Capital Flows and Risky Rural-Urban Structural Change*

Florian Trouvain
University of Oxford

December 2025

Abstract

I argue that rapid rural-to-urban structural change—and the inherent risk associated with households moving from traditional farming into urban economic activity—are central drivers of precautionary savings, generating capital outflows along the development path. I develop a tractable model that explores this argument. The key ingredients of the model are structural transformation away from agricultural production and households' uncertain, heterogeneous income-growth experience as they enter the urban sector. The model provides new insights into the link between consumption smoothing, precautionary savings, and the uneven distribution of catch-up growth across households. Consistent with the theory, I present evidence from China that highlights stark differences in savings behavior across rural and urban households.

Keywords: Capital Flows; Structural Transformation; Precautionary Savings; Inequality; Human Capital Risk; Urban–Rural Differences

*Many thanks to my advisors John Leahy, Linda Tesar, and Dmitriy Stolyarov. I thank Paco Buera, Chris Carroll, Steve Lugauer, John Laitner, Karol Mazur, Alessio Moro, Rachel Ngai, Pablo Ottonello, Damiano Sandri, and Yongs Shin for their insights. Funding from the Michigan Ross China Research Initiative is gratefully acknowledged. Email: florian.trouvain@economics.ox.ac.uk.

1 Introduction

In a seminal article, Lucas (1990) asks: "Why doesn't capital flow from rich to poor countries?" Even more troubling from the standpoint of standard macroeconomic theory is that capital tends to flow out of fast-growing emerging markets. These capital outflows constitute a major puzzle, as they seem to contradict a cornerstone of modern macroeconomics: the permanent income hypothesis (PIH). The PIH implies that households that are relatively poor today 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. Yet this prediction has been contradicted by virtually every miracle economy, including Japan, Taiwan, Korea, and, most recently, China.

The main contribution of the paper is to propose a novel theory in which structural change from traditional rural production into a modern urban sector generates precautionary savings pressure during a growth miracle, which can dominate consumption smoothing forces and induce capital outflows despite fast aggregate income growth. The key assumption is that urban income growth is risky and unevenly distributed, and rural-urban migrants do not know ex ante how successful they will be in the urban sector. I view this as a natural assumption as the economy undergoes structural reforms, which creates high average growth but simultaneously generates uneven outcomes across ex ante similar households.

Two main insights emerge from the model. First, modelling the growth miracle as an uneven and risky process reduces the standard consumption smoothing force by an order of magnitude, opening the door for risk to dominate consumption smoothing motives, generating savings pressure and capital outflows. This contrasts with the representative agent framework, where convergence growth leads to a powerful consumption smoothing motive that induces borrowing along the transition path. This pattern carries over to heterogeneous agent models, where households are assumed to grow at the same aggregate growth rate on average. However, once convergence growth itself is unevenly distributed, with ex ante uncertainty about who will be among the winners and losers as growth takes off, the effect of future aggregate growth on current consumption declines sharply.[1]

Second, the theory provides a simple explanation for the hump-shaped aggregate saving rates characteristic of miracle economies along the transition path. Representative agent models fail to capture this feature of the data. For standard preferences, income effects dominate, and a TFP-induced growth spurt leads to high consumption and little savings along the transition path.[2] In the present model, the aggregate strength of the precautionary motive depends on the fraction of agents accumulating precautionary savings and their income share in the economy, which is changing over time due to structural change. Initially,

[1]Uneven growth will induce an additional desirable feature in the form of a fast rise in tail inequality, consistent with the rapid rise of inequality in miracle economies. Such a rise in inequality in less than two or three decades is hard to generate in models where all agents grow at the same long-run growth rate, but easy with growth rate heterogeneity as explained in Gabaix et al. (2016).

[2]This ought not to be confused with the convergence dynamics induced by a one time shift in TFP, and the process of capital accumulation that follows. Here, I am focused on growth driven by incremental convergence in TFP over time, consistent with the data.

structural change adds to the savings pressure by reallocating households into the urban sector where they try to build up buffer stock savings given the risky nature of urban growth. Eventually, individual risk disappears, at which point the household's precautionary motive fades. When calibrated, the model generates transition dynamics for the current account consistent with the data.

I proceed in three steps. First, I motivate the theory by presenting macro and micro evidence consistent with the proposed mechanism. I show that rapid structural change is a defining feature of every growth miracle. In particular, the set of countries exhibiting fast growth and capital outflows in Gourinchas and Jeanne (2013) is precisely the set experiencing rapid reallocation toward non-agricultural activity. Second, I use micro data from China—a large miracle economy with high saving rates and capital outflows—to examine differences in asset accumulation across rural and urban households. I document that asset-to-income ratios and asset growth rates are systematically higher for urban households, both for total assets and for a narrower subset of safe assets. Combined with fast-paced structural change out of agriculture, these differences help rationalize the buildup of aggregate savings and safe assets during episodes of rapid catch-up growth. China is a case in point: while roughly 80 percent of workers were employed in agriculture in the early 1980s, this share fell to less than 25 percent by 2020. To my knowledge, this is the first paper to highlight the interaction between urban–rural differences in asset demand and rapid structural change as an explanation for the positive association between growth, savings, and capital outflows.

In the second step, I develop a tractable model of miracle growth that centers on the risky and uneven nature of the rural-urban transition. The environment is otherwise standard: infinitely lived households, perfect foresight over aggregate dynamics, and conventional preferences. I depart from the benchmark neoclassical model in two crucial ways. First, it features a two-sector economy with continuous structural change driven by productivity growth in the urban sector. Second, individual growth in the urban sector is risky, reflecting the high dispersion in returns to specialized human capital and the fact that households do not know ex ante where they will end up in the urban income distribution.[3] This assumption seems particularly appropriate during an episode of rapid structural transformation.

I model this risk rural-urban transition parsimoniously. New urban households experience fast income growth, exit this phase randomly according to a Poisson process, and then receive a permanent, mean-preserving human-capital shock. This stochastic process combines growth and risk while preserving analytical tractability. The model admits a simple sufficient statistic that signs the direction of capital flows along the transition paths, and intuitively summarizes how growth and risk shape consumption smoothing vs. precautionary savings.

The quantitatively important difference relative to the previous literature is that catch-

[3]This is not to say, however, that there are no shocks in the rural sector. Volatile weather conditions are one obvious source of risk for rural households. At the same time, in terms of idiosyncratic risk, a vast literature following Townsend (1994) has established that consumption insurance for rural households in developing economies appears almost perfect.

up growth itself is unevenly distributed: the duration of high growth on the household level is random, which substantially lowers expected risk-adjsuted lifetime income growth relative to aggregate growth. This weakens the consumption-smoothing motive that dominates representative-agent and canonical incomplete-markets models. Intuitively, given risk aversion, expected utility is largely shaped by low-growth realizations rather than the average growth experience. Aggregate growth, on the other hand, is much larger and driven by the emerging right tail of the urban income distribution populated by lucky households that stay in the high-growth regime for a long time.[4]

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 by a factor of 6 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.

Literature Review. The model draws heavily on insights developed in the literature on precautionary savings and incomplete markets (Bewley, 1977; Deaton, 1991; Imrohoroğlu, 1989; Huggett, 1993; Aiyagari, 1994; Carroll, 1997). I contribute to this literature by developing a tractable model that adds convergence growth and structural change. I stress that uneven growth is quantitatively important in taming the strong consumption smoothing motive during phases of catch-up growth.[5]

The focus on a dual-sector economy with differences in risk and opportunities across sectors relates to the classic stationary two-sector model of Harris and Todaro (1970) with urban unemployment risk,[6] and Lucas (2004) model of the rural-urban transition driven by human capital accumulation. I incorporate risky and uneven growth in the urban sector, which induces distinct transitional dynamics as countries embark on an episode of rapid convergence growth. Consistent with the idea that the urban sector is more risky is Townsend

[4]This stochastic process is able to quickly generate tail inequality due to growth-rate heterogeneity across households, see Gabaix et al. (2016), which is otherwise difficult to achieve in standard models where all households grow at the same rate.

[5]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 urban-rural differences and transitional catch-up growth. See Coeurdacier et al. (2019) for the difficulty of overcoming consumption smoothing motives during a phase of catch-up growth.

[6]In an unpublished manuscript, Carroll and Jeanne (2009) highlight the role of unemployment risk for fast-growing economies related to my framework. I stress that introducing uneven catch-up growth is quantitatively powerful, and conjecture that this modification would elevate their mechanism.

(1994), which finds a high degree of consumption insurance in the rural sector.[7] In contrast, uninsurable idiosyncratic income risk is large in modern market economies, (Heathcote et al., 2014), and even higher in fast-growing emerging markets (Chamon et al., 2013; Santaeulalia-Llopis and Zheng, 2018). Also relevant is the literature on the agricultural productivity gap (Restuccia et al., 2008; Caselli, 2005; Gollin et al., 2013). I will demonstrate how rural-urban wage gaps further help tame the consumption smoothing motive while boosting aggregate growth along the transition path.

