption and asset returns: An empirical analysis. *Journal of political Economy 99*(2), 263–286.

Fagereng, A., M. B. Holm, B. Moll, and G. Natvik (2019). Saving behavior across the wealth distribution: The importance of capital gains. Technical report, National Bureau of Economic Research.

Fan, J. and Ş. Kalemli-Özcan (2016). Emergence of asia: reforms, corporate savings, and global imbalances. *IMF Economic Review 64*(2), 239–267.

Fan, J. and B. Zou (2019). Industrialization from scratch: The" construction of third front" and local economic development in china's hinterland.

Fuchs-Schündeln, N. and M. Schündeln (2005). Precautionary savings and self-selection: Evidence from the german reunification "experiment". *The Quarterly Journal of Economics 120*(3), 1085–1120.

Gabaix, X., J.-M. Lasry, P.-L. Lions, and B. Moll (2016). The dynamics of inequality. *Econometrica 84*(6), 2071–2111.

Gan, L., Q. He, R. Si, and D. Yi (2019). Relocating or redefined: A new perspective on urbanization in china. Technical report, National Bureau of Economic Research.

Garriga, C., A. Hedlund, Y. Tang, and P. Wang (2020). Rural-urban migration and house prices in china. *Regional Science and Urban Economics*, 103613.

Gollin, D., D. Lagakos, and M. E. Waugh (2013). The agricultural productivity gap. *The Quarterly Journal of Economics 129*(2), 939–993.

Gollin, D., D. Lagakos, and M. E. Waugh (2014). The agricultural productivity gap. *The Quarterly Journal of Economics 129*(2), 939–993.

Gourinchas, P.-O. and O. Jeanne (2013). Capital flows to developing countries: The allocation puzzle. *Review of Economic Studies 80*(4), 1484–1515.

Gourinchas, P.-O. and J. A. Parker (2002). Consumption over the life cycle. *Econometrica 70*(1), 47–89.

Guvenen, F. (2007). Learning your earning: Are labor income shocks really very persistent? *American Economic Review 97*(3), 687–712.

Guvenen, F., F. Karahan, S. Ozkan, and J. Song (2015). What do data on millions of us workers reveal about life-cycle earnings risk? Technical report, National Bureau of Economic Research.

Guvenen, F., S. Schulhofer-Wohl, J. Song, and M. Yogo (2017). Worker betas: five facts about systematic earnings risk. *American Economic Review 107*(5), 398–403.

Hall, R. E. (1988). Intertemporal substitution in consumption. *Journal of political economy 96*(2), 339–357.

Harris, J. R. and M. P. Todaro (1970). Migration, unemployment and development: a two-sector analysis. *The American economic review 60*(1), 126–142.

Hausmann, R., L. Pritchett, and D. Rodrik (2005). Growth accelerations. *Journal of Economic Growth 10*(4), 303–329.

He, H., F. Huang, Z. Liu, and D. Zhu (2018). Breaking the "iron rice bowl:" evidence of precautionary savings from the chinese state-owned enterprises reform. *Journal of Monetary Economics 94*, 94–113.

Heathcote, J., K. Storesletten, and G. L. Violante (2014). Consumption and labor supply with partial insurance: An analytical framework. *American Economic Review 104*(7), 2075–2126.

Herrendorf, B., R. Rogerson, and Á. Valentinyi (2014). Growth and structural transformation. In *Handbook of Economic Growth*, Volume 2, pp. 855–941. Elsevier.

Hicks, J. H., M. Kleemans, N. Y. Li, and E. Miguel (2017). Reevaluating agricultural productivity gaps with longitudinal microdata. Technical report, National Bureau of Economic Research.

Hill, R. J. (2000). Measuring substitution bias in international comparisons based on additive purchasing power parity methods. *European Economic Review 44*(1), 145–162.

Hnatkovska, V. and A. Lahiri (2018). Urbanization, structural transformation and rural-urban disparities in china and india.

Hsieh, C.-T. (1999). Productivity growth and factor prices in east asia. *American Economic Review 89*(2), 133–138.

Hsieh, C.-T. and P. J. Klenow (2007). Relative prices and relative prosperity. *American Economic Review 97*(3), 562–585.

Huggett, M. (1993). The risk-free rate in heterogeneous-agent incomplete-insurance economies. *Journal of economic Dynamics and Control 17*(5-6), 953–969.

İmrohoroğlu, A. and K. Zhao (2017). Household saving, financial constraints, and the current account in china. *International Economic Review*.

İmrohoroğlu, A. and K. Zhao (2018). The chinese saving rate: Long-term care risks, family insurance, and demographics. *Journal of Monetary Economics 96*, 33–52.

Jones, C. I. and J. Kim (2018). A schumpeterian model of top income inequality. *Journal of Political Economy 126*(5), 1785–1826.

Jordà, Ò., M. Schularick, and A. M. Taylor (2017). Macrofinancial history and the new business cycle facts. *NBER macroeconomics annual 31*(1), 213–263.

Kaboski, J. P. and R. M. Townsend (2011). A structural evaluation of a large-scale quasi-experimental microfinance initiative. *Econometrica 79*(5), 1357–1406.

Kogan, L., D. Papanikolaou, and N. Stoffman (2020). Left behind: Creative destruction, inequality, and the stock market. *Journal of Political Economy 128*(3), 855–906.

Krueger, D., K. Mitman, and F. Perri (2016). Macroeconomics and household heterogeneity. *2*, 843–921.

Lagakos, D., S. Marshall, A. M. Mobarak, C. Vernot, and M. E. Waugh (2020). Migration costs and observational returns to migration in the developing world. *Journal of Monetary Economics 113*, 138–154.

Lagakos, D., A. M. Mobarak, and M. E. Waugh (2018). The welfare effects of encouraging rural-urban migration. Technical report, National Bureau of Economic Research.

Lagakos, D. and M. E. Waugh (2013). Selection, agriculture, and cross-country productivity differences. *American Economic Review 103*(2), 948–80.

Lane, P. R. and G. M. Milesi-Ferretti (2007). The external wealth of nations mark ii: Revised and extended estimates of foreign assets and liabilities, 1970–2004. *Journal of international Economics 73*(2), 223–250.

Levine, R. and D. Renelt (1992). A sensitivity analysis of cross-country growth regressions. *The American Economic Review*, 942–963.

Lucas, Jr, R. E. (2004). Life earnings and rural-urban migration. *Journal of political economy 112*(S1), S29–S59.

Lucas, R. E. J. (1990). Why doesn't capital flow from rich to poor countries? *American Economic Review 80*(2), 92–96.

Lucas, R. E. J. (1993). Making a miracle. *Econometrica: Journal of the Econometric Society*, 251–272.

Lucas, R. E. J. (2018). What was the industrial revolution? *Journal of Human Capital 12*(2), 182–203.

Luttmer, E. G. (2007). Selection, growth, and the size distribution of firms. *The Quarterly Journal of Economics 122*(3), 1103–1144.

Luttmer, E. G. (2011). On the mechanics of firm growth. *The Review of Economic Studies 78*(3), 1042–1068.

Machado, J. A., P. Parente, and J. Santos Silva (2020). Qreg2: Stata module to perform quantile regression with robust and clustered standard errors.

Mas-Colell, A., M. D. Whinston, J. R. Green, et al. (1995). *Microeconomic theory*, Volume 1. Oxford university press New York.

Mendoza, E., V. Quadrini, and V. Rios-Rull (2009). Financial integration, financial deepness and global imbalances. *Journal of Political Economy 117*(3), 371.

Modigliani, F. (1986). Life cycle, individual thrift, and the wealth of nations. *Science 234*(4777), 704–712.

Moll, B. (2014). Productivity losses from financial frictions: Can self-financing undo capital misallocation? *American Economic Review 104*(10), 3186–3221.

Morten, M. (2019). Temporary migration and endogenous risk sharing in village india. *Journal of Political Economy 127*(1), 1–46.

Piketty, T., L. Yang, and G. Zucman (2019). Capital accumulation, private property, and rising inequality in china, 1978–2015. *American Economic Review 109*(7), 2469–96.

Porzio, T., F. Rossi, and G. Santangelo (2020). The human side of structural transformation.

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

Qin, B. and Y. Zhang (2014). Note on urbanization in china: Urban definitions and census data. *China Economic Review 30*, 495–502.

