Missing Land Markets, Misallocation, Insurance, and Redistribution*

Joao Rodrigues†

January 1, 2022

CLICK HERE FOR LATEST VERSION

Abstract

Does replacing *use it or lose it* land institutions with land markets trigger productivity growth and reallocation of labor towards non-agriculture? How does uninsurable risk shape the benefits of these reforms? I use rich panel microdata from Malawi and document higher income risk in non-agriculture. I then build a model where households face risk in non-agriculture and maintain rights to agricultural land via the principle of *use it or lose it*. When markets are incomplete but land is easily accessible, households can partially self-insure by returning to agriculture when non-agricultural opportunities fail. As land becomes less accessible, productive farmers disproportionately forgo profitable opportunities in non-agriculture to limit their exposure to uninsurable risk. I use moments from the Malawi microdata and fixed effect estimates of income to estimate key model parameters via indirect inference. With only 1.3% of landless households accessing land each year in Malawi, the prospect of long landless spells pushes productive farmers to stay in agriculture. I use the model in a policy experiment where land is privatized and landholders at the time of the reform earn income on the land. While the policy yields an increase in aggregate output of 28%, patterns of entry and exit attenuate the productivity gains in agriculture by 16%. Productivity grows by 22% in agriculture and 15% in non-agriculture. Although landless households benefit from the higher wage in agriculture, their welfare gains are on average 10 times lower than those of landowners.

JEL Classification Codes: O11, E02, H11

Keywords: communal land, property rights, land misallocation, insurance and redistribution

---

*I am grateful to the support of my advisors Ellen McGrattan, Todd Schoellman, Alisdair McKay, and Mike Waugh. The views expressed herein are mine and not necessarily those of the Federal Reserve Bank of Minneapolis or the Federal Reserve System

†University of Minnesota and Federal Reserve Bank of Minneapolis E-mail: fons0029@umn.edu

1 Introduction

Increasing evidence suggests that missing land markets generate large productivity costs by preventing the reallocation of land to more productive farmers. Under communal land systems in developing countries, farmers have few opportunities to grow their operations through market transactions. In Africa, for instance, it is estimated that 90% of land on the continent is under customary designation (Deininger et al., 2003). Under customary rules, local traditional authorities restrict private transfers and reallocate unused land to landless farmers. Studies using microdata from China and Malawi argue that these practices disproportionately impact high-ability farmers – driving them either to operate small farms or to leave agriculture altogether (Adamopoulos, Brandt, Leight and Restuccia, 2021; Chen, 2017). They argue that land markets would reallocate labor across sectors and drive productivity gains in agriculture in the order of 3 to 4 fold.

Despite these productivity costs, rural institutions have long acted as a source of informal insurance in developing countries (Townsend, 1994; Udry, 1994; Morten, 2019). Under communal land systems, land rights operate under the principle of use it or lose it as local leaders reclaim unused land and reallocates them to landless households. Aware of this, households move to villages where land is available in times of economic stress. In low-income countries where the formal government plays a minor role in social insurance or redistribution, this provides an alternative avenue for poor households to access productive resources and smooth consumption. This suggests a tradeoff whereby misallocation emerges in part while providing potentially valuable insurance and redistribution.

This paper quantitatively explores this tradeoff and their implications for policies that promote more active land markets in low-income countries. I develop and quantify a model where land misallocation emerges from limits on private transfers. However, uninsurable risk interacts with land availability to select the best farmers to stay in agriculture. I use the model to introduce a market for land and show that: despite the growth in aggregate productivity, the selection effect attenuates the productivity gains in agriculture, and those on the land at the time of the reform benefit by a factor of 10 times more than the landless.

To conduct my analysis, I develop a model where households differ in their permanent agricultural and non-agricultural productivity as in Roy (1951) and face idiosyncratic uninsurable risk in non-agriculture as in Aiyagari (1994). Households access communal land by working in agriculture, and earning a wage as farm workers while landless. If land becomes available, the local leader assigns the available plot to a landless farm worker who then becomes a smallholder. In order to keep the land, smallholders have to operate a farm continuously. If the smallholder chooses to leave agriculture and earn income in non-agriculture, local leaders reclaim the land and reallocate it to a landless farm worker.

The main insight of my paper is that the degree of uninsurable risk in non-agriculture makes leaving agriculture more costly when land is more scarce. As land becomes more scarce, fewer plots become available for reallocation, and households can take longer to access land. It adds a new risk component to leaving agriculture: if opportunities in non-agriculture fail, regaining access to land may entail long spells in low-paid agricultural work. Since farmers get to keep their output, productive farmers earn more on their land. Landless households performing farm work earn the market wage independent of their agricultural productivity. Therefore, the costs of uninsurable risk are higher for productive farmers who have more to lose by leaving agriculture. As land access deteriorates, these costs rise further and push productive farmers disproportionately to stay in agriculture.

I use panel microdata from Malawi to discipline the parameters of my model. I exploit its longitudinal nature to estimate a permanent fixed effect measure of household productivity in each sector for each household. I use the estimated fixed effect and patterns of occupational mobility to discipline the dispersion and interdependence of permanent sectoral productivities. I extract the idiosyncratic transitory component of income and use it to discipline the degree of risk in non-agriculture. At last, I use the rate in which households access land through local leaders to discipline the degree of land availability. The model matches well patterns of employment, risk and mobility across sectors.

In the main quantitative exercise, I conduct a policy experiment where communal land is converted into a rental market. I do this by assigning property rights to households on the land at the time of the reform. After the reform, local leaders no longer play a role in allocating agricultural land. The new landowners can now earn an income on their land and allocate their labor to the sector where income is higher. The land gets reallocated as households rent land in and out until a rental price clears markets.

While the policy raises aggregate output, private markets induce entry and exit patterns that attenuate the productivity gains in agriculture. Aggregate output grows by 28%, and productivity rises by 22% in agriculture and 15% in non-agriculture. The increase in agricultural productivity results from improvements in land allocation across farmers. However, only 1.3% of landless households access land via local leaders in any given year in Malawi. This implies a degree of land scarcity that keeps productive farmers from leaving agriculture. Since markets also raise the labor return in agriculture, agricultural productivity gains are attenuated by too many productive farmers leaving agriculture and unproductive ones entering after the reform.

The pattern of selection in my model induced by markets play a different role in shaping the effects of communal land compared to the previous literature. Land scarcity interacts with uninsurable risk in non-agriculture to induce productive farmers to remain in agriculture. By contrast, previous work abstracting from incomplete markets find that farm size restrictions tend to lead productive farmers to exit the agricultural sector. For instance, in Adamopoulos, Brandt, Leight and Restuccia (2021) and Chen (2017), selection accounts for 60% to 75% of the productivity gains in agriculture from removing farm-level distortions. In other words, selection amplifies the effects of farm-level distortions by inducing more productive farmers to enter agriculture and unproductive ones to leave. In a static framework, households choose the sector that pays more, and face no dynamic tradeoffs. In my paper, restrictions on private transfers also function as a tax on productive farmers and a subsidy to unproductive ones. However, productive farmers are willing to give up the most temporary income in non-agriculture to prevent their exposure to uninsurable risk – an effect that is exacerbated by land scarcity. Private markets allow them to realize potentially large income gains in non-agriculture, causing a selection effect that is absent in static models.

The distributional consequences of the reform are sizable and leave landless households benefitting very little relative to landholders. The most significant welfare gains fall to low farm and low non-farm productivity on the land at the time of the reform. The reason for this is that if you are very unproductive, most of your farm income came from the wage in agriculture before the reform. However, after the reform, you can now earn rents on the land by renting it out. If you have low productivity and need very little of the land, now land income is a substantial share of household income. And since land is better allocated in agriculture, a more productive farm sector means higher wage and land rents in agriculture. Landless households, many of them in agriculture (9% in the data) only gain from the increased wage. In terms of consumption equivalent gains, this turns out to be on average 10 times less that the same measure for landholders.

The key departure in my paper from the literature on land misallocation come from the addition of uninsurable risk into an otherwise standard model of selection and farm-level distortions. My model includes the standard selection framework from Adamopoulos, Brandt, Leight and Restuccia (2021), but also adds land access rules that are common in communal land systems. My paper highlights an important interaction between the degree of land availability and uninsurable risk. So far, virtually all papers related to land misallocation and communal land abstract from incomplete insurance, and as a result, can only speak to welfare costs through efficiency gains (Adamopoulos, Brandt, Leight and Restuccia, 2021; Restuccia and Santaeulalia-Llopis, 2017; Chen, 2017; Gottlieb and Grobovšek, 2019; Chen, Restuccia and Santaeulàlia-Llopis, 2021).

I join a growing literature in macroeconomics trying to understand the implications of uninsurable risk to agricultural productivity in low-income countries (Donovan, 2020; Brooks and Donovan, 2020; Lagakos et al., 2018). Like these papers, the core of my model is a standard incomplete markets model in the tradition of Imrohoroğlu (1989), Bewley (1986), Huggett (1993), and Aiyagari (1994). In my paper, communal land access rules interact with incomplete markets, making households less willing to leave agriculture. As land becomes more scarce, risk in non-agriculture becomes more relevant as returning to communal land may entail long spells as a landless farm worker. This interaction determines the degree to which households are misallocated within and between sectors.

2 Context

Communal land systems are prevalent in the developing world through customary arrangements of access to land. It covers an estimated 90% of land in Africa, some indigenous regions of Latin America, as well as some countries in Asia (Deininger et al., 2003). Under these systems, rights on land for individual cultivation are secured through membership in the lineage that cleared the land. Although property rights on land fall outside the purview of formal law, land rights provided by these systems are often secure, long-term, and in most cases inheritable (Feder and Noronha, 1987). Once the family has control over the land, they pass it down to their children as they marry and form their own families.

In the quantitative exercise, I focus on the particular case of Malawi. Malawi is typical of low-income countries in that it is primarily rural and agricultural. The particular advantage of Malawi is that it has detailed panel microdata that speaks to households' income from farm and non-farm activities and how frequently they switch between them. Further, it includes information on how households acquire the land they use in agriculture.