Lastly, the paper relates to the literature on global imbalances. In particular, I tackle what Gourinchas and Jeanne (2013) coined the "allocation puzzle", referring to the fact that faster growing emerging markets display high saving rates and capital outflows, at odds with standard consumption smoothing forces. One strand of the literature argues that the flows occur due to developing economies' inability to produce safe assets or to manage risky ones (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. In contrast to their work, I focus on transitional growth miracles where the focus rests on the role of urban-rural differences in understanding the high savings pressure along the transition path. I consider a growth miracle that pushes up GDP per capita by a factor of six, consistent with the Taiwanese experience, and larger than what previous work has considered.[8]

An important contribution in this context is Song et al. (2011), which studies the Chinese growth miracle and generates current account outflows along the transition path. The key friction is the inability to channel domestic savings to high-return domestic enterprises, similar to Buera and Shin (2017).[9] I view my work as complementary to papers focusing on 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. Since households, on average, feature high income growth, any quantitative infinite horizon model will induce a strong borrowing motive for these households, in contrast to the data.[10] I view my focus on the risky rural-urban transition as complementary to the large literature on financial frictions.

[7]See Rosenzweig and Stark (1989) for the role of informal rural insurance networks in smoothing consumption. Bryan et al. (2014), Munshi and Rosenzweig (2016), Morten (2019), Meghir et al. (2022) and Lagakos et al. (2023), which study the interplay between migration and insurance.

[8]Caballero et al. (2008) and Mendoza et al. (2009) mostly abstract away from transitional growth spurts. Buera and Shin (2017) and Sandri (2014) consider transitional growth dynamics that are much smaller than the ones considered here.

[9]In China, Bai (2009) documents that the lion's share of savings is due to private households relative to savings by the government and corporations. This is important because it highlights that the lion's share of savings does not stem from corporations, while admitting that the distinction between firms and household savings is blurry.

[10]In the Chinese context, Chamon and Prasad (2010) considers several possible drivers of high household saving rates, highlighting that demographics appear relatively unimportant. See also Carroll and Summers (1991) for the unimportant role of demographics in household savings. See Coeurdacier et al. (2015) for an alternative take that stresses the role of demographics. İmrohoroğlu and Zhao (2017, 2018) argue how the one-child policy in China raised aggregate savings. Wei and Zhang (2011) posit that the high Chinese saving rates are driven by the gap in the sex ratio.

The rest of the paper proceeds as follows: Section 2 provides a set of stylized facts related to structural change and urban-rural differences in savings. Section 3 develops the theory that connects those facts and studies the tradeoff between catch-up growth and risk. Section 5 provides a quantification and computes the transition dynamics. Section 6 concludes.

2 Stylzed Facts

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 partially known, I relate them to the fast-paced structural change and in particular show that the "allocation puzzle" (Gourinchas and Jeanne, 2013) is entirely accounted for by emerging markets exhibiting fast-paced structural change. I then analyze household level panel data from China to provide micro evidence consistent with the urban-rural differences in saving behavior. In what follows I use the terms city versus countryside, urban versus rural, and agricultural versus non-agricultural interchangeably.

2.1 Growth, Capital Flows, and Structural Change

Figure 1 is a version of the main figure in the influential paper of Gourinchas and Jeanne (2013), which revisits the negative link between productivity growth and capital inflows. While the left panel plots the familiar puzzling negative relationship between productivity growth and capital inflows as in Gourinchas and Jeanne (2013), 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. Comparing left and right panel, it is easy to see that the allocation puzzle is entirely driven by countries that experience fast structural change out of agriculture. This pattern suggests that structural change is not merely correlated with capital outflows, but may fundamentally alter household saving incentives during catch-up growth.

The work of Gourinchas and Jeanne (2013) is also important as it justifies the focus of the present paper on the household side.[11] Gourinchas and Jeanne (2013) finds that investment is relatively undistorted in fast-growing economies, relative to the household savings. That is to say, investment is in sync with the standard neoclassical model, but household savings are counterfactually large.

The countries driving the allocation puzzle in Figure 1 are the so-called "miracle economies" (Lucas, 1993): the East Asian tigers—Japan, Korea, and Taiwan—and, more recently, China, all of which experienced large and sustained per capita growth. Figure 2 highlights the exceptional growth spurt sustained over several decades relative to the United States, which represents the technological frontier. The figure plots the time series for each miracle economy, and complements the simple cross-sectional plot in the previous figure. The figure highlights the co-movement between fast-paced structural change out of agriculture and

[11] I largely abstracts away from capital accumulation dynamics by maintaining a small open economy setup with fixed interest rate in the main part of the paper. See the appendix for an extension with capital.

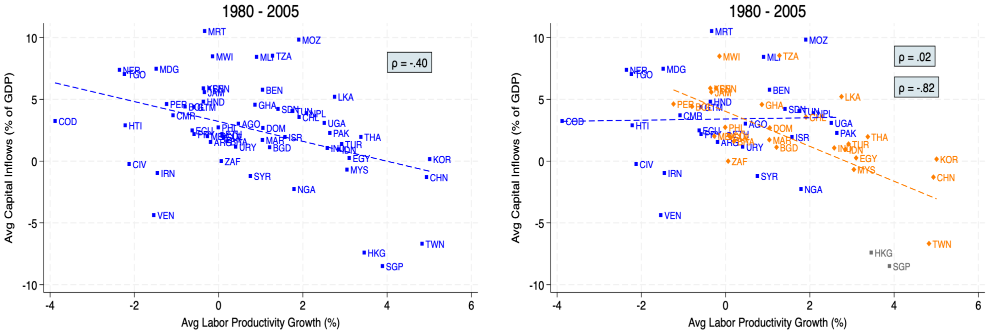

Figure 1: Relationship between average capital inflows and average labor productivity growth for a cross section of emerging markets following Gourinchas and Jeanne (2013). Orange diamonds represent emerging markets with above average decline of the agricultural employment share. Current-account-to-GDP 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. Hongkong and Singapur are dropped when computing the regression lines for the split sample due to their special role as financial centers. The correlation between capital inflows and productivity growth for the full sample is $-.42$. For the split sample, the correlation is $-.82$ and 0.02 for countries with fast and slow structural change, respectively.

convergence growth.[12]

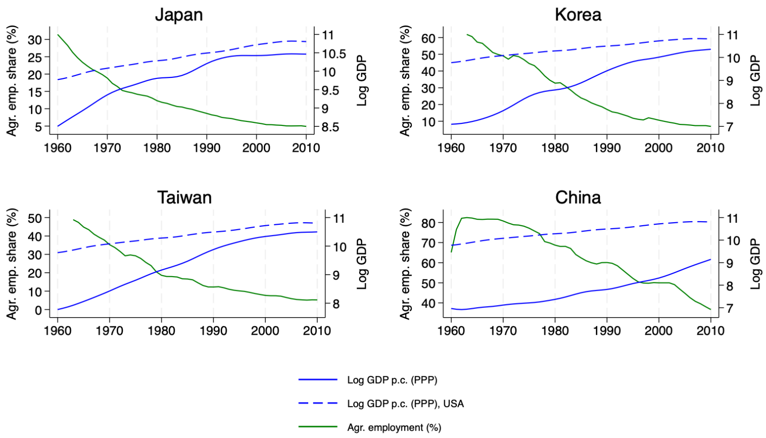

Figure 2: Relationship between agricultural employment share and convergence in GDP for miracle economies. GDP series in purchasing power parity taken from the Penn World Tables 9.1, and smoothed using an hp-filter with smoothing parameter of 8.5.

Figure 3 shows that miracle economies exhibit elevated saving rates that, for most countries and time periods, coincide with current account surpluses and hence net capital outflows. Aggregate saving rates display a characteristic hump-shaped pattern: although investment rates are high, they are consistently dominated by even higher saving rates. By the balance-of-payments accounting identity, these excess savings translate into current account surpluses. This hump-shaped saving profile—common across miracle economies and

[12]The 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.

also operating in China, where saving rates peaked in the 2010s and have since declined—is difficult to reconcile with standard preferences and benchmark models of convergence.[13]

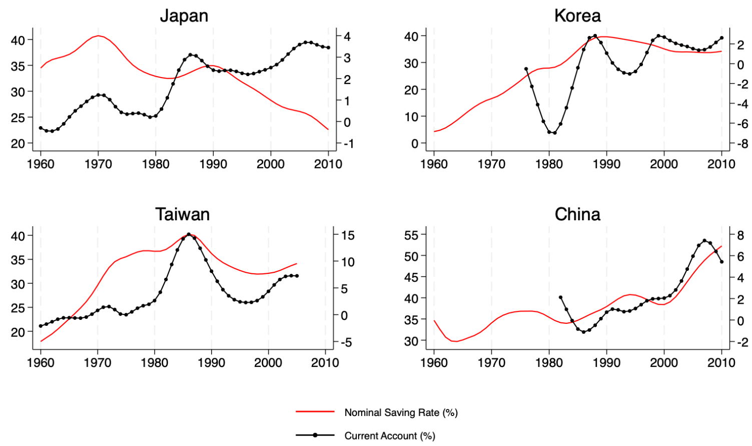

Figure 3: Aggregate nominal saving rate (left axis) and current account (right axis) for Japan, Korea, Taiwan, and China. Series are smoothed using an hp-filter with smoothing parameter of 8.5.

In the benchmark neoclassical model, rapid catch-up growth implies consumption smoothing, low saving rates, and large current account deficits. Neither of these predictions are borne out in the data.[14] My framework notes that fast productivity growth also induces rapid structural change, which, in turn, could be a risky and uneven process on the household level inducing strong precautionary savings motives. An immediate implication is that urban households should build up safe assets, relative to rural ones. I investigate this in the next section using household panel data from China.