Restuccia, D., D. T. Yang, and X. Zhu (2008). Agriculture and aggregate productivity: A quantitative cross-country analysis. *Journal of Monetary Economics 55*(2), 234–250.

Rosenzweig, M. R. and O. Stark (1989). Consumption smoothing, migration, and marriage: Evidence from rural india. *Journal of political Economy 97*(4), 905–926.

Sala-I-Martin, X. X. (1997). I just ran two million regressions. *The American Economic Review*, 178–183.

Sandri, D. (2014). Growth and capital flows with risky entrepreneurship. *American Economic Journal: Macroeconomics 6*(3), 102–23.

Santaeulalia-Llopis, R. and Y. Zheng (2018). The price of growth: consumption insurance in china 1989–2009. *American Economic Journal: Macroeconomics 10*(4), 1–35.

Schechtman, J. (1976). An income fluctuation problem. *Journal of Economic Theory 12*(2), 218–241.

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

Solow, R. M. (1956). A contribution to the theory of economic growth. *The quarterly journal of economics 70*(1), 65–94.

Song, Y. (2014). What should economists know about the current chinese hukou system? *China Economic Review 29*, 200–212.

Song, Z., K. Storesletten, and F. Zilibotti (2011). Growing like china. *American Economic Review 101*(1), 196–233.

Storesletten, K., C. I. Telmer, and A. Yaron (2004). Consumption and risk sharing over the life cycle. *Journal of Monetary Economics 51*(3), 609–633.

Tombe, T. and X. Zhu (2019). Trade, migration, and productivity: A quantitative analysis of china. *American Economic Review 109*(5), 1843–72.

Townsend, R. M. (1994). Risk and insurance in village india. *Econometrica 62*(3), 539–591.

Townsend, R. M. (2016). Village and larger economies: The theory and measurement of the townsend thai project. *Journal of Economic Perspectives 30*(4), 199–220.

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

Wei, S.-J. and X. Zhang (2011). The competitive saving motive: Evidence from rising sex ratios and savings rates in china. *Journal of Political Economy 119*(3), 511–564.

Xie, Y. and J. Hu (2014). An introduction to the china family panel studies (cfps). *Chinese sociological review 47*(1), 3–29.

Xu, H. and Y. Xie (2015). The causal effects of rural-to-urban migration on children's well-being in china. *European sociological review 31*(4), 502–519.

Yaari, M. E. (1965). Uncertain lifetime, life insurance, and the theory of the consumer. *The Review of Economic Studies 32*(2), 137–150.

Young, A. (1995). The tyranny of numbers: confronting the statistical realities of the east asian growth experience. *The Quarterly Journal of Economics 110*(3), 641–680.

Young, A. (2013). Inequality, the urban-rural gap, and migration. *The Quarterly Journal of Economics 128*(4), 1727–1785.

7 Theory Appendix

7.1 Simple Infinite Horizon Economy with Poisson Arrival of Type

Deriving equation (4.20):

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

$$\begin{aligned}
V_{t_m} &= \max \int_{t_m}^{\infty} \lambda exp\left(-\lambda[t-t_m]\right) \left[\int_{t_m}^{t} exp\left(-\rho[s-t_m]\right) u\left(c_s\right) ds + exp\left(-\rho[t-t_m]\right) \mathbb{E}_{\varphi} V\left(\varphi, a_{t}\right) \right] dt \\
&= \max \int_{t_m}^{\infty} exp\left(-\rho[s-t_m]\right) u\left(c_s\right) \left[\int_{s}^{\infty} \lambda exp\left(-\lambda[t-t_m]\right) dt \right] ds \\
&+ \int_{t_m}^{\infty} \lambda exp\left(-\lambda[t-t_m]\right) exp\left(-\rho[t-t_m]\right) \mathbb{E}_{\varphi} V\left(\varphi, a_t\right) dt \\
&= \max \int_{t_m}^{\infty} exp\left(-\left(\lambda+\rho\right)\left[s-t_m\right]\right) \left[u\left(c_s\right) + \lambda \mathbb{E}_{\varphi}\left[V\left(\varphi, a_s\right)\right]\right] ds.
\end{aligned}$$

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.[63] 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

$$\frac{\dot{X}_s}{X_s} = -(g - g^*) + \frac{\lambda}{\eta}\left\{\mathbb{E}_\varphi\left[\left(\frac{X_s}{[g^*(\eta-1)+\rho]\,Z_s+\varphi}\right)^\eta\right]-1\right\} \tag{7.1}$$

$$\frac{\dot{Z}_s}{Z_s} = -(g - r^*) + \frac{1}{Z_s}(1 - X_s). \tag{7.2}$$

The loci for equation (7.1) and (7.2) are given by

$$X_{**} = \left\{\frac{(g-g^*)}{\lambda}\eta + 1\right\}^{\frac{1}{\eta}}\left\{\mathbb{E}_\varphi\left[\left(\frac{1}{\varphi + [g^*(\eta-1)+\rho]\,Z_{**}}\right)^\eta\right]\right\}^{-\frac{1}{\eta}} \tag{7.3}$$

$$Z_{**} = \frac{1 - X_{**}}{g - r^*} \tag{7.4}$$

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 (7.3) is increasing in Z_{**} and (7.4) 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 (7.3) to be below the intercept of (7.4) 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 $\mathbb{E}_\varphi\left[\left(\frac{1}{\varphi}\right)^\eta\right] = \exp\left(\frac{\sigma^2}{2}\eta\,[1+\eta]\right)$. 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 of

[63] 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

a mean-preserving spread that can be applied in the context at hand. Note that Mas-Colell 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 $\{X : X < B\} = \{X : X \leq -B\} \cup \{X : -B <$

$X < B\} \supset \{X : B < X < -B\} = \{X : |X| < B\}$.

$$\begin{aligned}\mathbb{P}\left(\varphi_K < y\right) &> \mathbb{P}\left(\left|\hat{\mathbb{E}}_K \log(\epsilon_j) - \mathbb{E}[\log(\epsilon)]\right| < -\log(y)\left(\frac{1}{M} - \frac{1}{K}\right)\right) \\ &= 1 - \mathbb{P}\left(\left|\hat{\mathbb{E}}_K \log(\epsilon_j) - \mathbb{E}[\log(\epsilon)]\right| \geq -\log(y)\left(\frac{1}{M} - \frac{1}{K}\right)\right) \\ &\geq 1 - \frac{\mathbb{V}(\epsilon)}{K\left[-\log(y)\left(\frac{1}{M} - \frac{1}{K}\right)\right]^2}\end{aligned}$$

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

$$\frac{d\frac{\dot{X}_s}{X_s}}{dZ}|_{X_{**}} < 0 \tag{7.5}$$

$$\frac{d\frac{\dot{Z}_s}{Z_s}}{dX}|_{Z_{**}} < 0. \tag{7.6}$$

Figure 14 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.

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.

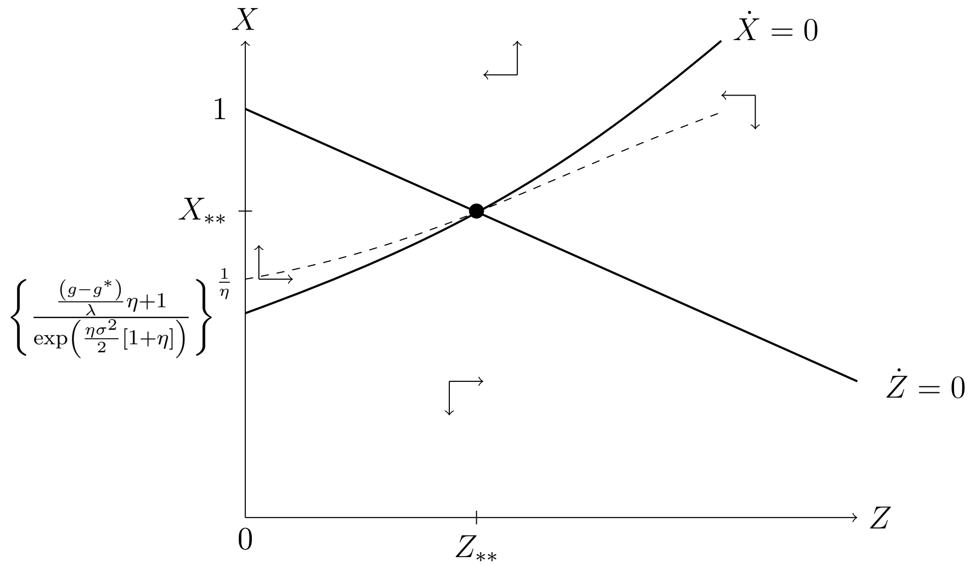

Figure 14: Pseudo steady state analysis of asset-to-income and consumption-to-income ratio in high-growth regime