Malawi is one of the poorest and most densely populated countries in Africa. About 82% of the population live in rural areas, only down from 91% in 1977 (Peters and Kambewa, 2007). This makes Malawi particularly relevant as land availability issues may be increasingly relevant in the absence of markets. Further, its customary tenure system shares many similarities with other countries in Africa; hence insights on their welfare costs and benefits may have broader implications.

Formally, there are three types of land ownership delineated in the Malawian constitution: customary, public, and leasehold. Under leasehold, the government grants long-term leases and collects royalties from private individuals who run what is commonly known as estate farms. Leasehold titles were originally granted to Malawi nationals from land previously owned by colonial powers. Since independence, the government issues new titles by converting customary land into leasehold land. This, however happens infrequently as it requires permission from local leadership, who is expected to only grant such permissions for unused land that the village no longer needs. In Malawi, estate farms are the primary cultivators of the country's cash crops: tobacco and tea.

I focus on customary land, which represents the primary land tenure system and where most of the population in Malawi lives. Customary land is an official designation of indigenous land that functions much like other communal land systems. The village chief (or headman) manages customary land by upholding lineage-based inheritance rules and allocating unused land to local landless households. White (1989) sums up best the role of village chiefs regarding land in his book on village life in Malawi.

> The land belonged to the village, not to the lineage. It was the headman who was responsible for its distribution. Land which had once been allocated could not be taken back by the headman unless it had been left uncultivated, but the headman remained responsible for finding land for fresh settlers and for arbitrating boundary disputes. (p. 164)

This description illustrates the incentives facing poor households in Malawi, where few sources of social safety net exist from the formal government. In a country where nearly all rural land is managed in this way, moving to villages where land is available is typical for overcoming adverse shocks.

3 Empirical Evidence

My model deals with the interaction of sector-specific productivity, sectoral risk, and occupational choice. I use the longetudinal nature of the data to separate the permanent from risk components of productivity across sectors. This section shows some key results about the empirical counterparts of these objects that motivate my analysis and discipline central parameters. There are two main takeaways:

- Permanent income dispersion is larger in non-agriculture
- Non-agriculture is riskier than agriculture

3.1 Data

I use data from Malawi's Integrated Household Survey (IHS), starting from 2010, including the World Bank's Integrated Surveys of Agriculture (LSMS-ISA). ISA added detailed information on agricultural production to Malawi's primary household survey. I use the cross-sectional data from 2010-2011 the panel subsamples from 2010-2011 to 2012-2013.

The cross-sectional sample is nationally representative and contain detailed information and all sources of household income and production. The 2010-2011 sample includes 12,271 households (56,397 individuals). In the data, I observe income generating activities of each household member: agricultural production, labor income, capital income, business income, assets, and transfers.

The panel subsample has the same information as the cross-section and includes both original and split-off households. In the panel data, the attrition rate was low. Only about 6% of eligible individuals from the baseline could not be traced back for follow-up in 2013. The 2012-2013 panel wave grew to 4,000 households as some were lost due to tracking issues or death, while others were created by splitting off from baseline households.

3.2 Income

In agriculture, I construct value added at the crop level by adding up revenues, the household's consumption value of unsold output, and subtracting costs. I follow Gollin, Lagakos and Waugh (2014) and measure value added from all activities in agriculture, including non-permanent crops, permanent crops, livestock, and livestock products. I aggregate all these measures first across seasons and then to the household level. To measure the value of unsold products, I use local prices whenever available. If the most local price is not available, I broaden the region category and use prices at a broader regional level (De Magalhães and Santaeulàlia-Llopis, 2018).

I compute non-farm income as the sum of formal labor income, informal labor income, and non-farm business income. The household module of the survey contains detailed information on enterprises run by the household, including their revenues, input costs, and labor hours. Formal labor income is recorded for primary and secondary occupations, while informal labor (ganyu) income is reported separately. Rural workers earning primarily ganyu and having no land are the poorest among Malawi's society. Total household income can be written as the sum of three categories of income: $y_i = y_{a,i} + y_{b,i} + y_{l,i}$, where $y_{a,i}$ is the value-added measure of agricultural output, $y_{b,i}$ is business income, and $y_{l,i}$ is labor income.

The next goal is to separate income by sector. For farm value added, business income and formal labor income, the sector is either clear or specified in the survey. For ganyu, it is not specified in which sector the income is earned and is a large share of labor income for the poor. I classify any ganyu earned in rural areas as agricultural labor income and non-agricultural labor income if earned in urban areas.

In addition to sectoral choice, the model also deals with households that differ in their access to land. Some agricultural households work as smallholders while others work as farm laborers. I classify as a smallholder those who earn at least 50% of their annual income from agricultural production. For other households, I label them as farm workers those earning more than 50% of their income from informal labor in rural areas. All other households are non-agricultural households. Table 1 summarizes income differences across these groups. Appendix C has more detail on the data construction of incomes in different sectors.

Table 1: Summary of Household Income, IHS 2010-11

	average	s.d.	p20	p40	p60	p80	p90
Farm worker	26.8	37.7	4.2	12.3	21.9	39.0	56.0
Smallholder	51.1	126.1	15.4	27.2	41.7	68.1	97.6
Non-farm	161.2	484.8	26.8	48.6	84.0	157.6	286.3

Notes: All income is measured in thousands of March (2010) Kwachas.

Table 1 provides initial evidence of the degree of risk and heterogeneity in non-agriculture relative to agriculture. The share of households that are classified as farm owners is 63.3%, as farm workers is 9.3% and as non-farm workers is 27.3%. In 2010, Malawi's GDP per capita was 72 thousand of 2010 Kwachas. That is just below the 60th percentile of the non-farm income, above the 80th percentile of the income of farm owners and above the 90th percentile of farm workers.

3.3 Productivity and Risk

My model deals with occupational choice due to productivity differences and risk in non-agriculture. In this section, I use the longitudinal nature of the data to separate permanent versus transitory sources of productivity. To this end, I use the classification of households as smallholders, farm workers, and non-farm workers as described above. To measure agricultural productivity, I use income from agricultural production $y_{a,i}$. To measure non-agricultural productivity, I only keep the business, labor income, and informal labor income in the city $y_{b,i} + y_{l,i,urban}$ as income in non-agriculture. I can then exploit the panel dimension of the data to estimate the permanent and risk components of income in the two sectors.

To get the permanent component of output, I follow Adamopoulos, Brandt, Leight and Restuccia (2021) and estimate the following equation

$$y_{i,v,t} = \mu_t + \mu_i + \varepsilon_{i,v,t} \tag{1}$$

where μ_t is the year fixed effect that captures time varying shocks to productivity that are common to all farmers; μ_i is the farm specific component that does not vary over time; and $\varepsilon_{i,v,t}$ captures idiosyncratic shocks specific to the household in a given year. In a second step, I follow their procedure to remove village specific effects that may not differ over time but differ across individuals. I remove these village specific effects by regressing the household fixed effect on village dummies and extracting the residual,

$$\mu_i = \mu_v + \eta_i \tag{2}$$

where η_i is a fixed farm component that captures a farm's permanent output that is constant across years and purged of village level factors. I then gather the residuals from equation 2 and they become the permanent measures of productivity (I do this for income in each sector). My measure of risk comes from the idiosyncratic residual component of income $\varepsilon_{i,v,t}$ in equation 1.

Table 2: Productivity Variance

	Permanent	Transitory
Farm	0.71	1.20
Non-farm	1.12	2.19

Notes: Dispersion is measured by the variance.

The analysis above uncovers large differences in terms of heterogeneity and risk in agriculture versus non-agriculture. The high relative risk in non-agriculture motivates to model risk in non-agriculture.

3.4 Productivity and Land Access

I now investigate the extent to which households gain access to agricultural land through the village chief and its relationship to farm productivity. I first report patterns of agricultural land access in the cross section. I then use the panel subsample to tie how households acquired their land to the estimated productivity fixed effect.

Table 3: Land Tenure in Malawi

	Family	Village chief	Rent/Own
Agricultural plots (%)	67	17	11

LSMS (2010/2011)

Table 3 displays the patterns of access to agricultural land in Malawi. Households are asked how each one of the plots they use in agriculture were acquired. Each column represents the percentage of agricultural plots acquired by mode of acquisition. Plots acquired either through allocation by traditional authorities or by family add up about 90% of agricultural plots. These represent the households living under customary land. The first column illustrates the dominance of the family-based system of land allocation, where land that was once allocated by traditional authorities continually gets passed down through the lineage. The chief (traditional authority) still dominates a large share of allocated land (17%). The rest is composed of short-term rentals and long-term leases. Under the rent/own category, I include both short-term rentals and owned land (or on leasehold). Short-term rentals are generally illegal by customary practices, but they happen anyway. Of the 11% in this category, 3% is reported to be on leasehold. These are regulated leases managed by the formal government and discussed in section 2.

I now turn to the relationship between farm productivity and land access through the village chief. I first separate households by whether their current land holdings come entirely from the village chief (306 households). I use land size divided by household size and the estimated farm-level productivity fixed effect. Table 4 displays the difference in levels for land size and productivity when households acquire their plot through the chief versus otherwise. It shows that households with chief allocated land (likely to be veterans in the village) have smaller farms and lower productivity.

Table 4: Land Access and Productivity

	Land size	Productivity
Allocated by chief	0.36	1.31
Non-allocated	0.47	1.42

To further investigate the role of the chief, I regress the log of land size on the log of the estimated farm productivity fixed effect and a dummy for whether the households' land is acquired entirely from the chief versus those whose land has no chief allocated land. The coefficient on productivity is 0.48, while the coefficient on the dummy is 0.12. While the coefficient on productivity significant, the one on the dummy is only significant at the 0.05 level. To put this in context, the correlation of farm size and productivity in U.S. data is 0.90 (Adamopoulos and Restuccia, 2014). By selecting households whose land is primarily allocated by the chief, they are likely to have been recently allocated. This is because the share of chief allocated land is likely to decrease as households acquire more plots through the lineage-based system. Figures 12 in the appendix C displays the distribution of land sizes and the correlations estimated in a scatter plot.

3.5 Sector Switching

I now investigate the extent to which households switch across sectors and the productivity patterns. Since I classify the occupation of the household based on the primary income earned, not all switchers into the farm sector are new land allocations.

In table 5 I take households in each occupation in 2010-2011 (by row) and see where they end up in 2012-2013. I then take the percentage by occupation of households in 2010-2011 that ends up in each of the occupations in 2012-2013.