2.1.1 Urban-Rural Differences in Asset Accumulation

I next document differences in asset accumulation across rural and urban households in China using data from the Chinese Family Panel Study (CFPS). The CFPS is a nationally representative household panel that contains detailed information on income, expenditures, assets, and demographics and is well suited to studying household saving behavior during China's growth transition. I use the CFPS to compare savings and asset accumulation across rural and urban households.

To assess whether urban households save more than rural ones, I focus on asset-to-income ratios following Carroll and Samwick (1997). This measure is useful for three reasons.

[13]In the benchmark neoclassical model, the dynamics tend to be monotone, see the textbook treatment by Barro and Sala-i-Martin (2003). If the convergence is driven by TFP and not capital accumulation, then the aggregate saving rate tends to be negative due to strong income effects for standard preferences. One classic explanation is consumption habit, see Carroll et al. (2000). However, consumption dynamics in the micro data seem to contradict this explanation, see Chamon and Prasad (2010).

[14]There is an important distinction between safe assets and more risky FDI flows, as well as private vs. public flows, see Aguiar and Amador (2011). My framework focuses on safe assets, which is driving the capital outflows out of fast growing emerging markets, and I abstract away from the distinction between private and public flows.

First, assets are stock variables that summarize long-run saving behavior and are less noisy than flow measures of savings (Dynan et al., 2004). Second, scaling by income removes the mechanical effect that richer households save more in levels. Third, in models with precautionary savings, buffer-stock wealth is proportional to permanent income, making asset-to-income ratios a natural statistic for inferring the role of risk (Carroll, 1997). The same relationship holds in the model developed in the next section.[15]

I leverage the panel dimension of the dataset by computing average household asset holdings and average household income over the sample period. This ensures that the income measure is closer to a notion of permanent income, and reduces measurement error. I use a broad concept of net labor income that includes business income, income from agricultural activity, and wage income after taxes and transfers consistent with the model. I then compute asset-to-income ratios. I distinguish between total assets –net household wealth including housing, business assets, and durables– and safe assets, defined as financial assets such as deposits, bonds, stocks, and cash. The focus on safe assets is motivated by the fact that this asset class is central to capital outflows from emerging markets (Bernanke et al., 2005; Aguiar and Amador, 2011; Gourinchas and Jeanne, 2013).

The analysis restricts attention to households whose head is between ages 27 and 62, participates in the labor force, appears in at least three survey waves, experiences no household splits, and remains in the same sector throughout the sample. These restrictions ensure that differences in asset accumulation are not driven by life-cycle effects or changes in household composition. Additional details and summary statistics are provided in the appendix.

I estimate the following cross-sectional specification

$$\frac{b_i}{y_i} = \alpha + \beta D_i + \Gamma' X_i + \epsilon_i, \tag{2.1}$$

where $\frac{b_i}{y_i}$ denotes the household asset-to-income ratio, D_i is an indicator for urban households, and X_i includes education, household size, and demographic controls. The coefficient β captures whether urban households hold higher asset-to-income ratios than rural households. I estimate this equation using median regressions, as ratio-based measures are highly skewed (Dynan et al., 2004; Fagereng et al., 2019).[16]

Figure 4 plots the implied median asset-to-income ratios for rural and urban households based on equation (2.1) without controls. Urban households hold substantially higher safe asset-to-income ratios, with a median of about 0.6 compared to 0.25 for rural households, despite urban income being more than twice as high. Thus, urban households not only earn more but also accumulate disproportionately larger precautionary asset buffers.

[15] To understand the tight link between asset-to-income ratios and household savings, consider a model where households save a constant fraction of their labor income y, labor income grows at rate g, and households start out with zero assets that earn return r, giving rise to the flow budget constraint $\dot{b}_t = rb_t + sy_t$. In this case, the asset-to-income ratio converges in the long run to $\frac{b}{y} = \frac{s}{g-r}$ for $r < g$. This simple model highlights how asset-to-income ratios are proportional to the saving rate.

[16] If assets and income were both normally distributed, the ratio would follow a Cauchy distribution. In that case, the mean would not be well defined while a median regression remains a consistent estimator for

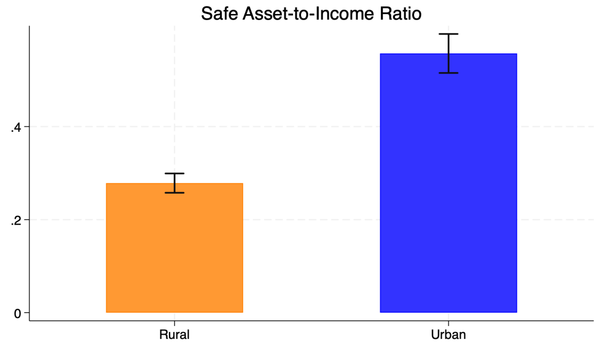

Figure 4: Cross sectional rural-urban median differences in safe asset-to-income ratios based on CFPS with 95% confidence intervals.

While rural–urban differences do not establish causality, the gap is robust to controlling for age, education, and other demographics, with results in table 1.[17] The rural-urban gap in safe asset-to-income ratio is robust and remains highly significant well below the 1 % level. Results for total assets, as well as estimates using sampling weights, are reported in the appendix and yield similar gaps. Through the lens of an incomplete market model, it appears that urban households have a stronger precautionary motive.

Table 1: Rural-Urban Median Differences for Safe Assets

	(1) safe_asset_income_ratio	(2) safe_asset_income_ratio	(3) safe_asset_income_ratio
urban dummy	0.279***	0.264***	0.168***
	(0.0214)	(0.0198)	(0.0229)
_cons	0.278***	-0.441**	-0.506**
	(0.0106)	(0.207)	(0.225)
age	No	Yes	Yes
sex and hh size	No	No	Yes
education	No	No	Yes
Observations	5957	5957	5945

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.

An implication of models with buffer-stock savings is that households that recently migrated to the urban sector should i) have a lower asset-to-income ratio, and ii) build up buffer-stock savings at a relatively higher rate, so as to catch up with households that have been in the urban sector for longer. I test this prediction using information on households' hukou status. The hukou status is either agricultural or non-agricultural, and access to public services is more or less tied to the local hukou. Initially, hukou status and rural-urban sectoral classification were highly correlated but both illegal migration and a relaxation of the stringency of the system allowed for households with agricultural hukou to move to urban areas. In contrast, changes in hukou status are rare as the hukou status is inherited from

differences across groups.

[17] A long-standing question in the literature is whether high-income households save more, see Dynan et al. (2004), De Nardi (2015), Straub (2019), and Fagereng et al. (2019). Since income is in the denominator, and urban households tend to have higher income, controlling for income amplifies the difference.

the mother. Consequently, the strict overlap between rural vs. urban and agricultural vs. non-agricultural hukou no longer exists although the correlation between these two measures in the CFPS is still 0.6.[18]

I focus on urban households, and use the agricultural hukou as an imperfect proxy for whether a household started out in the rural sector. I then test the following prediction: among urban households, the ones with a rural hukou are more likely to be newcomers, i.e., more recent rural-urban migrants. If so, and conditional on income and demographics, these households should accumulate assets at a higher rate, while lagging behind with their asset levels relative to urban incumbent households.[19] I explore this prediction by computing annualized asset growth rates using the first and last time a household appears in the sample, and regress these growth rates on the hukou status, where I control for age, gender, education, income, and income growth in a long difference specification. This specification partly addresses concerns that the previous differences were solely driven by selection as I am now leveraging differences driven by households that likely moved across sectors. Table 2 reports the results. Urban households with rural hukou exhibit faster asset growth, both for safe and total assets, conditional on income and income growth. The evidence is consistent with households self-insuring against income risk in the urban sector, with new entrants accumulating precautionary wealth that incumbents have already built. As rural households continue to transition into urban employment, this mechanism can generate sustained aggregate saving pressure and capital outflows.

Table 2: Asset Growth and Hukou Status in Urban

	(1)	(2)	(3)	(4)
	g_safe_asset	g_safe_asset	g_total_asset	g_total_asset
rural hukou dummy	0.0590**	0.0430*	0.0460***	0.0332**
	(0.0250)	(0.0257)	(0.0158)	(0.0166)
_cons	-0.0396**	0.435*	-0.0334***	0.446***
	(0.0167)	(0.228)	(0.01000)	(0.121)
controls	No	Yes	No	Yes
Observations	2267	2065	2267	2065

Note: The dependent variable is annualized growth in the respective asset class, i.e., $g_a = \frac{\log a_{t+k} - \log a_t}{k}$, and income growth is defined analogously. I only consider urban households in this regression, impose the same sample restrictions as before, and additionally drop the 2% of households with highest asset growth rates to ensure the results are not driven by outliers. Robust standard are errors in parentheses. *, **,*** denote statistical significance at 1, 5, and 10 percent level. The use of the logarithm makes clear that households with negative assets are dropped from the analysis, which concerns about 2% of the sample after applying the other selection criteria.

Evidence on Income Risk. A growing literature documents that the Chinese growth miracle is associated with unusually high household income risk (Chamon et al., 2013; He

[18] See Chan and Buckingham (2008) for further information, and a discussion of reforms in the hukou system in the 2000s.