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:

$$
\begin{aligned}
c\left(t_m, t\right) &= c_0\left(t-t_m\right) exp(g^* t_m) \\
a\left(t_m, t\right) &= a_0\left(t-t_m\right) exp\left(g^* t_m\right) \\
y\left(t_m, t\right) &= exp\left(g\left[t-t_m\right]\right) exp\left(g^* t_m\right) \\
&= y_0\left(t-t_m\right) exp\left(g^* t_m\right)
\end{aligned}
$$

Proof:

$$
X\left(s\right) = \frac{c\left(t, t_m\right)}{y\left(t, t_m\right)} = \frac{c\left(t+k, t_m+k\right)}{y\left(t+k, t_m+k\right)} \qquad \forall k \in [-t_m, \infty)
$$

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 $y(t_m) = A_t$, we can rewrite the equality

$$\frac{c(t, t_m)}{y(t - t_m, 0) A_{t_m}} = \frac{c(t - t_m, 0)}{y(t - t_m, 0)}$$

This implies that $c(t, t_m) = A_{t_m} c(t - t_m, 0)$ or short $c_0(t - t_m)\exp(g^* t_m)$. The case for assets is analogous. $\square$

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

$$V_{t_m} = \max_{c_s} \mathbb{E}_{\varphi,T} \int_{t_m}^{\infty} exp\left(-\rho[s - t_m]\right) \frac{c(\varphi, s, T)^{1-\eta}}{1-\eta} ds. \tag{7.7}$$

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 $\exp(g^* t_m)$. This allows us to rewrite the expression as follows

$$\begin{aligned} V_{t_m} &= \exp\left(-g^*[\eta - 1]t_m\right) \max_{c_s} \mathbb{E}_{\gamma,T} \int_{t_m}^{\infty} exp\left(-\rho[s - t_m]\right) \frac{c_0(\varphi, s - t_m, T - t_m)^{1-\eta}}{1-\eta} ds && (7.8) \\ &= \exp\left(-g^*[\eta - 1]t_m\right) V_0 && (7.9) \\ & && (7.10) \end{aligned}$$

This delivers the important result that

$$\dot{V_{t_m}} = -\left(g^*[\eta - 1]\right) V_{t_m} \tag{7.11}$$

Two comments are in order. First, 7.11 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.[64] 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 7.11 into the indifference condition (4.12) 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 that the overall wage rate in the urban sector keeps growing.

[64]Notes are available upon request.

$$
\begin{aligned}
\frac{(w_t^r)^{1-\eta}}{1-\eta} =& \tau^{\eta-1}\left(\rho V_t - \dot{V}_t\right) \\
\frac{(w_t^r)^{1-\eta}}{1-\eta} =& \tau^{\eta-1}\left(\rho V_t - g^*\left[1-\eta\right] V_t\right) \\
\frac{(w_t^r)^{1-\eta}}{1-\eta} =& \tau^{\eta-1} \exp\left(g^*\left[1-\eta\right] t\right) V_0 \left(\rho - g^*\left[1-\eta\right]\right) \\
w_t^r =& \exp\left(g^* t\right) \frac{\left[\left(1-\eta\right) V_0 \left(\rho + g^*\left[\eta-1\right]\right)\right]^{\frac{1}{1-\eta}}}{\tau}.
\end{aligned}
$$

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 (4.16) we get

$$
\frac{dL_t^u}{dt} = -\frac{dL_t^r}{dt}. \tag{7.12}
$$

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,

$$
\frac{dL_t^u}{dt} - \lambda M_{t,0} = \frac{dM_{t,0}}{dt}. \tag{7.13}
$$

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

$$dM_{t,1} = \lambda dt M_{t,0}. \tag{7.14}$$

using the Poisson arrival.[65]

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.[66] This would be agents in the set $M_{0,0}$. Using (7.13) together with (5.2) yields

$$\dot{M}_{t,0} + \lambda M_{t,0} = \frac{g^*}{1-\alpha} L_0^r \exp\left(-\frac{g^*}{1-\alpha}t\right). \tag{7.15}$$

Using $\exp(\lambda t)$ as integrating factor, this differential equation can be solved and yields

$$M_{t,0} = M_{0,0} exp(-\lambda t) + \frac{g^* L_0^r}{\lambda(1-\alpha) - g^*}\left[\exp(-\frac{g^*}{1-\alpha}t) - \exp(-\lambda t)\right]. \tag{7.16}$$

Similarly, the mass of households $M_{t,1}$ can be obtained

$$\begin{aligned} M_{t,1} &= \lambda \int_0^t M_{s,0} ds + M_{0,1} \\ &= M_{0,0}\left[1 - \exp(-\lambda t)\right] + \left(\frac{g^* L_0^r}{\lambda(1-\alpha) - g^*}\right)\left\{\frac{\lambda(1-\alpha)}{g^*}\left[1 - \exp(-\frac{g^*}{1-\alpha}t)\right] - \left[1 - \exp(-\lambda t)\right]\right\} + M_{0,1} \end{aligned}$$

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 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:

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 $Z_s < 0$.

Proof. Next, by contradiction, suppose that

$$\frac{\dot{c}}{c} = \frac{\lambda}{\eta}\left[\left(\frac{c_s}{y_s + [\rho + [\eta - 1]g^*]a_s}\right)^\eta - 1\right] + g^* < g^*$$

[65] There is a law of large numbers operating in the background here.

[66] I can relax that assumption very easily and will do so in a later section.

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 3:

Proof. 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 4:

Proof. Note that the proof for proposition 4 follows simply from the fact that proposition 2 and 3 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.

$$\begin{aligned}
V_{t,0} &= \int_t^\infty exp(-(\lambda+\rho)[s-t])\,\{log(y_s) + \lambda \mathbb{E}_\varphi V_{s,1}(\varphi, a_s)\}\,ds \\
&= \int_t^\infty exp(-(\lambda+\rho)[s-t]) \left\{[log(y_t) + g[s-t]] + \lambda \left[\frac{log(y_s)}{\rho} + \mathbb{E}_\varphi\left[\frac{log(\varphi)}{\rho}\right] + \frac{g^*}{\rho^2}\right]\right\} ds \\
&= \int_t^\infty exp(-(\lambda+\rho)[s-t]) \left\{[log(y_t) + g[s-t]] + \lambda \left[\frac{log(y_s)}{\rho} + \mathbb{E}_\varphi\left[\frac{log(\varphi)}{\rho}\right] + \frac{g^*}{\rho^2}\right]\right\} ds \\
&= \int_t^\infty exp(-(\lambda+\rho)[s-t]) \left\{[log(y_t) + g[s-t]] + \lambda \left[\frac{log(y_s)}{\rho} - \frac{\sigma^2}{2\rho} + \frac{g^*}{\rho^2}\right]\right\} ds \\
&= \int_t^\infty exp(-(\lambda+\rho)[s-t]) \left\{[log(y_t) + g[s-t]]\,(1+\frac{\lambda}{\rho}) + \lambda \left[\frac{g^*}{\rho^2} - \frac{\sigma^2}{2\rho}\right]\right\} ds \\
&= \left[\frac{log(y_t)}{\lambda+\rho} + \frac{g}{(\lambda+\rho)^2}\right](1+\frac{\lambda}{\rho}) + \frac{\lambda}{\lambda+\rho}\left[\frac{g^*}{\rho^2} - \frac{\sigma^2}{2\rho}\right] \\
&= \left[\frac{log(y_t)}{\lambda+\rho}\right](1+\frac{\lambda}{\rho}) + \frac{1}{\lambda+\rho}\frac{1}{\rho^2}[\lambda g^* + \rho g] - \frac{\lambda}{\lambda+\rho}\frac{\sigma^2}{2\rho} \\
&= \frac{log(y_t)}{\rho} + \frac{1}{\lambda+\rho}\frac{1}{\rho^2}[\lambda g^* + \rho g] - \frac{\lambda}{\lambda+\rho}\frac{\sigma^2}{2\rho}.
\end{aligned}$$