Table 6 shows the level and dispersion of farm productivity for different occupations in 2012-2013 for those who were smallholders in 2010-2011. It shows the mean of farm productivity for those who stayed in agriculture is similar to those who left agriculture.

Table 5: Occupational Mobility

	Farm workers	Smallholders	Non-farm workers
Farm workers	16	57	26
Smallholders	4	71	24
Non-farm workers	2	30	67

Notes: Percentages add up to 100 by row.

However, the variance of farm productivity for those who left agriculture was much larger.

Table 6: Farm Productivity of 2010-2011 Smallholders

	Mean	Variance
Farm in 2012-2013	0.016	0.54
Non-farm in 2012-2013	0.015	0.80

4 Dynamic Model of Communal Land

In this section, I formalize the details of the model of communal land. There are two sectors: agriculture and non-agriculture. Households choose which sector to work in and face income risk in non-agriculture. Landless households can gain access to communal land by working in agriculture. If they gain use rights on land, they combine land and labor to produce agricultural output. At any time, they can run a non-farm business in non-agriculture. Households with use rights on land can either continue to produce on their farm or walk away from their land to run a non-farm business in non-agriculture. If they walk out of their land, it gets reallocated to a landless household in agriculture. Households have access to a risk-free asset they can accumulate for self-insurance (Bewley, 1986; Huggett, 1993; Aiyagari, 1994). In this section, I describe in detail the choices faced by households in the model.

Demographics. There is a unit measure of dynastic households who discount utility at rate β and a village leader who has control over agricultural land, appropriates unused land, and grants use rights to landless agricultural households in the economy.

Preferences. Dynastic households maximize expected discounted utility.

$$U = \sum_{t=s}^{\infty} \beta^{t-s} u(c_t) \tag{3}$$

where we can abstract from sectoral differences in consumption goods as we assume they are perfectly substitutible. The utility function is CRRA with risk aversion parameter σ so per period utility can be written as: $u(c) = \frac{c_t^{1-\sigma}}{1-\sigma}$

Household endowments. Households are endowed with one unit of labor, which they supply inelastically. Their permanent productivities $\bar{z}_a$ in agriculture and $\bar{z}_n$ in non-agriculture are drawn from a joint log-normal distribution. Households face idiosyncratic persistent shocks $z_{n,t}$ in non-agriculture. When working in agriculture, households can be either landless or landholders. Agricultural households who are landholders are endowed with a common endowment of land ℓ_c and are called smallholder, while those who are landless are called farm workers.

Technology. Smallholder have access to farming technology based on their permanent productivities and produce $\bar{z}_a f(\ell_c, n_{a,t})$ by combining ℓ_c with hired labor. They earn profits and the return on their own labor $w_{a,t}$. At any point, the household can choose to work in non-agriculture and access a technology to produce according to $y_n(\bar{z}_n, z_{n,t}) = \bar{z}_n z_{n,t}$ where $z_{n,t}$ is drawn from an AR1 process that can be written as follows:

$$\log z_{n,t+1} = \rho_z \log z_{n,t} + \sigma_z \epsilon_t \tag{4}$$

where ϵ_t is an i.i.d random variable drawn from a standard normal distribution.

Household decisions. Households make dynamic decisions according to five individual states: landholding status, permanent productivity in each sector, persistent shocks in non-agriculture and assets. These individual states can be summarized according to $x \in \mathcal{L} \times \mathcal{B} \times \mathcal{Z}_n \times \bar{\mathcal{Z}}_n \times \bar{\mathcal{Z}}_a$ such that $\mathcal{L}$ is binary indicator that represents access to land while other states are continuous. Given its endowment and realizations of shocks, households first decide whether to work in agriculture or non-agriculture. Let W and V denote the value functions of landless and landholding households respectively. Define the value function of choosing agriculture as $\bar{W}^a(b, z_n, \bar{z}_n, \bar{z}_a) = \pi_a V^a(b, z_n, \bar{z}_n, \bar{z}_a) + (1 - \pi_a) W^a(b, z_n, \bar{z}_n, \bar{z}_a)$. Then we can write the sector choice of the households as:

$$W(b, z_n, \bar{z}_n, \bar{z}_a) = \max\left\{\bar{W}^a(b, z_n, \bar{z}_n, \bar{z}_a), W^n(b, z_n, \bar{z}_n, \bar{z}_a)\right\} \tag{5}$$

$$V(b, z_n, \bar{z}_n, \bar{z}_a) = \max\left\{V^a(b, z_n, \bar{z}_n, \bar{z}_a), W^n(b, z_n, \bar{z}_n, \bar{z}_n)\right\} \tag{6}$$

where π_a is the probability of land allocations by the village chief; V^a is the value function of being a landholder and working in agriculture as a smallholder; W^a is the value function of being landless and working in agriculture as farm worker; and W^n is the value function of working in non-agriculture. Note that the allocation of the village chief is directed towards landless households.

I now explain each of the underlying value functions in turn. Landless households choosing agriculture can either become a farm worker or a smallholder. Those choosing agriculture but not allocated a plot of land become landless farm workers. They earn the wage in agriculture and divide their assets between consumption and savings but are subject to a borrowing constraint. Let W^a be the value function of landless farm workers, their problem can be written as:

$$W^a(b, z_n, \bar{z}_n, \bar{z}_a) = \max_{c,b'} \quad u(c) + \beta \mathbb{E}_{z'_n|z_n} [W(b', z'_n, \bar{z}_n, \bar{z}_a)] \tag{7}$$

$$c + b' = w_a + Rb, \quad b' \geq -\phi \tag{8}$$

With probability π_a, the household is allocated a plot by the village chief and produce using the farm technology. Households face constraints on their borrowing which is limited to ϕ. Those who become smallholders now earn profits along with the return on their labor. Let V^a be the value function when being allocated a plot of land. Then, the consumption-savings problem of the household can be written as:

$$V^a(b, z_n, \bar{z}_n, \bar{z}_a) = \max_{c,b',n_a} \quad u(c) + \beta \mathbb{E}_{z'_n|z_n} [V(b', z'_n, \bar{z}_n, \bar{z}_a)] \tag{9}$$

$$c + b' = y_a(\bar{z}_a; w_a) + Rb, \quad b' \geq -\phi \tag{10}$$

where $y_a(\bar{z}_a; w_a)$ are the profits resulting from the period by period maximization of profits solved by the farmers which can be written as:

$$y_a(\bar{z}_a; w_a) = \max_{n_a} \bar{z}_a f(\ell_c, n_a) - (n_a - 1)w_a \tag{11}$$

When choosing to work in non-agriculture, the household faces income risk. There is no possibility of being allocated land when choosing to work in non-agriculture. Let W^n denote the value of non-farm work. These households now divided their risky income into consumption and savings according to:

$$W^n(b, z_n, \bar{z}_n, \bar{z}_a) = \max_{c,b'} \quad u(c) + \beta \mathbb{E}_{z_n'|z_n}[W(b', z_n', \bar{z}_n, \bar{z}_a)] \tag{12}$$

$$c + b' = y_n(\bar{z}_n, z_n) + Rb, \quad b' \geq -\phi \tag{13}$$

where summarizes the value functions of the three occupations in the economy.

Like the landless households, landholders choose whether to work in agriculture or non-agriculture. If they choose agriculture, they are smallholder farmers, have value function V^a and face the consumption savings problem described in equations 9 and 10. Households choosing to work in non-agriculture have value function W^n and face the consumption savings problems described in equations 12 and 13. A key feature of my model is that when landholders choose to work in non-agriculture, they lose their land rights and walk into the next period with the value function of a landless household W. Having described the decision problems facing all agents in the economy, I now describe the equilibrium concept I rely on for my analysis.

Equilibrium. I focus on a stationary equilibrium where the distribution of states in the economy are constant. A stationary equilibrium in this economy is an agricultural wage (w_a), a land allocation probability (π_a), and value functions $V(b, z_n, \bar{z}_n, \bar{z}_a)$, $W(b, z_n, \bar{z}_n, \bar{z}_a)$ such that:

- Landless and landholding households solve their individual problems
- Agricultural labor markets clear:

$$\begin{aligned}(n_a(w_a) - 1) &\left(\int_V \mathbb{I}_{(occ=a)}(x) dF(x) + \pi_a \int_W \mathbb{I}_{(occ=a)}(x) dF(x) \right) \\ &= (1 - \pi_a) \int_W \mathbb{I}_{(occ=a)}(x) dF(x)\end{aligned}$$

- Communal land used in agriculture production cannot exceed supply

$$\ell_c \int_V dF(x) = \ell_c N_c \leq L$$

- Stationarity implies the flow of new communal households equals the the flow of new landless households

$$\pi_a \int_W \mathbb{I}_{(occ=a)}(x)dF(x) = \int_V \mathbb{I}_{(occ=n)}(x)dF(x)$$

5 Model Mechanisms

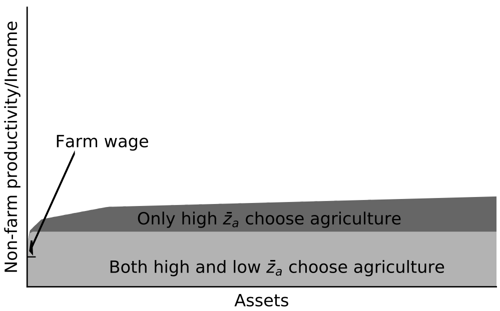
In this section, I demonstrate how my model captures the incentives households face under communal land and their implications for resource allocation, insurance, and redistribution. I begin by describing the incentives households face to choose their occupation and their implications for insurance and redistribution. I then compare the dynamic model to a nested static case that encompasses many of the features that are common in the land misallocation literature. All analysis in this section is performed in partial equilibrium where both w_a and π_a are given.