[19] Note that there is only a small number of rural-urban migrants in the sample insufficient for statistical inference.

et al., 2018; Santaeulalia-Llopis and Zheng, 2018). I build on this work by emphasizing two additional features. First, miracle growth is unevenly distributed: households do not share a common long-run growth rate. This form of growth-rate heterogeneity is essential for weakening consumption-smoothing forces and is consistent with the rapid rise in top-tail inequality observed in China (Piketty et al., 2019; Gabaix et al., 2016). Second, rural–urban transitions are intrinsically risky: migrants face substantial ex ante uncertainty about their long-run outcomes despite high average urban wages.[20]

I next develop a stylized and tractable model of the rural–urban transition that isolates the core mechanism at work. The model illustrates how income risk and growth-rate heterogeneity interact with structural change to generate elevated saving pressure along the transition path. Its purpose is not to tightly fit income and asset dynamics, but to provide transparency and intuition for how risk, uneven growth, and sectoral reallocation jointly help explain the high saving rates observed in miracle economies. More computationally intensive models can deliver a closer quantitative fit, but often obscure the underlying economic forces that this stylized framework is designed to highlight. I pursue such a richer model in follow up work.

3 Theoretical Framework

4 A tractable model of the rural-urban transition

This section presents a tractable model that isolates the mechanism emphasized in the data: (i) structural change (rural $\rightarrow$ urban reallocation) together with (ii) risky, uneven household-level catch-up growth generate strong precautionary saving motives, a hump-shaped aggregate saving rate, and net capital outflows. The model is intentionally stylized so that the core mechanism is explicit and derivations remain transparent.

4.1 Environment

Time is continuous, $t \geq 0$. A unit mass of infinitely lived households indexed by $i \in [0, 1]$ supply labor inelastically. There are two sectors: rural and urban. Rural households decide on whether to permanently migrate to the urban sector. Rural households are hand-to-mouth as in Moll (2014) while urban households solve a forward-looking consumption-saving problem. Let L_t^r and L_t^u denote the share of households in rural and urban sector so $L_t^r + L_t^u = 1$.

Technology and Market Structure.

Aggregate output equals the output of each sector, $Y_t = Y_t^r + Y_t^u$, where the price is normalized to one and each sector produces the same final output good. Rural production exhibits

[20] In this context, Baseler (2023) and Porcher et al. (2024) provide micro evidence on information frictions associated with migration in developing economies. In work-in-progress, I use a structural dynamic model, and find that permanent income risk is systematically higher in urban relative to rural South Africa, informal insurance is lower in urban South Africa, and that there is much uncertainty ex ante in terms of how well rural-urban migrants do in the urban sector. This latter insight is intuitively linked to the fact that not all households do better in the urban sector, despite the fact that average wages are much higher in urban.

diminishing returns due to fixed land (normalized to one)

$$Y_t^r = A_0^r (L_t^r)^\chi, \qquad \chi \in (0,1), \tag{4.1}$$

and rural per-worker compensation (including implicit land return) is $w_t^r = A_0^r (L_t^r)^{-(1-\chi)}$, in the spirit of Lewis (1954) where workers receive their average, not marginal, product, and there is no heterogeneity across workers as each of them is endowed with one efficiency unit of labor.

Urban production is constant-returns-to-scale in effective labor with productivity A_t^u

$$Y_t^u = A_t^u \int_{i \in L_t^u} h_t^i \, di, \tag{4.2}$$

where $h_{i,t}$ are the efficiency units of labor of household i in urban activity. Production is perfectly competitive so that workers earn their marginal product and the wage per efficiency unit equals $w_t^u = A_t^u$.

Preferences, budget constraint, and migration.

Households have CRRA flow utility with risk aversion $\eta > 0$ and discount rate ρ. Rural households consume their labor income each instant: $c_t^r = w_t^r$. Rural households face a migration decision. They are free to leave the rural sector, and do so whenever the expected value in urbane exceeds the value of staying in the rural sector net of a utility cost of migration τ. The optimality conditions have to satisfy the following two equations

$$\rho V_t^r = \frac{(w_t^r)^{1-\eta}}{1-\eta} + \dot{V}_t^r, \tag{4.3}$$

$$V_t^r \geq \frac{V_t^u}{\tau^{1-\eta}}. \tag{4.4}$$

where $\tau^{1-\eta}$ scales the utility cost appropriately such that $\tau > 1$ can be interpreted as a wedge in consumption equivalents. Note that rural households enter the urban sector with zero assets.

Urban households choose consumption c_t and asset holdings a_t to maximize utility

$$V_t^u = \mathbb{E}\left[\int_t^\infty e^{-\rho(s-t)} \frac{c_s^{1-\eta}}{1-\eta} \, ds\right] \tag{4.5}$$

subject to a standard intertemporal budget constraint and the no-Ponzi condition

$$\dot{b}_t = r b_t + y_t - c_t, \tag{4.6}$$

where b_t represents holdings of a risk-free internationally traded bond that yields return r, and y_t and c_t denote income and consumption of an urban household. The economy is small relative to the rest of the world (ROW) so that the interest rate is exogeneous.

Aggregate asset accumulation, i.e., aggregate savings, in the small open economy equal

$$\int_{i\in L_t^u} \dot{b}_{i,t}\, di = Y_t^u + rB_t - C_t^u, \tag{4.7}$$

where B_t is aggregate bond holdings of the small open economy.

Growth and Risk.

The productivity of the urban sector after the emerging market "reforms" at time zero in ways left unspecified grows at rate g, which is the same as for the ROW

$$A_t^u = A_0^u e^{gt}. \tag{4.8}$$

Given a fixed factor in the rural sector, productivity growth in the urban sector induces structural change consistent with the importance of pull-factors during early stages of the development process (Alvarez-Cuadrado and Poschke, 2011; Hnatkovska and Lahiri, 2018).[21]

A household that migrates at time $t_{m,i}$ enters a temporary high-growth phase during which its (log) income grows at a rate $g_h > g$. The duration of this high-growth phase is stochastic, with households exiting the fast-growth process at a Poisson arrival rate λ. Upon exit at time T_i, the household draws a multiplicative type $\varphi_i > 0$ from distribution $F(\varphi)$, which can be interpreted interchangeably as a human capital shock or an inequality shock. Following this draw, the household's income grows at the ROW rate g and is subject to no further idiosyncratic uncertainty. Thus log income is

$$\log y_t^i = \begin{cases} \log A_t^u + (g_h - g)(t - t_{m,i}), & t_{m,i} \le t \le T_i, \\ \log A_t^u + (g_h - g)(T_i - t_{m,i}) + \log \varphi_i, & t > T_i. \end{cases}$$

Remark. The stochastic process implies exponential, memoryless waiting times in the high-growth regime, which keeps the distributional dynamics tractable. This structure closely resembles a"Luttmer rocket" (Luttmer, 2011), here applied to household income growth rather than firm size. Upon exiting the high-growth regime, households draw a multiplicative type φ, generating a permanent inequality shock and cross-sectional dispersion in income levels. The resulting income process cleanly separates two forces: uneven growth during the transition and permanent inequality afterward. This stochastic process captures a key feature of structural transformation—human capital differences are largely irrelevant in agriculture but become first-order in urban production, where ex ante similar workers can experience sharply different outcomes.

Competitive equilibrium.

A competitive equilibrium consists of prices $\{w_t^r, w_t^u, r\}$, rural and urban shares $\{L_t^r, L_t^u\}$ consumption allocations and income $\{c_{i,t}, y_{i,t}\}$, migration decisions $\{t_{i,m}\}$ and joint time-

[21]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).

varying distribution over (t_m, T, b, φ) such that households optimize given prices and stochastic processes, firms maximize profits, labor markets clear, and the joint distribution evolves consistently with individual decisions and the stochastic process. This equilibrium will be generically non-stationary with declining rural share.

4.2 Solving the Model

The model can be conveniently solved backwards, focusing on the urban sector first, and then analyzing the rural migration problem. Note that the ROW grows at rate g and all uncertainty on the household level is already resolved, so that the standard Euler equation implies $r = \eta g + \rho$.

Urban household problem.

After the type draw ($t \geq T$) the household faces deterministic growth at rate g and no idiosyncratic risk, so consumption growth obeys the standard Euler equation, $\frac{\dot{c}_t}{c_t} = \frac{1}{\eta}(r-\rho) = g$. Using this, the post-draw value function can be written in closed form,

$$V^u(b_t, y_t(\varphi, T)) = \frac{\big(y_t + [g(\eta-1)+\rho]b_t\big)^{1-\eta}}{(1-\eta)\big(g(\eta-1)+\rho\big)}, \tag{4.9}$$

where $y_t(\varphi, T)$ makes clear that income depends on both the type draw φ as well as the time spent in the high-growth regime, T. Whilst the household is in the high-growth regime, I derive in the appendix that the household value function equals

$$V_t^u = \max_{\{c_s\}_{s\geq t}} \int_t^\infty e^{-(\rho+\lambda)(s-t)} \Big(\frac{c_{s-t}^{1-\eta}}{1-\eta} + \lambda\, \mathbb{E}_\varphi \left[V\left(b_{s-t}, y_{s-t}(\varphi, T)\right)\right] \Big) ds. \tag{4.10}$$

Applying standard optimal-control techniques after substituting (4.9) into (4.10) yields optimal consumption growth in the high-growth regime

$$\frac{\dot{c}_s}{c_s} = \frac{r-\rho}{\eta} + \frac{\lambda}{\eta}\Big\{ \mathbb{E}_\varphi \Big[\Big(\frac{[g(\eta-1)+\rho]a_s + \varphi y_s}{c_s} \Big)^{-\eta} \Big] - 1 \Big\}. \tag{4.11}$$

This Euler equation is key. Relative to the balanced-growth rate $\frac{r-\rho}{\eta} = g$, consumption growth while in the high-growth spell is adjusted by an expectation term that will capture both consumption smoothing forces due to elevated income growth, and precautionary motive induced by uncertainty embodied in φ. 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 where y_s is the left limit of the income process and $\lim_{\Delta\downarrow 0} y_{s+\Delta} = \varphi y_s$ is the right limit when the type shock hits at time s. Before I discuss several propositions that can be derived from this modified Euler equation, I solve the rural households problem.