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

$$\begin{aligned} log(w_t^r) &= \rho\left(\frac{log(y_t)}{\rho} + \frac{1}{\lambda+\rho}\frac{1}{\rho^2}\left[\lambda g^* + \rho g\right] - \frac{\lambda}{\lambda+\rho}\frac{\sigma^2}{2\rho}\right) - \dot{V}_t \\ &= log(y_t) + \frac{1}{\lambda+\rho}\frac{1}{\rho}\left[\lambda g^* + \rho g\right] - \frac{\lambda}{\lambda+\rho}\frac{\sigma^2}{2} - \dot{V}_t \\ &= log(y_t) + \frac{1}{\lambda+\rho}\frac{1}{\rho}\left[\lambda g^* + \rho g\right] - \frac{\lambda}{\lambda+\rho}\frac{\sigma^2}{2} - \frac{g^*}{\rho} \\ &= log(A_t) + \frac{1}{\lambda+\rho}\left(g - g^*\right) - \frac{\lambda}{\lambda+\rho}\frac{\sigma^2}{2}. \end{aligned}$$

7.1.1 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 et al., 2008; Gollin et al., 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 (7.17)

$$\log(w_t^r) = \log(A_t) + \frac{1}{\lambda+\rho}\left(g - g^*\right) - \frac{\lambda}{\lambda+\rho}\frac{\sigma^2}{2} - \log\left(\tau\right). \tag{7.17}$$

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 $\frac{1}{\lambda+\rho}(g - g^*)$. 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.

7.2 Derivation of income distribution

From the main text we obtain the following equation

$$\log(y(t, t_{m_i}, T_i, \varphi)) = 1(T_i \geq t \geq t_{m_i})[g - g^*][t - t_{m_i}] + 1(T_i < t)\{[g - g^*][T_i - t_{m_i}] + \log\varphi\} \tag{7.18}$$

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:

$$\begin{aligned}
P\left(t_{m_i} \leq z|T_i > t\right) =& P\left(t_{m_i} < z|T_i > t\right) && (7.19)\\
=& \frac{P\left(t_{m_i} < z \cap T_i > t\right)}{P\left(T_i > t\right)} && (7.20)\\
=& \frac{\mathbb{E}_{t_{m_i}}\left[P\left(T_i > t|t_{m_i}\right)\right] D\left(t_{m_i} < z\right)}{P\left(T_i > t\right)} && (7.21)\\
=& \frac{\mathbb{E}_{t_{m_i}}\left[P\left(T_i - t_{m_i} > t - t_{m_i}|t_{m_i}\right)\right] D\left(t_{m_i} < z\right)}{P\left(T_i > t\right)} && (7.22)\\
=& \frac{\int_0^z exp(-\lambda(t-s))dF\left(s\right)}{M_{0,t}} && (7.23)\\
=& \frac{M_{0,0}\exp\left(-\lambda t\right) + \int_0^z exp(-\lambda(t-s))\frac{g^*}{1-\alpha}\Lambda exp(-\frac{g^* s}{1-\alpha})ds}{M_{0,t}} && (7.24)
\end{aligned}$$

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

$$
\begin{aligned}
P\left(t_{m_i} \leq k | i \in M_{1,t}\right) =& \frac{\lambda \int_0^k M_{0,x} dx + \lambda \int_k^t M_{0,x} \frac{M_{0,0} \exp(-\lambda x) + \int_0^k \frac{g^*}{1-\alpha} \Lambda \exp(-\frac{g^* s}{1-\alpha}) \exp(-\lambda(x-s)) ds}{M_{0,x}} dx}{\lambda \int_0^t M_{0,x} dx} \\
=& \frac{M_{1,k} + \lambda \int_k^t M_{0,0} \exp(-\lambda x) dx + \lambda \int_k^t \int_0^k \frac{g^*}{1-\alpha} \Lambda \exp(-\frac{g^* s}{1-\alpha}) \exp(-\lambda(x-s)) ds dx}{M_{1,t}} \\
=& \frac{M_{1,k} + M_{0,0} \exp(-\lambda t) \left(exp(\lambda\left[t-k\right]) - 1\right) + \lambda \frac{g^*}{1-\alpha} \Lambda \int_k^t \exp\left(-\lambda x\right) \int_0^k \exp\left(\left[\frac{(1-\alpha)\lambda - g^*}{1-\alpha}\right] s\right)}{M_{1,t}} \\
=& \frac{M_{1,k} + M_{0,0} \exp(-\lambda t) \left(exp(\lambda\left[t-k\right]) - 1\right)}{M_1, t} \\
&+ \frac{\lambda \frac{g^*}{1-\alpha} \Lambda \int_k^t \exp\left(-\lambda x\right) \frac{1-\alpha}{(1-\alpha)\lambda - g^*} \left[\exp\left(\left[\frac{(1-\alpha)\lambda - g^*}{1-\alpha}\right] k\right) - 1\right] dx}{M_{1,t}} \\
=& \frac{M_{1,k} + \left(\exp\left(-k\lambda\right) - \exp\left(-\lambda t\right)\right) \left\{M_{0,0} + \frac{\Lambda g^*}{g^* - (1-\alpha)\lambda} \left[1 - \exp\left(\left[-\frac{g^* - (1-\alpha)\lambda}{1-\alpha}\right] k\right)\right]\right\}}{M_{1,t}}
\end{aligned}
$$

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 arrivales 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 approximatin 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 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.

$$
\begin{aligned}
m_k &= \int (\log y)^k \, dF\left(y|i \in M_{1,t}\right) \\
&= [g-g^*]^k \, \mathbb{E}\left[(T_i - t_{m_i})^k \, |i \in M_{1,t}\right] \\
&= [g-g^*]^k \int \mathbb{E}\left[(T_i - t_{m_i})^k \, |t_{m_i} = z\right] dF_{t_{m_i}}\left(z|i \in M_{1,t}\right) \\
&= [g-g^*]^k \int_0^t \int (T_i - z)^k \, \frac{\lambda \exp\left(-\lambda\left[x\right]\right)}{1-\exp\left(-\lambda\left[t-t_{m_i}\right]\right)} dx \, dF_{t_{m_i}}\left(z|i \in M_{1,t}\right) \\
&= [g-g^*]^k \int_0^t \int_0^{t-z} (x)^k \, \frac{\lambda \exp\left(-\lambda\left[x\right]\right)}{1-\exp\left(-\lambda\left[t-z\right]\right)} dx \, dF_{t_{m_i}}\left(z|i \in M_{1,t}\right) \\
&= [g-g^*]^k \int_0^t \int_0^{t-z} (x)^k \, \frac{\lambda \exp\left(-\lambda\left[x\right]\right)}{1-\exp\left(-\lambda\left[t-z\right]\right)} dx \, dF_{t_{m_i}}\left(z|i \in M_{1,t}\right)
\end{aligned}
$$

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.

$$
\begin{aligned}
&\frac{M_{1,t}}{M_{0,t}+M_{1,t}}\mathbb{E}_{t_m}\mathbb{E}_{\log\varphi}P_1\left(T_i-t_m<\frac{k-\log\varphi}{g-g^*}|t_m,\varphi\right)\\
=&\frac{M_{1,t}}{M_{0,t}+M_{1,t}}\mathbb{E}_{t_m}\mathbb{E}_{\log\varphi}P_1\left(T_i-t_m\leq\frac{k-\log\varphi}{g-g^*}\right)D\left(0\leq k-\log\varphi\leq(g-g^*)(t-t_m)\right)\\
+&\frac{M_{1,t}}{M_{0,t}+M_{1,t}}\mathbb{E}_{t_m}\mathbb{E}_{\log\varphi}P_1\left(T_i-t_m\leq\frac{k-\log\varphi}{g-g^*}\right)D\left(\frac{k-\log\varphi}{g-g^*}<0\right)\\
+&\frac{M_{1,t}}{M_{0,t}+M_{1,t}}\mathbb{E}_{t_m}\mathbb{E}_{\log\varphi}P_1\left(T_i-t_m\leq\frac{k-\log\varphi}{g-g^*}\right)D\left(\frac{k-\log\varphi}{g-g^*}>t-t_m\right)\\
=&\frac{M_{1,t}}{M_{0,t}+M_{1,t}}E_{t_m}E_{\log\varphi}P_1\left(T_i-t_m\leq\frac{k-\log\varphi}{g-g^*}\right)D\left(0\leq k-\log\varphi\leq t-t_m\right)\\
+&\frac{M_{1,t}}{M_{0,t}+M_{1,t}}E_{t_m}E_{\log\varphi}*0*D\left(\frac{k-\log\varphi}{g-g^*}<0\right)\\
+&\frac{M_{1,t}}{M_{0,t}+M_{1,t}}E_{t_m}E_{\log\varphi}*1*D\left(\frac{k-\log\varphi}{g-g^*}>t-t_m\right)
\end{aligned}
$$