5.1 Occupation Choice

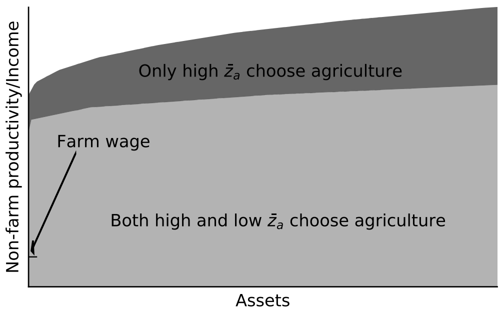
I begin by showing how occupation choices shape the distributions of farm productivities operating in agriculture. In the model, the only households eligible for new reallocations of communal land are landless farm workers. However, land is scarce, and there may not be available land for all those who want it. If a landholding household stops farming, they walk into the next period as a landless household and face risk in non-agriculture. In this section, I demonstrate the different incentives this creates for households and its implication for the distribution of operating farmers. As a result, throughout this section, I fix the permanent non-agricultural productivity and focus instead on other state variables.

I first consider a household with low and high farm productivity and study their occupation decisions with respect to other state variables: assets and non-farm shocks. Figure 1 plots the households who choose agriculture (shaded) and non-agriculture (unshaded) as a function of assets and non-farm productivity for landless households (figure 1a) and landholders (figure 1b). The light-shaded region represents the area under which both high and low farm-productivity households choose agriculture. The dark shaded area represents the region under which only the high farm productivity household choose agriculture. Regardless of their productivity, both households face the same return on their labor (farm wage). Their productivity determines the income they generate beyond the return on their wage. Finally, for any point on the graph, the corresponding point on the y-axis represents the household income in non-agriculture.

Figure 1: Occupational Choice and Agricultural Productivity



(a) Landless



(b) Landholder

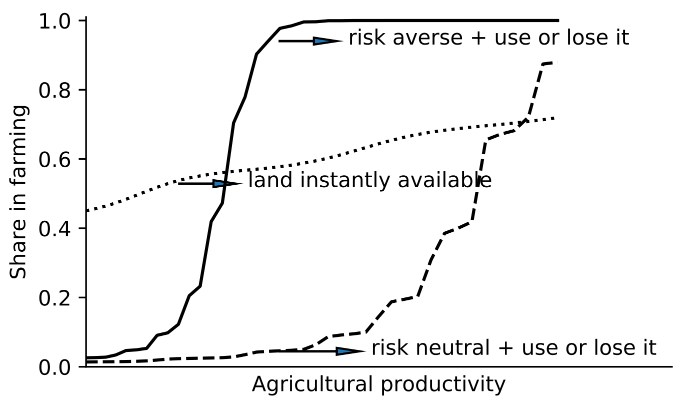
Notes: While $\bar{z}_n$ is fixed, non-farm productivity includes both components of productivity $\bar{z}_n z_n$. Since this also coincides with non-farm earnings when choosing non-farm, the y-axis has the same scale as earnings.

Landless households balance their current possibilities in non-agriculture with the potential income gains from acquiring land and the costs associated with spending time as a farm worker. The household only becomes a smallholder when allocated a plot of land by the village chief. As a farm worker, the household earns the prevailing wage in agriculture. As a smallholder, they earn both the wage and the farm profits. This implies more productive farmers generate more income on the farm and, henceforth, are willing to forgo more non-farm income for the chance of being allocated land. The dark shaded region represents the level of assets and non-farm shocks at which the unproductive farmer would be giving up too much non-farm income by staying in agriculture. The curvature on assets illustrates the role of borrowing constraints in limiting how low asset households can afford to work as landless farm workers while waiting for communal land. Finally, as $\pi_a \to 0$ and households have no chance to acquire land, the upper bound on the shaded region converges to the farm wage level and households become less willing to go into agriculture.

Landholders balance their current income in agriculture with their possibilities in non-agriculture, the degree of risk, and the degree of land availability. Those choosing to leave agriculture earn risky income and become landless households in the following period. The degree of land availability determines the period households spend on average as a farm worker before being allocated a plot of land. This implies that productive farmers face larger consumption loss if their non-farm productivity drops. This amplifies the degree to which productive farmers are willing forgo higher income in non-agriculture in order to keep their land (as illustrated by the dark shaded region in figure 1b). In contrast with landless households, as $\pi_a \to 0$, the shaded area expands, and landholders become less willing to leave agriculture.

In the next exercise, I study the implications for the distribution of operating farmers that emerge from these occupation choices. To this end, I compute the occupation and savings decisions of households in my model. I then simulate a panel of households for each level of permanent farm productivity $\bar{z}_a$ until the rate of households entering agriculture is the same as those leaving (stationary distribution). This allows me to measure in the model the mass of smallholders that emerge in a stationary distribution at each level of farm productivity. I now use this exercise to illustrate how the new mechanisms I emphasize matter for the distribution of operating farmers and, ultimately, for aggregate productivity.

Figure 2: Distribution of Farm productivity, Uninsurable Risk and Land Availability

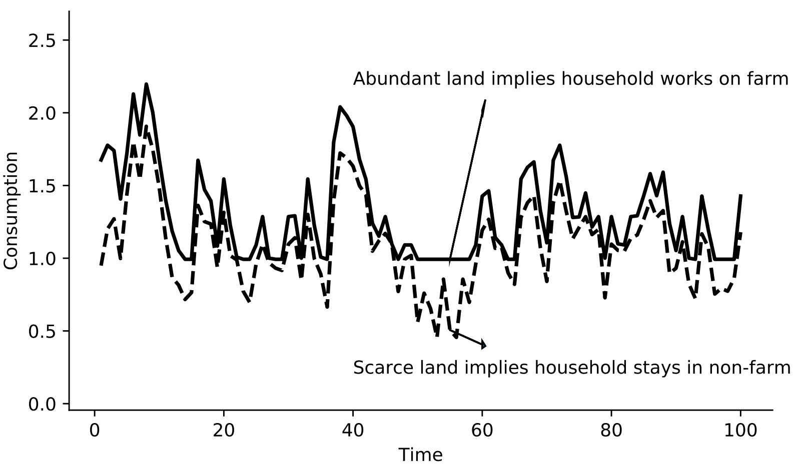


I now illustrate how each additional feature I introduce in my paper affects the distribution of operating farmers. In figure 2, I plot at each level of farm productivity the share of households that end up as smallholder farmers in the stationary distribution. In a standard Roy model with permanent productivity differences and no transitory shocks, this will be 1 or 0. The dotted line represents a model where I shut down the land scarcity channel by setting $\pi_a = 1$. Now, households compare their transitory income in non-agriculture to that on the farm and choose the occupation that pays more. It results in a mass of households experiencing positive shocks and moving to non-agriculture and a mass experiencing negative shocks in moving to agriculture. The dashed line represents the case in which I shut down risk by setting risk aversion parameter $\sigma = 0$ and introducing the use it or lose it mechanism by setting $\pi_a = 0.05$. This distorts occupation choice along the non-farm shock dimension. Because households are risk-neutral, they compare lifetime incomes under each occupation and choose the higher with no insurance concerns. This removes the curvature on the occupation cutoff curve along the asset dimension. This results in fewer unproductive farmers and more productive farmers ending up in agriculture. As we make households risk averse, they now have insurance concerns and dislike consumption volatility but are borrowing constrained. The prospect of becoming a landless farm worker implies risk averse households are willing to give up even more non-farm income to go/stay in agriculture. As land becomes more scarce, the costs of risk rise and households already on the land become less and less willing to give it up by switching sectors.

5.2 Insurance and Redistribution

In this section, I focus on the redistribution and insurance properties of communal land. The main parameter that affects insurance and redistribution in my model is π_a, determining the extent to which households can access land. To this end, I vary π_a and study occupation and consumption patterns across time and in the cross-section.

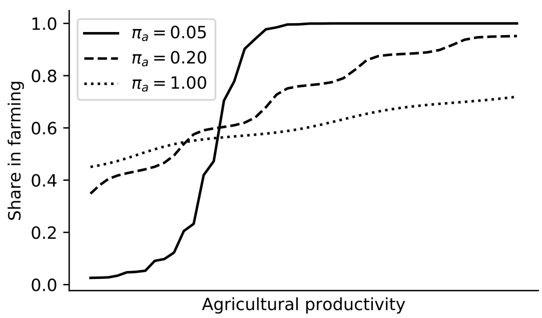
Figure 3: Consumption Path of Household under Scarce vs Abundant Communal Land



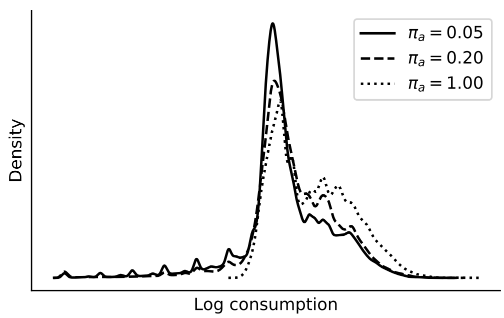
In the first exercise, I fix the permanent productivity and simulate a path of non-farm shocks that households experience under a low π_a and under high π_a. I first compute optimal individual decisions taking the wage $w_a = 0.1$ as given. I then simulate these households until a stationary equilibrium is reached. From the stationary distribution, I simulate the household over 100 periods. The non-agricultural income is a combination of their permanent and transitory components and move over time $\bar{z}_n z_{n,t}$. The household's farm income if they access land is a function of the permanent farm productivity $y(\bar{z}_a) = \pi(\bar{z}_a) + w_a$.

Consider first the consumption path for a household when $\pi_a = 1$ is represented by the solid line. The consumption level of this household never drops below their income from smallholder farming. They move in and out of agriculture several times, experience higher consumption levels as a result, and avoid states of the world that would bring their income below the farm income. This demonstrates how higher levels of π_a allow the household to enjoy higher consumption levels. The effect on volatility is less clear. For landless households, a higher π_a allows them to avoid low-income levels by giving them quick access to land. In contrast, it will enable smallholders to go out and take advantage of higher productivity in non-agriculture more often. In other words, while the redistribution effect of higher π_a is obvious by allowing households to enjoy higher consumption levels, its impact on consumption volatility is less clear.

Figure 4: Agricultural Productivity and Redistribution Tradeoff



(a) Land Scarcity and Productivity



(b) Land Scarcity and Redistribution

In figures 4a and 4b I study the distributional consequences of π_a. As in figure 2, in figure 4a, I plot the share of households that end up as landholders in the stationary distribution. This shows how a higher π_a leads more unproductive farmers to be on communal land and fewer productive farmers to be landholders. Figure 4b plots the cross sectional distribution of log consumption along the same three levels of π_a. It shows how higher π_a shifts consumption to the right along the entire distribution by allowing households to access communal land more often. This means households both enter and exit communal land more often.