Rural household problem.

Note that the value of moving to the urban sector, up to the constant growth terms g, always looks the same, and I can write this urban expected value in normalized form, $V_t^u = v^u e^{(1-\eta)gt}$. Intuitively, households undergo the same expected transition in the urban sector up to the constant growth term. Next, note that in any equilibrium with rural-urban structural change, it must be that $V_t^u = V_t^r$. By contradiction, note that if $V_t^u > V_t^r$, a positive mass of households would instantaneously move to the urban sector up until the value of moving and staying is equalized. If $V_t^u < V_t^r$, households would stay in the rural sector, but since productivity growth makes the urban sector more attractive, there would eventually be a point when the two values are equal, and from then on, households are indifferent between staying or migrating, which is sustained by a continous rural outflow that induces wage growth in the rural sector in lock-step with productivity growth in the urban sector. The value of staying in the rural sector thus equals

$$V_t^r = \frac{(w_t^r)^{1-\eta}}{(1-\eta)\big(g(\eta-1)+\rho\big)}. \tag{4.12}$$

The rural outflow necessary that keeps households indifferent between staying or migrating simply follows from finding the decline in the rural share that delivers wage growth at rate g given by

$$\frac{\dot{L}_t^r}{L_t^r} = -\frac{g}{1-\chi}. \tag{4.13}$$

Lastly, note that I can match any rural share at time zero, and any rural-urban wage gap, by using the free parameters τ and A^r.[22] Figure 5 sketches out the evolution of income for a random household that migrates from rural to urban, which summarizes the key elements of the model.

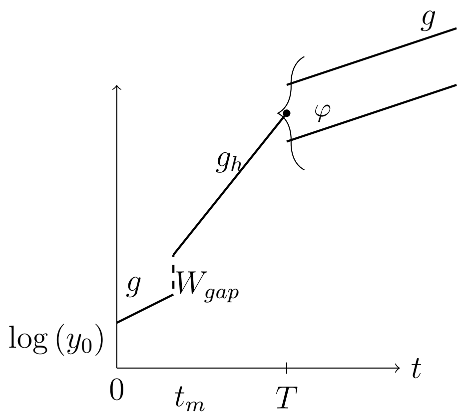

Figure 5: Income process from household starting in the agricultural sector

I next analyze theoretical whether this model features capital inflows or outflows along the transition path.

4.3 Precautionary Savings vs. Consumption Smoothing

The first result, although well known (Schechtman, 1976; Gourinchas and Parker, 2002), is worth pointing out again. For CRRA preferences with coefficient of relative risk aversion

[22]Finding the right parameters clearly depends on the value of v^u, but for the purposes of the paper, it does not matter what these are, since L_0^r and $\frac{w^u}{w^r}$ are directly taken from the data.

above one, the marginal utility of consumption at zero is infinity. Therefore, agents will never borrow.[23]

Proposition 1. *Let $\underline{\varphi}$ be the lower bound of the support of φ. Then, agents' borrowing decisions in the high growth regime respects the following inequality:* $\frac{b_t}{y_t}\left[(\eta-1)\,g+\rho\right] > -\underline{\varphi}$.

Proof. By contradiction, suppose an agent borrows above the borrowing level. Then there is a range of values $\varphi \in [\underline{\varphi}, -\frac{b_t}{y_t}\left[(\eta-1)\,g+\rho\right]]$ where the agent would have to consume weakly below zero. 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.[24]

Proposition 2. *Consumption growth for households with inequality shock and growth rate uncertainty is larger than consumption growth for households with only growth rate uncertainty, which in turn is larger than consumption growth for households without any risk.*

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 using Jensen's inequality.[25] 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.

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

[23]The differential equations, and especially the law of motion of consumption (**??**) 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.

[24]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.

[25]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.

Proof. See appendix. □

Proposition 3 shows that random convergence alone cannot generate the capital flow patterns we observe in the data. It highlights the need for an additional source of risk in order to explain the puzzle. The intuition for this result is as follows: If there is no human capital risk in the form of the type draw, then the only risk that households are exposed to is the random time they spend in the high growth regime. This type of risk, however, only represent upside risk and is therefore not sufficient to induce precautionary savings. At any point in time, a household will be better off in the future, no matter how long they are in the high growth regime and thus will want to borrow against future income. In contrast, in a model with human capital risk, some households can actually be worse off, at least for some time, despite strong convergence growth. Only this type of risk can leads to a precautionary savings motive that can dominate the consumption-smoothing motive.

The next proposition deals with the case when expectations about convergence growth are biased. One way to shut down the consumption smoothing motive of households is to make them believe that there is no convergence growth. My model allows me to consider this case effortlessly. Let $\tilde{\lambda}$ be the households' belief about the arrival rate of the low-growth regime, i.e. the relevant parameter for the household Euler equation. A high $\tilde{\lambda}$ relative to the correct λ represents "pessimistic" expectations in terms of convergence growth since households think their convergence period is, on average, shorter than it actually is.[26]

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. As long as the future looks 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. Once human capital risk is incorporated, however, downward biased expectations help to generate capital outflows as it weakens the consumption smoothing motive.

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 if and only if the following inequality holds*

$$\frac{g_h - g}{\lambda}\eta < \mathbb{E}_\varphi\left[\left(\frac{1}{\varphi}\right)^\eta\right] - 1 \tag{4.14}$$

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

Proof. See appendix. □

Proposition 5 is the main result of the theory section. 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 5 represents the consumption smoothing force, while the right hand side reflects human capital risk. Stronger convergence growth governed by the convergence rate $g_h - g$ 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 4.14 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.[27] In the next subsection I provide a simple calibration of the model that assumes a log normal distribution with $\log(\varphi) \sim N(-\frac{\sigma^2}{2}, \sigma^2)$ but the proposition generalizes beyond this case.

There exist no closed form solution for the transitional dynamics during the high-growth phase of a household. 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 (**??**), similar to the analysis of the neoclassical growth model in continuous time. 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 5 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. 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.

4.4 Difference to Canonical Precautionary Savings Models

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

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

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

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.[28] 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 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 8 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

[28] 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).

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.

5 Calibration

I next quantify the model to solve for the aggregate trajectory of the economy along the transition path. 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 upon migration, after netting out the effect of the urban-rural wage gap, equals

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

where I use both the independence of the type draw, as well as the fact that the Poisson process leads to Pareto-distributed income due to catch-up growth. 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%![29]

I set the coefficient of relative risk aversion to $\eta = 2$. I model the inequality shock as a draw from a Log-Normal distribution, i.e. $log(\varphi) \sim N(-\frac{\sigma^2}{2}, \sigma^2)$. This particular representation ensures that $\mathbb{E}[\varphi] = 1 \, \forall \sigma$, 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.[30] 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 that the variance of the log of income for the stationary distribution is $\left(\frac{g_h - g}{\lambda}\right)^2 + \sigma^2$, 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).[31]

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	$g_h = 7\%$
urban-rural wage gap	$W_{gap} = 100\%$
elasticity of agr. output with respect to labor	$\chi = \frac{5}{9}$
initial agr. share	$L_0^r = 75\%$
initial share of agents that know their type	$M_{0,1} = 17.5\%$

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. 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

[29]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%.

[30]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.

[31]Actually, the relevant statistic here is provided by De Magalhães and Santaeulàlia-Llopis (2018b) 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.

this mass point would dominate the dynamics of aggregate savings completely.[32]

Lastly, I need to set the parameter χ 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 χ through the lens of the model. Recall that $L_t^r = L_0^r \exp\left(-\frac{g}{1-\chi}t\right)$ describes the evolution of the rural share. This law of motion of agricultural employment leads to the following estimating equation,

$$\log\left(L_{t,c}^r\right) = \beta_{0,c} + \beta_1 * t + \epsilon_{t,c}, \tag{5.1}$$

where I added a random error term. The estimating equation (5.1) 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 **??** 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 $\hat{\chi} \approx \frac{5}{9}$.[33] Figure 6 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).[34]

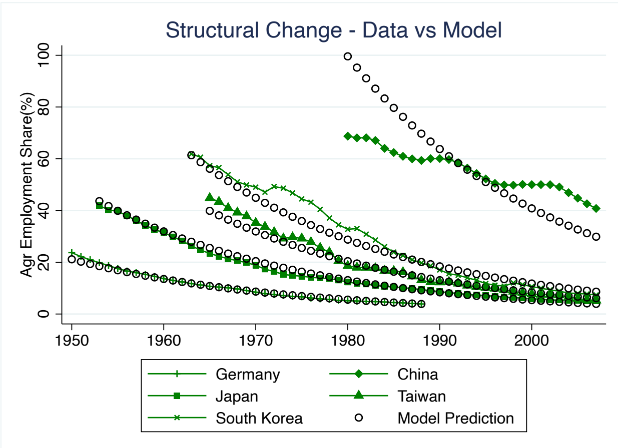

Figure 6: Prediction based on random effects model.