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.

$$
\begin{aligned}
=&\frac{M_{1,t}}{M_{0,t}+M_{1,t}}E_{t_m}\int_{[k-(t-t_m)(g-g^*)]}^{k} P_1\left(T_i - t_m \leq \frac{k-\log\varphi}{g-g^*}\right) dF\left(\log\varphi\right) \\
&+\frac{M_{1,t}}{M_{0,t}+M_{1,t}}E_{t_m}\Phi\left(\frac{k-\left(g-g^*\right)\left(t-t_m\right)}{\sigma}+\frac{\sigma}{2}\right) \\
=&\frac{M_{1,t}}{M_{0,t}+M_{1,t}}E_{t_m}\int_{[k-(t-t_m)(g-g^*)]}^{k} \frac{1-\exp\left(-\lambda\left[\frac{k-\log\varphi}{g-g^*}\right]\right)}{1-\exp\left(-\lambda\left[t-t_m\right]\right)} dF\left(\log\varphi\right) \\
&+\frac{M_{1,t}}{M_{0,t}+M_{1,t}}E_{t_m}\Phi\left(\frac{k-\left(g-g^*\right)\left(t-t_m\right)}{\sigma}+\frac{\sigma}{2}\right) \\
&=\frac{M_{1,t}}{M_{0,t}+M_{1,t}}E_{t_m}\frac{1}{1-\exp\left(-\lambda\left[t-t_m\right]\right)}\int_{\{k-(t-t_m)(g-g^*)\}}^{k} 1-\exp\left(-\lambda\left[\frac{k}{g-g^*}\right]\right)\exp\left(+\lambda\left[\frac{\log\varphi}{g-g^*}\right]\right) dF\left(\log\right. \\
&+\frac{M_{1,t}}{M_{0,t}+M_{1,t}}E_{t_m}\Phi\left(\frac{k-\left(g-g^*\right)\left(t-t_m\right)}{\sigma}+\frac{\sigma}{2}\right) \\
&=\frac{M_{1,t}}{M_{0,t}+M_{1,t}}E_{t_m}\frac{1}{1-\exp\left(-\lambda\left[t-t_m\right]\right)}\left[\Phi\left(\frac{k}{\sigma}+\frac{\sigma}{2}\right)-\Phi\left(\frac{k-\left(g-g^*\right)\left(t-t_m\right)}{\sigma}+\frac{\sigma}{2}\right)\right] \\
&-\frac{M_{1,t}}{M_{0,t}+M_{1,t}}E_{t_m}\frac{\exp\left(-\lambda\left[\frac{k}{g-g^*}\right]\right)}{1-\exp\left(-\lambda\left[t-t_m\right]\right)}\int_{\{k-(t-t_m)(g-g^*)\}}^{k} \exp\left(\frac{\lambda}{g-g^*}\log\varphi\right) dF\left(\log\varphi\right) \\
&+\frac{M_{1,t}}{M_{0,t}+M_{1,t}}E_{t_m}\Phi\left(\frac{k-\left(g-g^*\right)\left(t-t_m\right)}{\sigma}+\frac{\sigma}{2}\right)
\end{aligned}
$$

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 ($https://en.wikipedia.org/wiki/Truncated\_normal\_distribution$).

$$
\mathbb{E}[x^t|a \leq x \leq b] = \exp\left(\mu t + \frac{\sigma^2 t^2}{2}\right)\left[\frac{\Phi\left(\beta-\sigma t\right)-\Phi\left(\alpha-\sigma t\right)}{\Phi\left(\beta\right)-\Phi\left(\alpha\right)}\right]
$$

with $\alpha = \frac{a-\mu}{\sigma}$ and $\beta = \frac{b-\mu}{\sigma}$. Using this result yields

$$
\begin{aligned}
&=\frac{M_{1,t}}{M_{0,t}+M_{1,t}}E_{t_m}\frac{1}{1-\exp\left(-\lambda\left[t-t_m\right]\right)}\left[\Phi\left(\frac{k}{\sigma}+\frac{\sigma}{2}\right)-\Phi\left(\frac{k-\left(g-g^*\right)\left(t-t_m\right)}{\sigma}+\frac{\sigma}{2}\right)\right] \\
&-\frac{M_{1,t}}{M_{0,t}+M_{1,t}}E_{t_m}\exp\left(-\frac{\lambda}{g-g^*}\left(k-\frac{\sigma^2}{2}\left(\frac{\lambda}{g-g^*}-1\right)\right)\right)\left[\Phi\left(\frac{k}{\sigma}+\frac{\sigma}{2}-\sigma\frac{\lambda}{g-g^*}\right)-\Phi\left(\frac{k-(t-t_m)(g-g^*}{\sigma}\right.\right. \\
&+\frac{M_{1,t}}{M_{0,t}+M_{1,t}}E_{t_m}\Phi\left(\frac{k-\left(g-g^*\right)\left(t-t_m\right)}{\sigma}+\frac{\sigma}{2}\right)
\end{aligned}
$$

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

$$
\begin{aligned}
&= \frac{M_{1,t}}{M_{0,t}+M_{1,t}} \int_0^t \frac{1}{1-\exp\left(-\lambda\left[t-t_m\right]\right)} \left[\Phi\left(\frac{k}{\sigma}+\frac{\sigma}{2}\right) - \Phi\left(\frac{k}{\sigma}+\frac{\sigma}{2}-\frac{\left(g-g^*\right)\left(t-t_m\right)}{\sigma}\right)\right] dF_1(t_m) \\
&- \frac{M_{1,t}}{M_{0,t}+M_{1,t}} \int_0^t \exp\left(-\frac{\lambda}{g-g^*}\left(k-\frac{\sigma^2}{2}\left(\frac{\lambda-(g-g^*)}{g-g^*}\right)\right)\right) \\
&* \left[\Phi\left(\frac{k}{\sigma}+\frac{\sigma}{2}-\sigma\frac{\lambda}{g-g^*}\right) - \Phi\left(\frac{k}{\sigma}+\frac{\sigma}{2}-\sigma\frac{\lambda}{g-g^*}-\frac{(t-t_m)(g-g^*)}{\sigma}\right)\right] dF_1(t_m) \\
&+ \frac{M_{1,t}}{M_{0,t}+M_{1,t}} \int_0^t \Phi\left(\frac{k-\left(g-g^*\right)\left(t-t_m\right)}{\sigma}+\frac{\sigma}{2}\right) dF_1(t_m)
\end{aligned}
$$

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.

7.3 Derivation of f_0, f_1

The conditional density for household income in the high growth regime reads

$$
f_0\left(k\right) = \frac{1}{M_{0,t}} \frac{g^*}{\left(g-g^*\right)\left(1-\alpha\right)} L_t k^{-\left(\frac{\lambda}{g-g^*}-\frac{g^*}{(1-\alpha)(g-g^*)}\right)-1} \tag{7.25}
$$

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

$$
f_1\left(k\right) = \frac{1}{M_{1,t}} \frac{\lambda}{\left(g-g^*\right)} k^{-\frac{\lambda}{(g-g^*)}-1} \left\{1 - M_{1,0} - L_t^r \left[k^{\frac{g^*}{(g-g^*)(1-\alpha)}}\right]\right\} \tag{7.26}
$$

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}.[67] Then,

[67] Note that I keep everything normalized by the constant growth rate g^* to study the stationary distributions.

$$
\begin{aligned}
P\left(y_{0,t}^i \le k\right) &= P\left(\log y_{0,t}^i \le \log k\right) \\
&= P\left(\left(g-g^*\right)\left(t-t_{m_i}\right) \le \log k\right) \\
&= P\left(\left(g-g^*\right)\left(t-t_{m_i}\right) \le \log k\right) \\
&= P\left(t-t_{m_i} \le \frac{\log k}{g-g^*}\right) D\left(k \le \exp\left(\left(g-g^*\right)t\right)\right) \\
&= P\left(t-t_{m_i} \le \frac{\log k}{g-g^*}\right) D\left(0 \le \frac{\log k}{g-g^*} \le t\right)
\end{aligned}
$$