5.3 The Nested Static Case

I now turn to the nested static case that emerges by turning off risk and making land available to anyone who wants to farm. I begin by showing the essential components of the static model that differ from the dynamic model presented in section 4. I calibrate the static model with moments derived in section 3. I then add risk and different degrees of land availability. I conclude the section with a discussion of key features of my model that rationalize the variation of moments arising from changes in the degree of land availability when households face risk in non-agriculture.

The nested static economy entails turning off risk and making land available to all households who want it by setting $\pi_a = 1$ and $\sigma_z = 0$. This nested case in my model is a simplified version of Chen (2017) and speaks to the same issues discussed in Adamopoulos, Brandt, Leight and Restuccia (2021). In particular, land distribution is egalitarian and function as a subsidy to unproductive farmers and a tax on productive ones. The model captures permanent sector specific productivity differences and access to communal land that is available to all households. In a static model with no risk, households no longer accumulate assets for self insurance. Unlike the model in Chen (2017), this nested version of my model abstracts from differences in land and capital endowments as well as imperfect substitutability in the consumption of agricultural and non-agricultural goods.

The primary difference between the model delineated in section 4 and this simplified version lies in the occupation choice. In the static model, households have no insurance concerns and choose occupation based on the income they can earn in each sector. Since the allocated land is common across households, the occupation choice depends on the household's permanent productivity in each sector and model parameters, i.e.

$$y(\bar{z}_a, \bar{z}_n) = \max\{y_a(\bar{z}_a), y_n(\bar{z}_n, 1)\} \tag{14}$$

So for each permanent state $\bar{z} = (\bar{z}_a, \bar{z}_n)$, all households holding this state choose the same occupation and earn the same income.

Calibration. I calibrate the static model according to Adamopoulos, Brandt, Leight and Restuccia (2021) with one difference: to pin down the employment share in agriculture, they use a non-homothetic term in the consumption of agricultural goods. Instead, I use the mean μ_n of the non-agricultural productivity distribution $\bar{z}_n$. Otherwise, I proceed as them and target the variance of permanent agricultural income of operating farmers, the variance of permanent non-agricultural income of non-farm workers, and the convariance

Table 7: Model and Data Moments

Statistic	Model	Data
Share of employment in agriculture	0.73	0.73
STD of income of non-farm workers	1.05	1.05
STD of income of farmers	0.81	0.84
COV of income of farmers and potential income in non-farm	0.08	0.09

of agricultural income and potential non-agricultural income of operating farmers. The calibration matches well these four moments summarized in table 7 and estimated from the micro data in section 3.

Table 8 displays the calibrated population parameters of the joint distribution of agricultural and non-agricultural productivity that summarizes the static economy. These parameter are also calibrated in the dynamic economy where the same model objects are targeted. This allows us to track precisely how the new ingredients I add change these population parameters once we match the degree of risk and land availability in the Malawian economy.

Table 8: Calibration Parameters of Static Model

Statistic	Expression	Value
Mean of permanent non-farm productivity	μ_n	-0.51
STD of permanent non-farm productivity	σ_n	1.86
STD of permanent farm productivity	σ_a	1.22
COV of agricultural and non-farm productivity	$\sigma_{a,n}$	-0.35

5.4 Static Versus Dynamic Model

To understand how the dynamic features I introduce affect the household's occupation decisions I introduce them one at a time to the static model presented above. I proceed in this section in partial equilibrium where the wage in fixed and π_a changes to demonstrate its effects. I set $\rho_z = 0.4$, $\sigma_z = 1$, and vary π_a to demonstrate how risk interacts with the degree of land availability to change moments influenced by population parameters. I then simulate the model forward until arriving at a stationary distribution where aggregates are constant and the flow of entrants into communal land is equal to the flow of exiters. In a stationary distribution in partial equilibrium, I can derive the same model simulated moments as those in the static economy. I also derive new moments that only make sense through the lens of a dynamic economy but also shed light on the importance of dynamic incentives for the allocation of farmers.

In table 9, I report model simulated moments from the dynamic model under different levels of risk aversion and land accessibility. The first column describes the statistic description, and other columns report the value of this statistic for a risk averse versus risk neutral household across different levels of π_a. The first four rows report the same statistics as those in table 8 while the last two rows report output in agriculture and non-agriculture. If $\pi_a = 1$ and there was no risk $\sigma_z = 0$, the statistics would coincide with those targeted by the static model and reported in table 8.

When households can self-insure against income shocks below their potential income in agriculture, assets play no role in the occupation choice. Risk in non-agriculture imply some households face positive shocks and move to non-agriculture while others face negative shocks and move into agriculture. When $\pi_a = 1$, the household's occupation choice only depends on the income they earn in each occupation, i.e

$$y(z_n, \bar{z}_a, \bar{z}_n) = \max\{y_a(\bar{z}_a), y_n(\bar{z}_n, z_n)\} \tag{15}$$

Households merely flow to the sector that pays more. This is illustrated by the identical policies that arise when $\pi_a = 1$ independent of whether the household is risk averse ($\sigma = 0$ vs $\sigma = 2$). With incomplete markets, households still have insurance concerns but when $\pi_a = 1$, the household always works in the occupation that pays more. This increases the standard deviation of non-farm income as productive workers face positive shocks and unproductive ones face negative shocks.

When $\pi_a < 1$, incomplete markets imply households are willing to forgo profitable opportunities in non-agriculture in order keep their land. Without risk aversion, households only compare the lifetime income in the two occupations rather than the risk incurred in non-agriculture. This is illustrated by the last two columns in table 9. As $\pi_a \to 0$, employment in non-agriculture grows as fewer households are willing to wait to access land via local leaders. When households are risk averse, as $\pi_a \to 0$, incomplete markets imply fewer households experiencing positive shocks in non-agriculture move to non-agriculture. Since productive farmers earn higher income from their land, this particularly affects them who stay in agriculture to avoid long-lasting income losses in a risky non-agricultural sector. This implies land scarcity ensure output stays high in agriculture and is driven down in non-agriculture as productive farmers experiencing positive shocks in non-agriculture disproportionately choose to stay in agriculture.

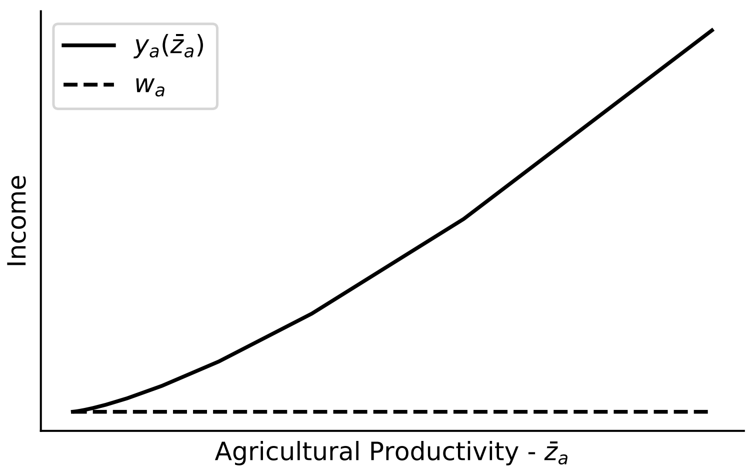
Table 9: Model Simulated Moments

	$\sigma = 2$			$\sigma = 0$		
Moment description	$\pi_a = 1$	$\pi_a = .1$	$\pi_a = .01$	$\pi_a = 1$	$\pi_a = .1$	$\pi_a = .01$
STD of perm. farm inc	0.84	0.86	0.86	0.84	0.87	0.87
STD of perm. non-farm inc	1.28	1.21	1.19	1.28	1.21	1.21
COV of non-farm and farm inc	0.03	0.21	0.29	0.03	0.19	0.30
Share of operating farmers	0.66	0.56	0.52	0.66	0.53	0.46
Mean farmer productivity	1.73	1.69	1.66	1.73	1.61	1.53
Y_a	2.23	2.23	2.20	2.23	2.14	2.06
Y_n	5.40	5.29	5.31	5.40	5.38	5.42

Notes: Income is measured in logs. For the covariance, I use the income of operating farmers. The farm income is what the farmer actually earns and non-farm income is the opportunity cost of farming.

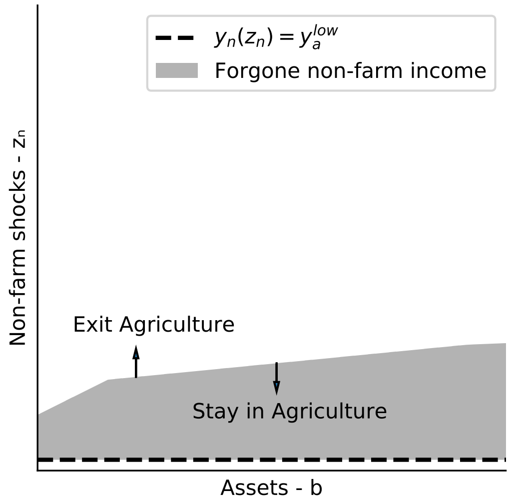
The last two rows of table 9 show that as $\pi_a \to 0$ the average productivity of entrants relative to incumbent farmers widen. This happens because productive farmers who earn more from their land relative to unproductive ones become less inclined to leave their farm behind. As shown in figure 5, the income from operating a farm is increasing in the permanent farm productivity while the income from farm labor is constant and equal to the income of a farmer with $\bar{z}_a = 0$. This implies, productive farmers have more to lose by leaving agriculture. As $\pi_a \to 0$, the exposure to uninsurable risk rises in the non-farm sector and productive farmers respond by sticking to agriculture. This raises the average productivity of farmers with access to communal land and rationalizes the gap between entrants and incumbent farmers.