[32]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.

[33]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.

[34]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.

5.1 Growth dynamics

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

$$g_{agg}(t) = \underbrace{(g_h - g)\frac{Y_{0,t}}{Y_t} + \left(\frac{g}{1-\chi}\right)\frac{W_{gap}L_t^r}{Y_t}}_{\text{catch-up growth}} + \underbrace{g}_{\text{long run growth}} \tag{5.2}$$

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 rural-urban wage gap contributes to convergence growth. In the long run, the aggregate growth rate will of course be equal to g. Figure 7 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, consistent with findings in the literature (Santaeulalia-Llopis and Zheng, 2018). 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 7 also shows the convergence in terms of log output per capita relative to the United States.

5.2 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 is given by $\frac{g_h - g}{\lambda}$.

Figure 8 shows the CDF of the log of normalized income for different decades. Importantly, while the distribution fans out overall, more and more weight is being shifted to the right tail that is composed of households that remained in the high growth regime for a relatively long time. The normalized stationary distribution of the log of income is given by the exponentially modified Gaussian distribution. This emerges as the sum of two independent random variables, one of which is normal (what I call the type draw) and one of which is exponentially distributed (time spent in high growth regime scaled by the growth rates). Armed with this CDF I can compute the log variance of income for non-agricultural households displayed in table 3. The results are broadly inline with the fast rise of inequality

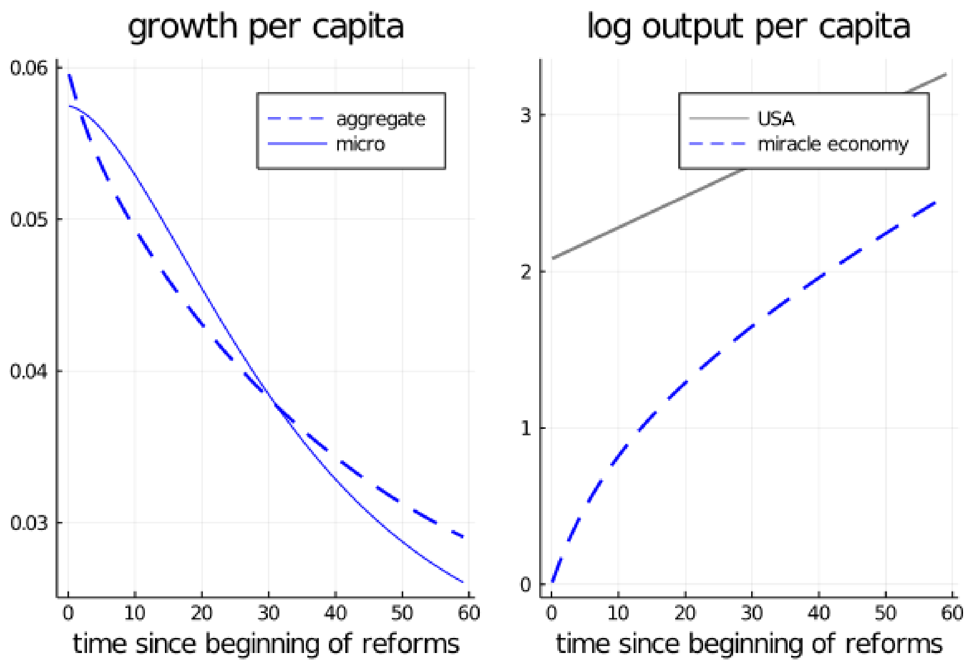

Figure 7: Per capita growth rates and log output computed using the values provided in table 5. Per capita GDP is assumed to be 8 times larger in the US at time zero.

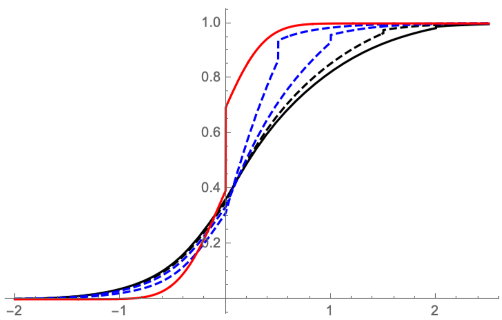

Figure 8: Simulated CDF of log of income, normalized by gt. The figure plots the distribution at time 0, 10, 20, 30, and 40 years after the reforms started. Lines to the right are later periods compared to the left. The CDF displays jumps that stem from the initial share of households that start growing fast. This mass point disappears over time as more and more agents are pulled out of the high growth regime.

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 3: 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 ($\frac{\lambda}{g_h - g}$) and top income inequality. For finite aggregate convergence we also need $\lambda > g_h - g$. 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 $\lambda \leq g_h - g$. A smaller and smaller share of fast-growing agents would eventually account for 100% of GDP, leading to a long-run growth rate of g_h. 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.

5.3 Capital Outflows

Finally, I compute the capital flows of the economy along the transition path. To do so, I need to first solve for the transitional consumption and asset accumulation dynamics in the high-growth regime. This is simple, however, since the problem of the household in the high growth regime always looks the same (up to some linear scaling factor), no matter if a household enters the urban economy at time zero or a thousand years in. Intuitively, households always go through the same dynamics, albeit at different starting wages w_t^u. All choice variables are then simply scaled by income but otherwise unchanged.

Figure 9 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 consumption in the fast-growth regime. When computing aggregate savings, then, I use the consumption function of figure 9 together with knowledge about how much time households spent in the high-growth regime, to get the right aggregate asset position.

Going back to the empirical exercise in section 2, 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. High urban asset-to-income ratios, then, are indicative of high human capital risk in modern production. The parameterization already reveals that the growth miracle is going to be accompanied by capital outflows, since households accumulate assets in the high growth regime. This is by no means guaranteed, and for more "even" growth miracles this would not be the case.

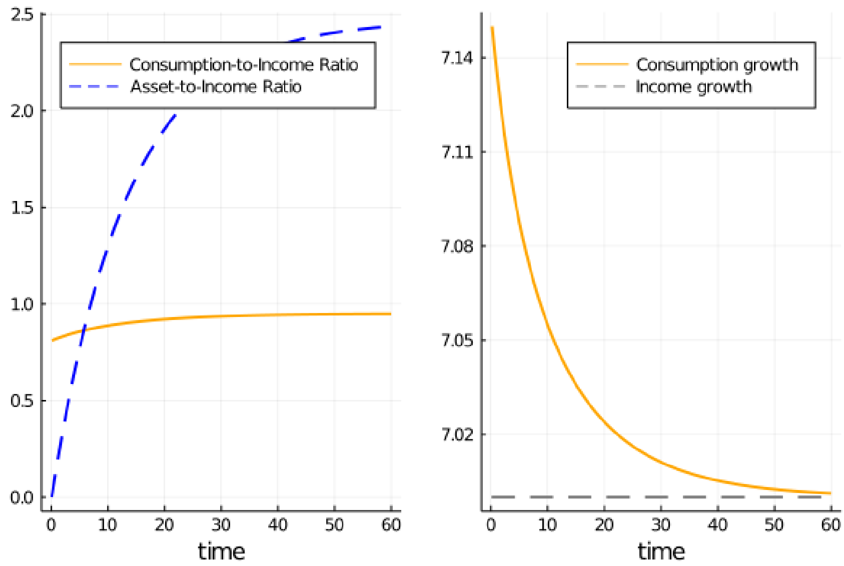

Figure 9: Transitional dynamics in high-growth regime using parameters from table 5. Time here is measured as time passed since the household switched sectors.

In a final step, I use the distribution of income together with the solution to the optimal consumption path in the high-growth regime to back out what the aggregate saving rate along the transition path. Recall that this aggregate saving rate is comparable to the current account since there is no capital in the model. Putting the pieces together we get a trajectory for the aggregate saving rate displayed in figure 10.

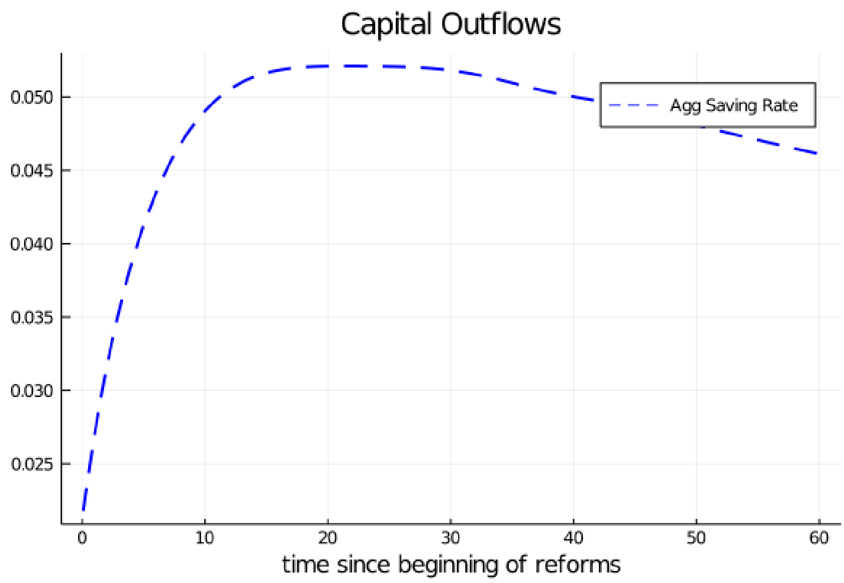

Figure 10: Aggregate saving rate over time. Since there is no capital in the model this coincides with the Current Account in the small open economy.