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 (7.24). 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

$$
\begin{aligned}
f_0(k) &= \frac{1}{k\left(g-g^*\right)} f_x\left(\frac{\log k}{g-g^*}\right) \\
&= \frac{g^*}{k\left(g-g^*\right)(1-\alpha)} \frac{L_0^r \exp\left(-\frac{g^*}{1-\alpha}t\right)\exp\left(-\left(\frac{\lambda}{g-g^*} - \frac{g^*}{(1-\alpha)(g-g^*)}\right)\log k\right)}{M_{0,t}} \\
&= \frac{1}{M_{0,t}} \frac{g^*}{\left(g-g^*\right)(1-\alpha)} L_t k^{-\left(\frac{\lambda}{g-g^*} - \frac{g^*}{(1-\alpha)(g-g^*)}\right)-1}.
\end{aligned}
$$

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

$$
\begin{aligned}
P\left(y_t \le k | T_i < t\right) = & \quad P\left(\log y_t \le \log k | T_i < t\right) \\
= & \quad P\left(\left(g-g^*\right)\left(T_i - t_{m_i}\right) \le \log k | T_i < t\right) \\
= & \quad \mathbb{E}_{t_m} P\left(\left(T_i - t_{m_i}\right) \le \frac{\log k}{\left(g-g^*\right)} | T_i < t, t_m\right)
\end{aligned}
$$

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.

$$
\begin{aligned}
P\left(y_t \leq k | T_i < t\right) &= \mathbb{E}_{t_m} P\left(T_i - t_{m_i} \leq \frac{\log k}{(g-g^*)} | T_i < t, t_m\right) \\
&= \mathbb{E}_{t_m}\left\{\frac{1-\exp\left(-\lambda\left[\frac{\log k}{(g-g^*)}\right]\right)}{1-\exp\left(-\lambda\left[t-t_{m_i}\right]\right)} D\left(\frac{\log k}{(g-g^*)} \leq t - t_{m_i}\right) + 1 * D\left(\frac{\log k}{(g-g^*)} > t - t_{m_i}\right)\right\} \\
&= \mathbb{E}_{t_m}\left\{\frac{1-\exp\left(-\lambda\left[\frac{\log k}{(g-g^*)}\right]\right)}{1-\exp\left(-\lambda\left[t-t_{m_i}\right]\right)} D\left(\frac{\log k}{(g-g^*)} \leq t - t_{m_i}\right)\right\} \\
&+ \int_0^t \left\{D\left(\frac{\log k}{(g-g^*)} > t - t_{m_i}\right)\right\} dF\left(t_m | i \in M_{1,t}\right).
\end{aligned}
$$

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 $f_1(t_m|T^i < t)$,

$$
\begin{aligned}
&= \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) \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. \\
&+ \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}} \\
&= \frac{1}{M_{1,t}} \int_0^{t-\frac{\log k}{(g-g^*)}} \left\{1-\exp\left(-\lambda\left[\frac{\log k}{(g-g^*)}\right]\right)\right\} \left(\frac{g^* L_0^r}{1-\alpha}\right) \exp\left(-\frac{g^*}{1-\alpha}t_m\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}} \\
&= \frac{L_0^r}{M_{1,t}} \left\{1-\exp\left(-\lambda\left[\frac{\log k}{(g-g^*)}\right]\right)\right\} \left\{1-\exp\left(-\frac{g^*}{1-\alpha}\left(t-\frac{\log k}{(g-g^*)}\right)\right)\right\} \\
&+ \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}} \\
&= \frac{L_0^r}{M_{1,t}} \left\{1-\exp\left(-\frac{\lambda}{(g-g^*)}\log k\right)\right\} \left\{1-\exp\left(-\frac{g^*}{1-\alpha}t\right)\exp\left(\frac{g^*}{(g-g^*)\left(1-\alpha\right)}\log k\right)\right\} \\
&+ \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}} \\
&= \frac{L_0^r}{M_{1,t}} \left\{\exp\left(-\frac{g^*}{1-\alpha}t\right)\exp\left(\frac{g^*-\lambda\left(1-\alpha\right)}{(g-g^*)\left(1-\alpha\right)}\log k\right) - \exp\left(-\frac{g^*}{1-\alpha}t\right)\exp\left(\frac{g^*}{(g-g^*)\left(1-\alpha\right)}\log k\right)\right\} \\
&+ \frac{L_0^r + 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}}
\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

$$
\begin{aligned}
f(y|y \in Y_1) =& \frac{1}{M_{1,t}} \left\{ (1-M_{1,0}) \left[\frac{\lambda}{(g-g^*)} k^{-\frac{\lambda}{(g-g^*)}-1} \right] \right\} \\
&- \frac{1}{M_{1,t}} \left\{ \frac{L_t^r}{(g-g^*)(1-\alpha)} \left[(\lambda(1-\alpha)-g^*) k^{-\frac{\lambda(1-\alpha)-g^*}{(g-g^*)(1-\alpha)}-1} + g^* k^{\frac{g^*}{(g-g^*)(1-\alpha)}-1} \right] \right\} \\
&+ \frac{1}{M_{1,t}} \frac{g^* L_0^r}{k(g-g^*)(1-\alpha)} \left[\exp\left(-\frac{g^*}{1-\alpha} \left(t - \frac{\log k}{g-g^*} \right) \right) - \exp\left(\frac{\lambda(1-\alpha)-g^*}{1-\alpha} \left(t - \frac{\log k}{g-g^*} \right) \right) \exp \right. \\
=& \frac{1}{M_{1,t}} (1-M_{1,0}) \left[\frac{\lambda}{(g-g^*)} k^{-\frac{\lambda}{(g-g^*)}-1} \right] \\
&- \frac{1}{M_{1,t}} \frac{L_t^r}{(g-g^*)(1-\alpha)} \left[(\lambda(1-\alpha)-g^*) k^{-\frac{\lambda(1-\alpha)-g^*}{(g-g^*)(1-\alpha)}-1} + g^* k^{\frac{g^*}{(g-g^*)(1-\alpha)}-1} \right] \\
&+ \frac{1}{M_{1,t}} \frac{g^* L_t^r}{(g-g^*)(1-\alpha)} \left[k^{\frac{g^*}{(g-g^*)(1-\alpha)}-1} - k^{-\frac{\lambda(1-\alpha)-g^*}{(g-g^*)(1-\alpha)}-1} \right] \\
=& \frac{1}{M_{1,t}} (1-M_{1,0}) \left[\frac{\lambda}{(g-g^*)} k^{-\frac{\lambda}{(g-g^*)}-1} \right] \\
&- \frac{1}{M_{1,t}} \frac{L_t^r}{(g-g^*)(1-\alpha)} \left[(\lambda(1-\alpha)-g^*+g^*) k^{-\frac{\lambda(1-\alpha)-g^*}{(g-g^*)(1-\alpha)}-1} + (g^*-g^*) k^{\frac{g^*}{(g-g^*)(1-\alpha)}-1} \right]
\end{aligned}
$$

hence we obtain

$$
f(y_0|y_0 \in Y_1) = \frac{1}{M_{1,t}} \frac{\lambda}{(g-g^*)} k^{-\frac{\lambda}{(g-g^*)}-1} \left\{ 1 - M_{1,0} - L_t^r \left[k^{\frac{g^*}{(g-g^*)(1-\alpha)}} \right] \right\}
$$

for the continuous part of the density.

7.4 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.[68] Consequently, the optimality condition simplifies to

$$\exp\left(\eta\left[g - g^*\right]T\right) = \mathbb{E}_{\varphi}\left[\left(\varphi^i + \exp\left(\left[r^* - g\right]T\right)\left[\frac{r^* - g^*}{g - r^*}\left[\exp\left(\left[g - r^*\right]T\right) - 1\right] - \left[1 - \exp\left(\left[g^* - r^*\right]T\right)\right]\right]\right)^{-\eta}\right]$$

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.[69] 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 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.

[68] 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.

[69] 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 (Krueger et al., 2016).

7.5 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,

$$\begin{aligned} k_i &= s y_i \\ &= s A^r X^{\alpha-1} \end{aligned}$$

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 bufferstock 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

$$y_t^i = k_t^\beta \left(A_t \varphi h_t\right)^{1-\beta} \tag{7.27}$$

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.