Figure 5: Agricultural Income by Land Status

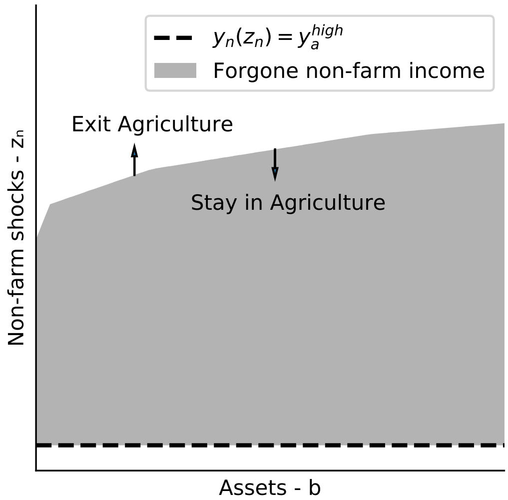


To better understand this, in figure 6, I fix $\bar{z}_n = \bar{z}_n^*$ and plot the occupation choice and non-farm income prospects of a low and high productivity farmer. Note that the more productive farmer earn higher income on the farm as shown by the dashed lines. The upper edge of the shaded region represents the values of z_n and b such that $V^a(b, z_n, \bar{z}_n^*, \bar{z}_a^{low}) = V^n(b, z_n, \bar{z}_n^*, \bar{z}_a^{low})$ in figure 6a and $V^a(b, z_n, \bar{z}_n^*, \bar{z}_a^{high}) = V^n(b, z_n, \bar{z}_n^*, \bar{z}_a^{high})$ in figure 6b when $\pi_a = 0.01$. When $\pi_a = 1$, the shaded region collapses and the households choose the occupation that pays more as in equation 15. This show that as $\pi_a \to 0$ not only does it drive productive farmers to stick to agriculture, it also means they give up more temporary income in non-agriculture to keep their land. This turns out to be very important in the policy experiment when we allow productive farmers to earn income from their land and realize income gains in non-agriculture when the opportunity presents itself.

Figure 6: Income and Occupation by Farm Productivity



(a) Low Productivity Farmer



(b) Higher Productivity Farmer

6 Calibration

In this section, I explain in detail how I discipline the model parameters with microdata. Guided by the last section and the relevance of parameters related to risk and land scarcity, I describe the moments in the data that inform the patterns of occupations and productivity I discussed. I use simulated method of moments to match the model to panel microdata from Malawi. To this end, I use a mix of both micro and macro

Table 10: Calibration Parameters

	Expression	Value
Predetedermined Parameters		
Interest rate	R	0.90
Discount Factor	β	0.96
Risk Aversion	σ	2.00
Land share	α	0.50
Span of control of farmer	ζ	0.50
Elasticity of substitution between land and labor	γ	1.10
Calibrated Parameters		
Mean of permanent non-agricultural productivity	μ_n	-1.80
STD of permanent non-agricultural productivity	σ_n	1.90
STD of permanent agricultural productivity	σ_a	1.25
COV of agricultural and non-agricultural productivity	$\sigma_{n,a}$	-0.40
Probability of land allocation	π_a	0.01
Persistence of non-agricultural productivity process	ρ_z	0.40
STD of non-agricultural productivity process	σ_z	1.00

moments estimated in section 3. I use the IHS years of 2010-2011 and 2012-2013. With moments from the data in hand, the calibration works as follows: I begin from a guess of parameters, solve the model, simulate a panel of households, compute moments from the model, and iterate on parameters until model and moments are matched. Table 10 has the list of parameters that I calibrate along with their values. Below I describe how they were chosen in detail.

Parameters set exogenously. There are five parameters to set exogenously, 4 of which are set to standard values in the literature. I set $\beta = 0.96$ since the model is in yearly frequency and $\sigma = 2$ as in Lagakos et al. (2018). I have three technology parameters to set: ζ determines the span of control of the farmer, γ is the elasticity of substitution between land and labor, and α determines the land share. I set the land share to $\alpha = 0.5$, the span of control parameter to $\zeta = 0.5$ and the elasticity of substitution to $\gamma = 1.1$. Since asset markets are in partial equilibrium as in Lagakos et al. (2018), I also set the interest on assets exogenously. Since my model is set in annual frequency, I set it to $R = 0.90$.

Parameters chosen jointly. There are seven parameters to target. I leave the supply of land L to be consistent with an equilibrium where all of the land is being used, ℓ_c, and μ_a to be normalized. Therefore, I am left to calibrate the probability of gaining access to communal land π_a, the variance and the persistence of the non-agricultural productivity process, σ_z and ρ_z, the variance of the distribution of productivity in agriculture, σ_a, the mean and variance of the distribution of permanent productivity in non-agriculture, μ_n and σ_n and the covariance of permanent agricultural and non-agricultural productivity $\sigma_{a,n}$.

My model speaks to sectoral productivity heterogeneity, risk, and communal land in agriculture. Each of the parameters above has an intuitive mapping to moments in the data. In what follows, I describe how moments are constructed in both the data and the model.

Individual Productivities. To calibrate the parameters of the two-dimensional log-normal distribution of productivities, I use the permanent income estimates from section 3. Each parameter has an intuitive mapping to the estimated parameters from the data. To calibrate the two dispersion parameters σ_a and σ_n, I match the variance of the log of the permanent component of income in the data to its analogue in the model. In the model, each household has a permanent component of income in farm and non-farm. I use the variance of the permanent component of income for those choosing the sector. To calibrate the covariance parameter $\sigma_{a,n}$, I match the covariance of the permanent component of income for those moving from agriculture in 2010 to non-agriculture in 2013 in the data to its analogue in the model. In the model, I save the households in ag in the first period and check who ends up in non-agriculture in the last period. I then take the covariance of their permanent component of income. For the mean of the non-agricultural productivity μ_n, I chose this level such that employment in the non-agriculture sector in the model matches the data.

Risk: The parameters on the non-agricultural shocks σ_z and ρ_z determine the degree of risk in the economy. To calibrate these parameters, I use the estimated transitory component of income in the two sectors. Since I focus on non-agricultural risk, I use the variance of the idiosyncratic component of switchers in agriculture in 2010 and move to non-agriculture in 2013. I take those in agriculture in the first period in the model and follow them for three years. I then use the variance of the transitory portion of income for these households. For the persistence parameter of the non-agricultural income, I use the share of workers in agriculture in 2010 who switched to non-agriculture in 2013. I can compute these shares both in the data and in the model.

Land availability: The degree of land availability in the communal land system is determined by π_a. I use this parameter to pin down the share of farm workers in the model. This parameter plays a key role in the landless decision to work in agriculture as it determines the chances that the landless become landholders.

Table 11: Data and Model Moments

Targeted Moments	Model	Data
Share of smallholders farmers (2010)	0.69	0.64
Share of farm workers (2010)	0.06	0.09
Share in farming in 2010 in non-agriculture in 2013	0.06	0.25
Variance of transitory non-farm income of switchers	0.96	2.10
Variance of permanent non-farm productivity in 2010	0.84	1.20
Variance of permanent farm productivity in 2010	0.78	0.71
Covariance of farm and non-farm productivity of switchers	0.18	0.09

7 Quantitative Analysis

I now use the model to understand the impacts of creating a land market. This experiment is analogous to removing the operation scale restrictions that punish productive farmers and subsidize unproductive ones. It provides a clear distinction between the channels I emphasize and those present in literature.

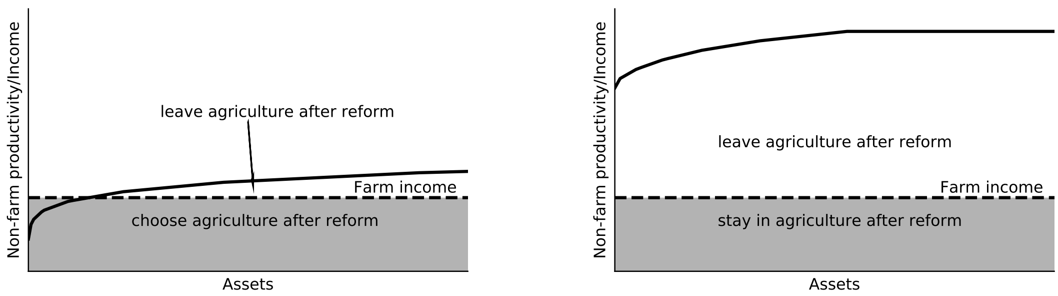
7.1 A Land Rental Market

In the primary policy experiment in the paper, I convert the communal land system into a private rental market. This policy experiment is helpful in two ways. First, many of the land titling reforms throughout Africa give ownership rights to current holders of the land. One of the primary goals of these reforms is to generate an active rental market. Second, this is a common exercise in the literature on static land misallocation, which has been the primary tool used to quantify the benefits of market-oriented land policies in low-income countries. The policy experiment can be implemented analogously in the nested static economy and provides a clear contrast with the features I add in this paper.

The policy takes the landholders in the stationary distribution, gives them ownership over the land. The ownership comes in the form of rental rights. Landowners can choose their operation size $\ell \neq \ell_c$ and earn rents $q\ell_c$ regardless of their occupation choice. In particular, the reform "detaches" the household from the land in the sense that they can now allocate their labor to where the return is higher and still earn a return on their land. Landless households who would like to farm but had no land can now enter the agricultural sector by renting land from landholders. The value functions that define the new economy are described in detail in the appendix A.

The occupation choice is no longer dependent on the ownership of land and instead, it is solely a function of labor earnings in the two sectors. Further, because households do not use assets in production or incur adjustment costs when switching sectors, occupation choice is independent of asset position. As a result, at any time, and for any combination of $\bar{z}_a$ and $\bar{z}_n$, there is a cutoff point $z_n^*(\bar{z}_a, \bar{z}_n)$ such that households go into non-agriculture if transitory shocks fall above this point and into agriculture if it falls below.

Figure 7: Sector Choice and Household Earnings



(a) Landless (b) Landholder

Notes: While $\bar{z}_n$ is fixed, non-farm productivity includes both components of productivity $\bar{z}_n z_n$

In figure 7 I plot the old and new occupation choices of households following the reform. The shaded area represents the levels of assets and non-farm productivity where households after the reform choose agriculture. The solid line represents the occupation cutoff decision of these households before the reform. This implies that households with assets and non-agricultural productivity levels below this curve and above the shaded region are leaving agriculture as a result of the reform. The dashed line represents the new farm income level of this household when they choose to run a farm as well as the new occupation cutoff after the reform. Note that land ownership or assets no longer matter for occupational choices of households. Since the land can be rented, households only compare their income prospects across sectors and choose the sector that pays the higher income.