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. Two hundred years in, the saving rate stabilizes around 3.5%.

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 11 shows the declining share of agriculture. The hump-shaped saving rate can only emerge as a compositional effect. That is, there needs to be an increasing share of precautionary savers relative to total output for some time. This is precisely what happens as figure 11 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, χ and λ as these govern inflow and outflow into $M_{t,0}$ and $M_{0,1}$.

Figure 10 shows that this model can solve the capital flow puzzle for the right parameter values. The heterogeneity in the growth process and rural-urban differences in human capital risk are key deviations from the representative agent neoclassical model that make this feasible.

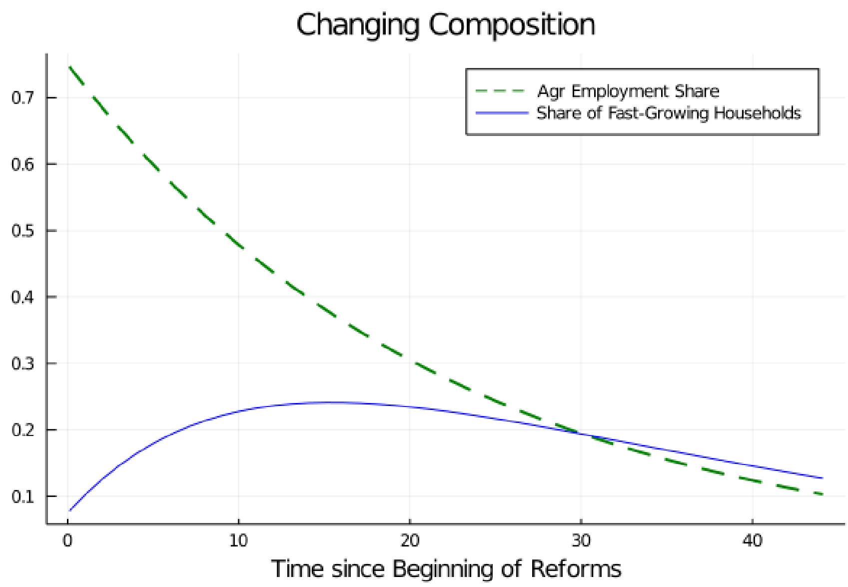

Figure 11: Changing share of population in agriculture, and in the high-growth regime.

6 Conclusion

This paper argues that rapid structural change from traditional agriculture to modern urban production fundamentally alters household saving behavior during episodes of catch-up growth. When households move into the urban sector, they face substantial ex ante uncertainty about their long-run income prospects despite high average growth. This risk weakens the consumption-smoothing motive implied by the permanent income hypothesis and generates strong precautionary saving pressure, leading to capital outflows even in fast-growing economies.

I document empirical patterns consistent with this mechanism using household-level data from China. Urban households hold substantially higher asset-to-income ratios than rural households, particularly in safe assets, and recent rural–urban migrants accumulate assets at faster rates than urban incumbents. Combined with the scale and speed of structural change, these differences help explain the rise and subsequent decline of aggregate saving rates observed during growth miracles.

To interpret these facts, I develop a tractable model of the rural–urban transition in which households experience fast but uneven income growth after migration. The key departure from standard frameworks is that convergence growth itself is risky and unevenly distributed across households. Together with a large rural–urban wage gap, this substantially weakens the consumption-smoothing force induced by fast aggregate growth that otherwise dominates standard incomplete-markets models. More broadly, incorporating uneven growth and sectoral reallocation appears to be a promising avenue for improving our understanding of aggregate savings and capital flows in fast-growing economies. The stylized model presented here is a first step in that direction.

References

Aguiar, M. and M. Amador (2011). Growth in the shadow of expropriation. *The Quarterly Journal of Economics 126*(2), 651–697.

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

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.

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.

Bai, Chong-En, Q. Z. (2009). Who is squeezing out household income? an analysis of the national income distribution in china. *Social Sciences in China*.

Barro, R. J. and X. I. Sala-i-Martin (2003). *Economic Growth* (2nd ed.). Cambridge, MA: The MIT Press.

Baseler, T. (2023). Hidden income and the perceived returns to migration. *American Economic Journal: Applied Economics 15*(4), 321–352.

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.

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.

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. and O. Jeanne (2009). A tractable model of precautionary reserves, net foreign assets, or sovereign wealth funds. Technical report, National Bureau of Economic Research.

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.

Carroll, C. D. and L. H. Summers (1991). Consumption growth parallels income growth: some new evidence. In *National saving and economic performance*, pp. 305–348. University of Chicago Press.

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., S. Guibaud, and K. Jin (2015). Credit constraints and growth in a global economy. *American Economic Review 105*(9), 2838–2881.

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

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.

De Nardi, M. (2015). Quantitative models of wealth inequality: A survey. Technical report, National Bureau of Economic Research.

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

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.

Dynan, K. E., J. Skinner, and S. P. Zeldes (2004). Do the rich save more? *Journal of political economy 112*(2), 397–444.

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

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.

Gollin, D., D. Lagakos, and M. E. Waugh (2013). 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.

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.

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

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

Imrohoroğlu, A. (1989). Cost of business cycles with indivisibilities and liquidity constraints. *Journal of Political economy 97*(6), 1364–1383.

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.

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., A. M. Mobarak, and M. E. Waugh (2023). The welfare effects of encouraging rural–urban migration. *Econometrica 91*(3), 803–837.

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

Lewis, W. A. (1954). Economic development with unlimited supply of labour. *The Manchester School 22*(2), 139–199.

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.

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

Meghir, C., A. M. Mobarak, C. Mommaerts, and M. Morten (2022). Migration and informal insurance: Evidence from a randomized controlled trial and a structural model. *The Review of Economic Studies 89*(1), 452–480.

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

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.

Munshi, K. and M. Rosenzweig (2016). Networks and misallocation: Insurance, migration, and the rural-urban wage gap. *American Economic Review 106*(01), 46–98.

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.

Porcher, C., E. Morales, and T. Fujiwara (2024). Measuring information frictions in migration decisions: A revealed-preference approach. Technical report, National Bureau of Economic Research.

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.

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.

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.

Straub, L. (2019). Consumption, savings, and the distribution of permanent income. *Unpublished manuscript, Harvard University*.

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.

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.

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.

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 (??):

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.[35] His insight generalizes to the framework at hand as well. Let $X_s := \frac{c_s}{y_s}$ denote the consumption-to-income ratio, and let $Z_s := \frac{a_s}{y_s}$ denote the asset-to-income ratio. This leads to the following differential equations

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

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

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

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

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

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 $\mathbb{E}_\varphi\left[\left(\frac{1}{\varphi}\right)^\eta\right] = \exp\left(\frac{\sigma^2}{2}\eta\left[1+\eta\right]\right)$. For $\eta = 2$, any σ^2 above .37 will suffice. For $\eta = 4$, any σ^2 above 0.16 is sufficient to dominate the consumption smoothing motive.

However, the argument extends beyond the log-normal case. What is needed is the notion of

[35] 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 $\{X : X < B\} = \{X : X \leq -B\} \cup \{X : -B <$

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

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

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 12 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.

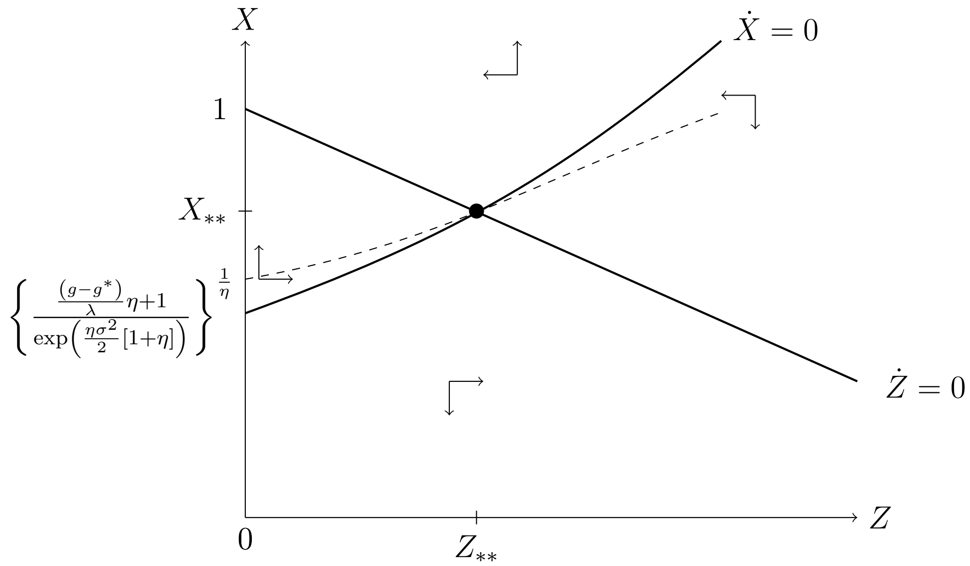

Figure 12: Pseudo steady state analysis of asset-to-income and consumption-to-income ratio in high-growth regime

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_m^i, 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:

$$c(t_m, t) = c_0(t - t_m)\, exp(g^* t_m)$$

$$a(t_m, t) = a_0(t - t_m)\, exp(g^* t_m)$$

$$\begin{aligned} y(t_m, t) &= exp(g[t - t_m])\, exp(g^* t_m) \\ &= y_0(t - t_m)\, exp(g^* t_m) \end{aligned}$$

Proof:

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

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.[36] 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 (**??**) 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.