8 Data Appendix

8.1 Aggregate Time Series

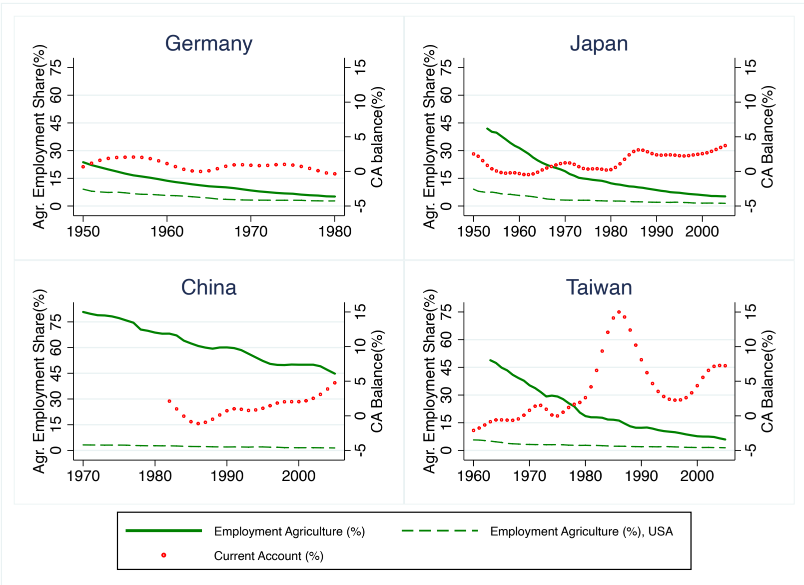

Figure 15: Relationship between agricultural employment share and current account for Germany, Japan, China, and Taiwan. Current Account series is smoothed using an hp-filter with smoothing parameter of 8.5. Current account data is from the WDI, Taiwanese Statistical office, and the historical macro database from Jordà et al. (2017) (for Germany and Japan).

8.2 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 et al., 2019). As mentioned in the introduction, my model does not necessarily 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 5 in urban areas, and omit the rural counterpart.

Table 4: 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 8.2 .

8.3 Aggregate Inequality

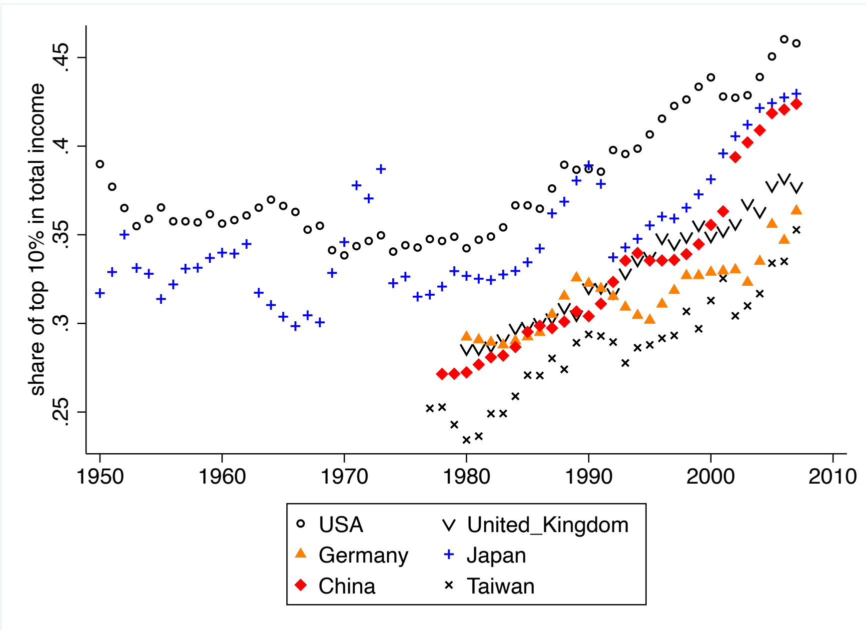

Figure 16: Share of income of top 10% earners to total income from the World Inequality Database. Raw data.

16 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 et al. (2013) for an analysis of the rise in West German wage inequality.

8.4 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 et al., 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 et al. (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$_a$sset 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 ap-*

pliances. 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 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, 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.[70]

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.

Qin and 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 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 Zhang (2014) both argue that much of the growth in urban areas is driven by internal

[70]Does someone know more about this than I do?

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 5 display mean and standard deviations of the raw sample, i.e. before I restrict on age etc.

Table 5: 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 6: 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.

8.5 Additional robustness for standard and quantile regressions

Table 7 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 7: 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 8 offers results for asset-to-consumption ratio, which might be a better proxy for permanent income.

Table 10 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 11 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 12, which is very similar to table 11. 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 agricultural and non-agricultural employment.

Table 8: Median regression using dummy for agricultural occupation for CHIP 1995

	(1) no_prod_assets_consum	(2) no_prod_assets_consum	(3) no_prod_assets_consum	(4) 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 9: 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 10: 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.

Table 11: Median regression with agricultural dummy for CFPS 2012

	(1) fin_asset_consum	(2) fin_asset_consum	(3) fin_asset_consum	(4) fin_asset_consum	(5) 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 12: Median regression using urban-rural dummy for CFPS 2012

	(1) fin_asset_consum	(2) fin_asset_consum	(3) fin_asset_consum	(4) fin_asset_consum	(5) 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 13: Median regression using urban-rural dummy for CFPS 2012

	(1) fin_asset_income	(2) fin_asset_income	(3) fin_asset_income	(4) fin_asset_income	(5) 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 14: Median regression with agricultural dummy for CFPS 2012

	(1) total_asset_income	(2) total_asset_income	(3) total_asset_income	(4) total_asset_income	(5) 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.

8.6 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.[71]

[71] The higher total-asset-to-income ratio is coming from the relatively lower income of switchers, and is driven by housing wealth.

Table 15: 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.

Table 16: Median regression using urban-rural dummy in CFPS 2012

	(1) total_asset_consum	(2) total_asset_consum	(3) total_asset_consum	(4) total_asset_consum	(5) total_asset_consum
urban_cfps	1.391***	0.589**	0.592***	-0.120	-0.101
	(0.214)	(0.246)	(0.228)	(0.260)	(0.243)
_cons	5.097***	4.470***	7.088***	5.433***	2.244
	(0.0936)	(0.167)	(1.408)	(1.434)	(2.475)
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-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 17: 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.

8.6.1 structural change regression results

Table 18: Hukou Median Regression for CFPS 2012

	(1) fin_asset_income	(2) fin_asset_income	(3) fin_asset_income	(4) fin_asset_income	(5) 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 19: Regression for CFPS 2012 – 2016

	(1) g_total_asset	(2) g_total_asset	(3) g_total_asset	(4) g_total_asset	(5) 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 20: Regression for CFPS 2012 – 2016

	(1) g_arc_total_asset	(2) g_arc_total_asset	(3) g_arc_total_asset	(4) g_arc_total_asset	(5) g_arc_total_asset
hukou_switcher	0.0823*	0.0700	0.0823*	0.0941**	0.0438
	(0.0439)	(0.0439)	(0.0446)	(0.0452)	(0.0463)
_cons	0.373***	0.322***	0.975**	0.793*	0.948**
	(0.0303)	(0.0323)	(0.413)	(0.419)	(0.432)
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.

8.7 Cross-Country Regression Evidence

The model has implications for aggregate savings, current accounts, and trade surpluses that I investigate in this section using cross-country data. While identifying causal relationships from aggregate time series is very challenging, cross-country regressions have been used extensively to learn about the "correlates" of growth (Levine and Renelt, 1992; Sala-I-Martin, 1997). In the same spirit, I hope to provide some additional evidence to the reader that structural change out

Table 21: Regression for CFPS 2012 – 2016

	(1) g_arc_fin_asset	(2) g_arc_fin_asset	(3) g_arc_fin_asset	(4) g_arc_fin_asset	(5) 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.

Table 22: Structural Change Regression

	(1) random effects	(2) fixed effects
Year	-0.0446	-0.0447
	(0.00339)	(0.000675)
r2		0.978
r2_w	0.958	

Data from the 10-Sector database from the University of Groeningen. Countries and time periods considered: China(1979–2007), Germany (1950–1989), Japan(1955–2007), Korea(1963–2007), Taiwan(1965–2007). Results are robu to different starting dates of the growth miracle. Hongokong and Singapur are excluded as their emergence as financial centers gave rise to different patterns of structural change. Robust Standard errors.

of agriculture is robustly correlated with rising saving rates.