In what follows in this section, I will first report the aggregate changes from the reform. I will then explain in detail how misallocation in agriculture in reduced. I will then move towards the welfare implications of the reform both in the aggregate and along the distribution of permanent ability differences.

7.1.1 Aggregate Implications

Table 12 reports the aggregate impacts of a land rental market. The first two columns reports aggregate impacts under the static economy while the last two columns reports results for the dynamic economy. I focus on output and employment in each sector. Agricultural employment is split into into smallholder farmers $N_{a,f}$ and farm workers $N_{a,\ell}$. I make this distinction because farmers contribute their permanent productivity towards production while workers do not. Note also that landless is a temporary outcome and no such notion exist in the static model.

The aggregate impacts of a rental market in terms of efficiency are sizable. Productivity rises in agriculture by 22% and in non-agriculture by 15%. The rise in non-agricultural productivity takes place as households are able to rent out their land and enjoy temporary income gains in non-agriculture. The increase in agricultural productivity is affected by reallocation of land across farmers and patterns of entry and exit of workers. Note that despite lower output gains in the static model, agricultural productivity rises significantly more than in the dynamic model.

Table 12: Aggregate Impacts

Statistic	Static		Dynamic	
	Baseline	Post-reform	Baseline	Post-reform
Y_a	1.00	1.41	1.00	1.41
N_a	0.69	0.67	0.72	0.83
Y_a/N_a	1.00	1.45	1.00	1.23
$Y_a/(N_a + N_{landless})$			1.00	1.43
Y_n	1.00	1.01	1.00	1.40
Y_n/N_n	1.00	0.96	1.00	1.29

7.1.2 Selection

In the land misallocation literature, selection plays a central role in driving productivity gains in agriculture. The idea is that limits on private transfers subsidize bad farmers and taxes good farmers. This implies that when private markets are put in place, it triggers the exit of bad farmers and entry of good farmers, propelling large productivity gains in agriculture. In (Adamopoulos, Brandt, Leight and Restuccia, 2021; Chen, 2017) for instance, selection accounts for 2/3 to 3/4 of the agricultural productivity gains in agriculture. In other words, selection amplifies the effects of farm-level distortions.

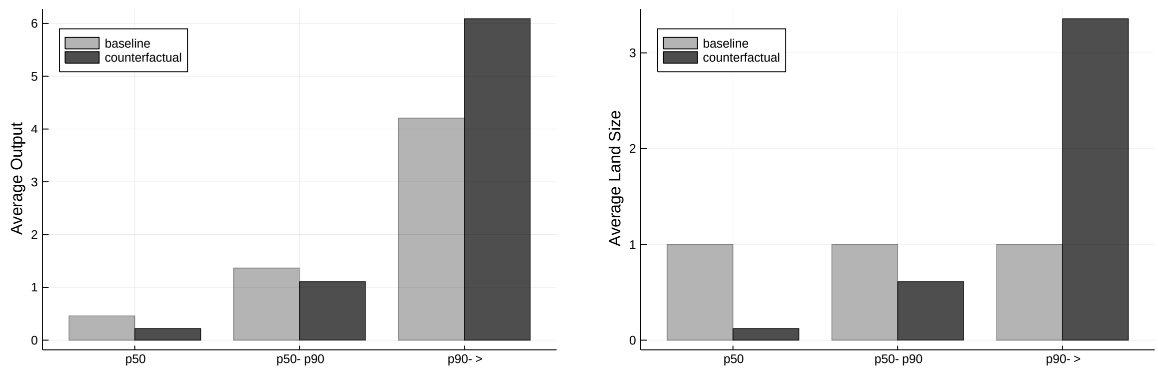
In this section, I study the role of selection in driving productivity gains in my dynamic incomplete markets framework relative to the nested static economy. I do this by conducting an exercise where I first fix occupational choices made at the baseline, and then allow current farmers on land to rent in and out according to the market clearing prices in the counterfactual. In the dynamic economy, agricultural output rises from 2.43 to 3.49. Since employment does not change in this exercise, agricultural productivity is $3.49/0.71 = 4.92$. Agricultural productivity grows to $3.48/0.83 = 4.19$ in the counterfactual and as a result, selection attenuates productivity growth in agriculture by $\log 4.92 - \log 4.19 = 16\%$. This happens because land is severely scarce in Malawi so productive farmers find it very unappealing to leave the farm despite profitable temporary opportunities.

7.1.3 Micro to Macro

I now explore the microeconomic channels by which agricultural productivity rises. Two important changes take place after the reform. First, the land becomes better allocated across those in agriculture. This is a key feature of many studies of land misallocation and is also operative here. Second, patterns of exit and entry change the distribution of abilities in each sector and shape the productivity effects of the reform.

Figure 8a shows the average output in different quantiles of the distribution of output in agriculture. It indicates that unproductive farmers produce less and less of the total output while productive ones produce more. Figure 8b, I plot quantiles of the land size distribution. This illustrates how output changes are driven by productive farmers scaling up their operation by increasing their land size while unproductive ones scale down.

Figure 8: Post Reform Allocation of Land and Output

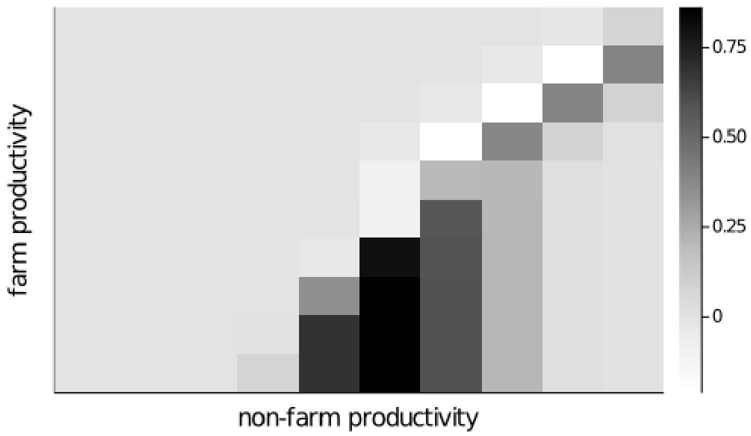


(a) Average Output (b) Average Land Size

Notes: Each bar represents the average of each measure within the percentiles delineated

In figure 9, I plot a heatmap with the proportion change of operating farmers for each productivity level. When the color approaches white, it means that farmers with those productivities are leaving while black implies farmers they are entering the farm-sector. This demonstrates how unproductive farmers are net entrants in this group while the most productive in agriculture are net exiters. This mechanism is only operative because of dynamic incentives that force farmers to give up non-farm income to keep their land. This will counteract the amplification channel of selection on agricultural productivity that is well documented in static models of farm-level misallocation (Adamopoulos, Brandt, Leight and Restuccia, 2021). Further, by allowing former farmers to earn risky income in non-agriculture, this channel leads to gains in the non-farm sector that static models miss.

Figure 9: Net Entrants After Reform

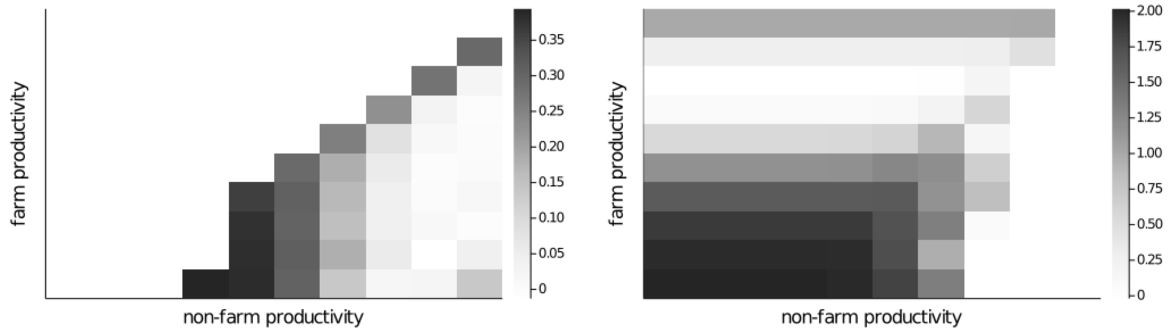


7.1.4 Distributional Implications

In this section, I delve deeper into how the benefits of privatization are distributed across the population. Because the policy experiment is from one steady state to another, I will focus on welfare benefits across the permanent states as well as land status. As a result, the welfare benefits I derive will be an average across assets and stochastic non-farm shocks.

To compare the benefits of privatization I consider households that begin the period with and without land under the private versus communal land regime. Let $k \in \{b, c\}$ denote baseline and counterfactual, and define $W^k(\bar{z}_a, \bar{z}_n)$ and $V^k(\bar{z}_a, \bar{z}_n)$ as the lifetime utility of a household that enters period 0 with and without land respectively. Note that starting in the baseline as a landholder may change over time while in the counterfactual it does not. The measurement of welfare used will be the consumption gain that the household is willing to accept under the baseline to make them indifferent between the baseline and the policy experiment. Hence, a positive value implies the household prefer life under the reform. Details on this computation are written in more detail in appendix B.

Figure 10: Sector Choice and Household Earnings



(a) Landless (b) Landholder

In figure 10, I plot a heatmap of the average consumption equivalent gains for households starting as landless versus landholders. The x-axis represents the permanent productivity in agriculture while the y-axis represents the the permanent productivity in non-agriculture. The blank portions in the graph represent the productivity levels with no households in that land status category. For instance, there are no landless households at the lowest levels of non-agricultural productivity.

The highest gains fall for landholders with the lowest and highest productivity in agriculture. For the low productivity farmers, their income rises substantially. Before the reform, they earned the agriculture wage and very little from their land. After the reform, the wage increases and their earn rents on the land that pays substantially as it is being rented quite heavily by the most productive farmers. High productivity farmers also gain a lot for two reasons. Some of these farmers have low productivity shocks in non=agriculture but their farm income increases as they can rent in some land and grow their operation. On the other hand, those experiencing high productivity shocks in non-agriculture can not leave the farm sector and without the fear of losing their land.