[36] 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 (**??**) 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.[37]

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

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

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

[38] 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

$$log(w_t^r) \stackrel{!}{=} \rho V_t - \dot{V}_t.$$

Given the conjectured equilibrium, I obtain a closed form solution for the value function as follows.

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

Now I can differentiate this expression with respect to t, and plug it back into the arbitrage condition that keeps agents on the country side indifferent between staying and moving, hence

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

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 $\frac{1}{\lambda+\rho}\left(g - g^*\right)$. Note that without the wedge, assuming that convergence growth and precautionary savings cancel, it can be shown that the urban wage at time t_m would be smaller than the rural wage at t_m. This happens because rural households need to be compensated for the lack of high-growth opportunity in equilibrium. Given log utility, this force dominates the risk adjustment, that pushed down the rural wage. This is why we need $\tau > 1$, i.e. there needs to be an additional wedge to migration.

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 2. 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.

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

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

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

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

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

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

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

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}.[39] Then,

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

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

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 $f_X\left(x\right)=\frac{\frac{g^{*}}{1-\alpha}L_{0}^{r}exp\left(-\frac{g^{*}}{1-\alpha}t\right)exp\left(-\left(\lambda-\frac{g^{*}}{1-\alpha}\right)x\right)}{M_{0,t}}$ yields

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

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

$$
\begin{aligned}
P\left(y_{t}\leq k|T_{i}<t\right) &= P\left(\log y_{t}\leq\log k|T_{i}<t\right)\\
&= P\left(\left(g-g^{*}\right)\left(T_{i}-t_{m_{i}}\right)\leq\log k|T_{i}<t\right)\\
&= \mathbb{E}_{t_{m}}P\left(\left(T_{i}-t_{m_{i}}\right)\leq\frac{\log k}{\left(g-g^{*}\right)}|T_{i}<t,t_{m}\right)
\end{aligned}
$$

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

$$
f_{t_{m}|T<t}\left(k\right)=\frac{1}{M_{1,t}}\left(\frac{g^{*}L_{0}^{r}}{1-\alpha}\right)\left[\exp\left(-\frac{g^{*}}{1-\alpha}k\right)-\exp\left(\frac{\lambda\left(1-\alpha\right)-g^{*}}{1-\alpha}k\right)\exp\left(-\lambda t\right)\right].
$$

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

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

$$+ \frac{\left(1-\exp\left(-\lambda t\right)\right)M_{0,0}}{M_{1,t}} \frac{1-\exp\left(-\lambda\left[\frac{\log k}{(g-g^*)}\right]\right)}{1-\exp\left(-\lambda t\right)}$$

$$+ F_1\left(t\right) - F_1\left(t - \frac{\log k}{g-g^*}\right) + \frac{M_{1,0}}{M_{1,t}}$$

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

$$+ \frac{M_{0,0}}{M_{1,t}}\left(1-\exp\left(-\frac{\lambda}{(g-g^*)}\log k\right)\right)$$

$$+ F_1\left(t\right) - F_1\left(t - \frac{\log k}{g-g^*}\right) + \frac{M_{1,0}}{M_{1,t}}$$

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

$$+ \frac{M_{0,0}}{M_{1,t}}\left(1-\exp\left(-\frac{\lambda}{(g-g^*)}\log k\right)\right)$$

$$+ F_1\left(t\right) - F_1\left(t - \frac{\log k}{g-g^*}\right) + \frac{M_{1,0}}{M_{1,t}}$$

$$= \frac{L_0^r}{M_{1,t}}\left\{1-\exp\left(-\lambda\left[\frac{\log k}{(g-g^*)}\right]\right)\right\}\left\{1-\exp\left(-\frac{g^*}{1-\alpha}\left(t-\frac{\log k}{(g-g^*)}\right)\right)\right\}$$

$$+ \frac{M_{0,0}}{M_{1,t}}\left(1-\exp\left(-\frac{\lambda}{(g-g^*)}\log k\right)\right)$$

$$+ F_1\left(t\right) - F_1\left(t - \frac{\log k}{g-g^*}\right) + \frac{M_{1,0}}{M_{1,t}}$$

$$= \frac{L_0^r}{M_{1,t}}\left\{1-\exp\left(-\frac{\lambda}{(g-g^*)}\log k\right)\right\}\left\{1-\exp\left(-\frac{g^*}{1-\alpha}t\right)\exp\left(\frac{g^*}{(g-g^*)\left(1-\alpha\right)}\log k\right)\right\}$$

$$+ \frac{M_{0,0}}{M_{1,t}}\left(1-\exp\left(-\frac{\lambda}{(g-g^*)}\log k\right)\right)$$

$$+ F_1\left(t\right) - F_1\left(t - \frac{\log k}{g-g^*}\right) + \frac{M_{1,0}}{M_{1,t}}$$

$$= \frac{L_0^r}{M_{1,t}}\left\{\exp\left(-\frac{g^*}{1-\alpha}t\right)\exp\left(\frac{g^*-\lambda\left(1-\alpha\right)}{(g-g^*)\left(1-\alpha\right)}\log k\right) - \exp\left(-\frac{g^*}{1-\alpha}t\right)\exp\left(\frac{g^*}{(g-g^*)\left(1-\alpha\right)}\log k\right)\right\}$$

$$+ \frac{L_0^r + M_{0,0}}{M_{1,t}}\left[1-\exp\left(-\frac{\lambda}{(g-g^*)}\log k\right)\right]$$

$$+ F_1\left(t\right) - F_1\left(t - \frac{\log k}{g-g^*}\right) + \frac{M_{1,0}}{M_{1,t}}$$

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.[40] 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.[41] 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.

[40] 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.

[41] 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 Empirical Appendix

8.1 CFPS

The CFPS is a household panel dataset that comprises detailed information on family structure, income, expenditure, assets, and demographics. The survey was launched in 2010 by the Institute of Social Science Survey (ISSS) of Peking University, China.[42] 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.

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).

[42]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.

Descriptive Statistics CFPS

Table 4 and 5 display mean and standard deviations of the raw sample, i.e. before I restrict on age, labor market attachment, and household splits, and for the selected sample, respectively.

Table 4: Descriptive Statistic CFPS

	Rural		Urban	
	mean	sd	mean	sd
Broad labor income	38915.2	85560.3	64739.3	81798.0
Years of schooling	6.1	4.3	9.0	4.7
Age	50.8	13.8	49.3	15.1
Male share	0.7	0.4	0.7	0.5
Household size	4.0	1.9	3.4	1.7
Total net assets	236532.5	835413.0	713546.6	1582119.2
Safe assets	21651.6	72873.4	74205.3	249065.8
Observations	25556		24179	

Note: Mean and standard deviation for the raw sample, i.e. before selection.

Table 5: Descriptive Statistic CFPS – Selected Sample

	Rural		Urban	
	mean	sd	mean	sd
Broad labor income	38660.9	42847.5	65632.3	86969.4
Years of schooling	6.2	4.1	9.4	4.4
Age	46.3	9.1	45.2	9.1
Male share	0.7	0.4	0.7	0.5
Household size	4.3	1.7	3.7	1.5
Total net assets	242997.9	698359.2	711700.6	1361748.5
Safe assets	21988.5	68107.3	79621.3	275307.8
Observations	12132		10468	

Note: Mean and standard deviation for the selected sample.

Additional Quantile Regression: total assets and sample weights

Table 6: Rural-Urban Median Differences for Safe Assets

	(1)	(2)	(3)
	safe_asset_income_ratio	safe_asset_income_ratio	safe_asset_income_ratio
urban dummy	0.208^{***}	0.191^{***}	0.128^{***}
	(0.0296)	(0.0285)	(0.0288)
_cons	0.330^{***}	0.806^{**}	0.357
	(0.0155)	(0.394)	(0.247)
age	No	Yes	Yes
sex and hh size	No	No	Yes
education	No	No	Yes
Observations	5384	5384	5384

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. I apply cross-sectional weights for 2012 for this regression.

Table 7: Rural-Urban Median Differences for Total Assets

	(1) tot_asset_income_ratio	(2) tot_asset_income_ratio	(3) tot_asset_income_ratio
urban dummy	2.780***	2.849***	2.494***
	(0.204)	(0.206)	(0.232)
_cons	5.843***	6.874***	7.635***
	(0.102)	(2.498)	(2.370)
age	No	Yes	Yes
sex and hh size	No	No	Yes
education	No	No	Yes
Observations	5949	5949	5937

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.

Table 8: Rural-Urban Median Differences for Total Assets

	(1) tot_asset_income_ratio	(2) tot_asset_income_ratio	(3) tot_asset_income_ratio
urban dummy	2.151***	2.154***	1.780***
	(0.267)	(0.250)	(0.264)
_cons	6.085***	14.65***	14.41***
	(0.148)	(3.566)	(3.585)
age	No	Yes	Yes
sex and hh size	No	No	Yes
education	No	No	Yes
Observations	5384	5384	5384

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. I apply cross-sectional weights for 2012 for this regression.