Of course, cross-country regressions come with multiple caveats. Classic threats to identification are omitted variable bias and reverse causality. Savings and changes in the agricultural share are jointly determined, and a regression therefore does not consistently estimate the partial effect of a decrease in the agricultural share on the saving rate.[72]

8.7.1 Description of the data

I use data from the Penn World tables (PWT), version 9.0, and match it with data from the GGDC data project in Groningen that contains long time-series on agricultural shares to estimate the reduced form relationship between structural change and saving rates. This unfortunately means that the number of countries in the estimation goes down to 40. The gain is that these are very long time series relative to the World Development Indicators (WDI). Unlike most researchers who use the PWT I am not interested in comparing GDP across countries, and I can instead rely on the national accounts data and domestic (non-PPP-adjusted) price indices.[73]

[72] Another issue with fixed effect regressions is that it requires strict exogeneity of the error term, i.e. the error has to be uncorrelated with all future and past realizations of the regressors. Forward and backward looking behavior of optimizing agents usually implies the failure of this assumption. See more on this issue in Barro (2012) who is skeptical of using a fixed effects framework for cross-country growth regressions.

[73] Comparing GDP across countries is an very challenging measurement exercises. The PWT rely on a version of the Geary-Khamis price index to map nominal GDP into real GDP using purchasing-power-

Measurement of savings may seem straightforward as a residual quantity from the national accounts after subtracting private and government consumption. Importantly, I measure savings as the nominal share of GDP that is not spend on private or government consumption, instead of "real savings".[74] In addition, I drop the smallest and largest 1% of data points in terms of their gross saving rate to make sure my results are not driven by extreme outliers as well as countries with a population below 1 million.

8.7.2 Cross-country evidence

The first specification estimates the effect of a decrease in the agricultural share on the gross saving rate in an economy using time and country fixed effects.

$$s_{i,t} = \alpha_i + \tau_t + \beta_1 g_{i,t} + \beta_2 Y_{it} + \beta_3 agrshare_{i,t} + \beta_4 distfrontier_{i,t} + \delta agrshare_{i,t} * distfrontier_{i,t} + e_{i,t} \tag{8.1}$$

The parameter of interest is δ, which captures the effect of an influx of agricultural workers into the formal economy interacted with a measure of backwardness. α_i is a country-fixed effect, τ_t is a time fixed effect, g_h and Y are real growth and real GDP, respectively. $distfrontier$ is the log difference between US GDP and the GDP of country i at time t.[75] The idea is that changes in the agricultural share matter when the country is "catching up" to US living standards, i.e. when $distfrontier$ is high, a decrease in the agricultural share should push up the saving rate. There is also variation in the agricultural shares of developed economies, but these changes probably do not constitute the same fundamental transformation that miracle economies undergo when converging to the frontier economy. This is why I focus on the interaction term, and my prior is that δ should be negative, i.e. a decline in the agricultural share, when the country is catching up, should push up the saving rate.

The coefficient of the interaction term is marginally significant at the 5% level, and supportive of my initial hypothesis. According to my regression results, a country whose real GDP per capita is one-fourth of US GDP increases its saving rate, on average, by roughly 2 % if the agricultural share drops by 10%. I run a robustness exercise where I use data from the WDI. I obtain very similar point estimates, although the standard errors are larger. The WDI contains more countries, but has a shorter time dimension. The limiting variable here is the agricultural share, for which the WDI relies on estimates from the international labor organization (ILO). Controlling for the share

adjusted exchange rates. This approach overestimates the GDP of less developed economies through the well-known substitution bias in consumption, see Hill (2000).

[74] In their influential paper Hsieh and Klenow (2007) argue that once measurement of investment (which is highly correlated with savings even in open economies) is carried out using domestic price indices instead of PPP adjusted measures the association between investment rates and economic growth disappears undermining the findings of Levine and Renelt (1992) or Sala-I-Martin (1997). They also mention in footnote 7 that saving rates are more closely related to income than investment rate. Saving rates are measured in real terms, which seems the wrong concept when focusing on precautionary savings. To get a sense of whether an agent saves much or little, we need not ask what fraction of income they put aside for a rainy day – this is nominal savings over nominal income. Similarly, whether a country is building up their net foreign asset position or not depends on the nominal trade balance. This is at odds with much of the literature in development accounting and growth accounting which insists on using "real" measures (Caselli, 2005; Alcalá and Ciccone, 2004).

[75] In general, I avoided the use of the Penn World tables PPP-adjusted GDP for the most part, but when computing the distance of a country to the frontier economy I need to rely on the estimates of real GDP in the PWT. The reason is that this comparison should be based on PPP to be comparable across countries.

Table 23: cross-country regression: saving rate

Regressor	$\hat{\beta}$	p-value
$realgrowth$	.1727 (.0442)	0.000
$realGDP$	.0001 (.0000)	0.486
$agriculturalshare$	−.0060 (.1177)	0.960
$distfrontier$	−.5387 (3.2132)	0.868
$agrshare * distfrontier$	−.0683 (.0349)	0.058
Country fixed effects	Yes	—
Time fixed effects	Yes	—
N = 1887		

Note: The dependent variable is the saving rate in the economy. Heteroskedasticity-robust and clustered standard errors (on the country level) are reported in parentheses. The data sources are the World Development Indicators, the Penn World Tables 9.0, and the GGDC data from the University of Groeningen.

of households below 30 years and above 65 years does not change the sign of the interaction term but the p-value increases to .45.

Much of the variation is contaminated by short run fluctuations. Therefore, I estimate another specification with 5-year non-overlapping averages. Instead of the interaction term, I now use lagged changes of the agricultural share as main explanatory variable. The idea is that households that leave the agricultural sector have initially a low weight in the aggregate because they are relatively poor, but since they catch up fast and save much of their income, they impact the aggregate saving rate with a lag.

$$s_{i,t}^{5y} = \alpha_i + \tau_t^{5y} + \beta_1 g_{i,t}^{5y} + \beta_4 distfrontier_{i,t}^{5y} + \beta_3 \Delta(agrshare)_{i,t}^{5y} + \delta\Delta(agrshare)_{i,t-5}^{5y} + e_{it}^{5y} \tag{8.2}$$

The results are much more convincing, and robust to controlling for demographics, which brings down δ to $-.4235$, which is still significantly different from zero at a significance level of 1%.

Similar specifications for trade balance and current account flows turned out to be less successful. This is not too surprising because i) trade balances are hard to measure. For example, compare the trade balance from the PWT with the data constructed by Jordà et al. (2017) and you will find substantial discrepancies, especially for Germany and Japan. Moreover, there is much variation in the trade balance for developed economies, with little variation in the agricultural share. To make matters worse, poor countries, especially in Africa, at times display very negative trade balances (easily below -20%), partly reflecting the effect of foreign aid inflows. All of this makes the trade balance a tricky measure.

Similarly, there is a measurement problem with current account flows and net foreign asset

Table 24: cross-country regression: saving rate, 5 year avg

Regressor	$\hat{\beta}$	p-value
$realgrowth$	$.5434$ $(.1786)$	0.004
$distfrontier$	-2.9446 (1.7600)	0.103
$\Delta(agrshare)$	$.0083$ $(.1944)$	0.966
$\Delta(agrshare)_{t-5}$	$-.4528$ $(.1545)$	0.006
Country fixed effects	Yes	—
Time fixed effects	Yes	—
N = 311		

Note: The dependent variable is the 5 year average of the saving rate in the economy. Heteroskedasticity-robust and clustered standard errors (on the country level) are reported in parentheses. The data sources are the United Nations Population Population Division, the Penn World Tables 9.0, and the GGDC data from the University of Groeningen.

positions. Unfortunately, current account flows are not available for most countries before 1990. Even the impressive dataset constructed by Lane and Milesi-Ferretti (2007) only starts in the 1970s and only for some countries.

Given the difficulty of detecting statistical relationships in cross-country regression, I am inclined to interpret the robust correlation between agricultural share and aggregate saving rates in favor of the main mechanism proposed in this paper: structural change is important to understand savings pressure in miracle economies. Next, I discuss other popular explanations that have served as an explanation for the relationship between growth, savings, and capital flows and contrast them with my framework.