7.1.5 Comparative Statics

Now suppose we have a higher supply of land. In earlier times when land was plentiful, very few families ended up landless as most could access land after trying out their skills in non-agriculture. In this exercise we will increase the land supply from 0.71 to 0.81. This results in $\pi_a = 0.95$

7.1.6 Implications for Structural Change

A key feature of low-income countries is that a larger fraction of the population work in agriculture, and they spend more of their income on food. Reforms like those in my paper often lead to some income growth and shift resources towards non-agriculture. In this section I will briefly discuss the features needed to capture this structural change and their implications.

In order to capture structural change, consumption must display non-homothetic behavior. In particular, as incomes rise, people spend less on food and fewer resources are needed in agriculture. This is generally captured by considering a floor on food consumption that each individual must satisfy. The idea is that if land markets trigger income growth across the board, it will also lead to lower relative expenditure on food. In static models of land misallocation, this force drives down the price of food and pushes unproductive farmers out of agriculture and productive ones back into agriculture. This entry and exit in agriculture play a key role in amplifying the productivity gains from privatizing land in this literature.

Through the lens of my model, the patterns of selection that arise may limit the extent of structural change resulting from land reform. For unproductive farmers, the use it or lose it aspect of farmland has no bite because most of their earnings come from working in other farms and they earn relatively less from their own land. This implies that the occupation choice of unproductive farmers are driven more by static income gains. If the wage growth in agriculture (from less misallocated land) is high enough, unproductive farmers will be swayed to move/stay in agriculture. If their return in non-agriculture is attractive enough and they can rent out their land, they may move to the non-farm sector after the reform. For productive farmers however, dynamic concerns feature more prominently in their occupation choice. In particular, productive farmers leave lots of static temporary income on the table to limit their exposure to uninsurable risk, and they do so increasingly as land grows more scarce. These dynamic concerns drive their exit from agriculture and reduce the potential productivity gains from land markets.

Even if some structural change does take place after the reform, it may come with higher exposure to uninsurable risk for for those leaving agriculture. The occupation choice of households after the reform is driven entirely by labor income, farm, and non-farm revenues. While reduced land misallocation will drive up the wage in agriculture, lower relative demand for food will drive down the price of agricultural goods and consequently revenues on the farm. In static models, income gains from reallocation to the occupation that pays more inevitably makes households better off. In dynamic models like mine, reallocation may subject more households to risk in non-agriculture by reducing the relative value of farm revenues.

8 Conclusion

In this paper I study the implications of risk and incomplete insurance for privatizing communal land systems in developing countries. While privatization delivers economy-wide benefits to all in the economy, it disproportionately benefits those with low productivity in agriculture and non-agriculture who happen to be on the land at the time of the reform. While this has significant welfare benefits by boosting the income of some of the poorest households in the economy, those without land are largely left out of these benefits. Under communal systems, as land becomes scarce, the landless population can be quite sizable and promoting policies that also benefits them may have welfare benefits beyond those of blanket privatization policies. I take up normative issues of land policies in future research.

References

Adamopoulos, T., Brandt, L., Leight, J. and Restuccia, D. (2021), Misallocation, selection and productivity: A quantitative analysis with panel data from china, Technical report, National Bureau of Economic Research.

Adamopoulos, T. and Restuccia, D. (2014), 'The size distribution of farms and international productivity differences', *American Economic Review* **104**(6), 1667–97.

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

Benabou, R. (2002), 'Tax and education policy in a heterogeneous-agent economy: What levels of redistribution maximize growth and efficiency?', *Econometrica* **70**(2), 481–517.

Bewley, T. (1986), 'Stationary monetary equilibrium with a continuum of independently fluctuating consumers', *Contributions to mathematical economics in honor of Gérard Debreu* **79**.

Brooks, W. and Donovan, K. (2020), 'Eliminating uncertainty in market access: The impact of new bridges in rural nicaragua', *Econometrica* **88**(5), 1965–1997.

Brumm, J. and Scheidegger, S. (2017), 'Using adaptive sparse grids to solve high-dimensional dynamic models', *Econometrica* **85**(5), 1575–1612.

Chen, C. (2017), 'Untitled land, occupational choice, and agricultural productivity', *American Economic Journal: Macroeconomics* **9**(4), 91–121.

Chen, C., Restuccia, D. and Santaeulàlia-Llopis, R. (2021), 'The effects of land markets on resource allocation and agricultural productivity', *Review of Economic Dynamics* .

De Magalhães, L. and Santaeulàlia-Llopis, R. (2018), '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.

Deininger, K. W. et al. (2003), *Land policies for growth and poverty reduction*, World Bank Publications.

Donovan, K. (2020), 'The equilibrium impact of agricultural risk on intermediate inputs and aggregate productivity', *The Review of Economic Studies* .

Feder, G. and Noronha, R. (1987), 'Land rights systems and agricultural development in sub-saharan africa', *The World Bank Research Observer* **2**(2), 143–169.

Floden, M. (2001), 'The effectiveness of government debt and transfers as insurance', *Journal of Monetary Economics* **48**(1), 81–108.

Gollin, D., Lagakos, D. and Waugh, M. E. (2014), 'Agricultural productivity differences across countries', *American Economic Review* **104**(5), 165–70.

Gottlieb, C. and Grobovšek, J. (2019), 'Communal land and agricultural productivity', *Journal of Development Economics* **138**, 135–152.

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.

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

McGrattan, E. R. (1996), 'Solving the stochastic growth model with a finite element method', *Journal of Economic Dynamics and Control* **20**(1-3), 19–42.

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

Peters, P. E. and Kambewa, D. (2007), 'Whose security? deepening social conflict over 'customary'land in the shadow of land tenure reform in malawi', *The Journal of Modern African Studies* **45**(3), 447–472.

Restuccia, D. and Santaeulalia-Llopis, R. (2017), Land misallocation and productivity, Technical report, National Bureau of Economic Research.

Roy, A. D. (1951), 'Some thoughts on the distribution of earnings', *Oxford economic papers* **3**(2), 135–146.

Townsend, R. M. (1994), 'Risk and insurance in village india', *Econometrica: journal of the Econometric Society* pp. 539–591.

Udry, C. (1994), 'Risk and insurance in a rural credit market: An empirical investigation in northern nigeria', *The Review of Economic Studies* **61**(3), 495–526.

White, L. (1989), *Magomero: portrait of an African village*, Cambridge University Press.

A Counterfactual Private Economy

This section describes the counterfactual economy that I use in order to quantify the insurance and redistribution value of communal land. This economy is analogous to having the share of the population in the stationary equilibrium on communal land with permanent rights on communal land. All functional forms are the same. The bellman equations have the same interpretation, however the binary land state is now permanent.

Individual policies Therefore the occupation choices can be written as:

$$W(b, z_n, \bar{z}_n, \bar{z}_a) = \max \{V^a(b, z_n, \bar{z}_n, \bar{z}_a), W^n(b, z_n, \bar{z}_n, \bar{z}_a)\}$$
$$V(b, z_n, \bar{z}_n, \bar{z}_a) = \max \{V^a(b, z_n, \bar{z}_n, \bar{z}_a), W^n(b, z_n, \bar{z}_n, \bar{z}_a)\}$$

where these value functions for each activity $k \in \{a, n\}$ can be written as follows:

$$V^k(b, z_n, \bar{z}_n, \bar{z}_a) = \max_{c,b'} \quad u(c) + \beta \mathbb{E}_{z'|z} [V(b, z_n, \bar{z}_n, \bar{z}_a)]$$
$$c + b' = y_k(z_n, \bar{z}_n, \bar{z}_a) + q\ell_c + Rb, \quad b' \geq -\phi$$

$$W^k(b, z_n, \bar{z}_n, \bar{z}_a) = \max_{c,b'} \quad u(c) + \beta \mathbb{E}_{z'|z} [W(b, z_n, \bar{z}_n, \bar{z}_a)]$$
$$c + b' = y_k(z_n, \bar{z}_n, \bar{z}_a) + Rb, \quad b' \geq -\phi$$

where $y_k(z) = z$ if $s = n$ and $y_k(z) = w_a$ if $s = a$. Further, q is the rental rate on land.

Equilibrium. A stationary equilibrium in this economy is an agricultural wage (w_a), a rental rate on land (q), and value functions $V((b, z_n, \bar{z}_n, \bar{z}_a), W((b, z_n, \bar{z}_n, \bar{z}_a)$ such that:

- Households solve their individual problems
- Household farms choose labor and land optimally
- Agricultural labor markets clear
- Agricultural land markets clear

B Welfare Decomposition

In this section I describe in detail the decomposition as described in Floden (2001). It's a slightly simplified version of what is in his paper as I don't have leisure in my economy. The goal is to devise a method that allows me to compare the welfare properties of the economy under communal land and one in which communal land is privatized. Undoubtedly, this change in policy will affect levels of consumption, distribution of income and uncertainty households face. We want a method that can capture the relative effects separately. We can start by defining lifetime utility as

Definition 1 *Lifetime utility V can be written as:*

$$V(\{c_s\}_{s=1}^{\infty}) = \sum_{s=t}^{\infty} \beta^{s-t} u(c_s) \tag{16}$$

Definition 2 *Certainty-equivalent consumption bundle $\{\bar{c}\}$ fulfills*

$$V(\{\bar{c}_s\}_{s=1}^{\infty}) = E_t V(\{c_s\}_{s=t}^{\infty}) \tag{17}$$

Now let C be the average consumption, $C = \int c d\lambda$ and let $\bar{C}$ denote the certainty equivalent consumption $C = \int \bar{c} d\lambda$. We are now ready to describe the utilitarian social welfare, U

Definition 3 *The utilitarian social welfare function is defined as:*

$$U = \int E_t V(\{c_s\}_{s=t}^{\infty}) d\lambda \tag{18}$$

And finally we are ready to define the welfare gain of a particular policy compared to a baseline.

Definition 4 *The utilitarian welfare gain of a policy change, ω_U is defined by*

$$\int E_t V(\{(1+\omega_U)c_s\}_{s=t}^{\infty}) d\lambda^A = \int E_t V(\{c_s\}_{s=t}^{\infty}) d\lambda^B \tag{19}$$

We can think of ω_U as the percent of lifetime consumption agents in economy A are prepared to give up to get the policy change. This is the classic measure of consumption equivalence from Lucas's calculations of the costs of business cycles. We can compute this for different groups. For example, I calculate this for different quintiles of the non-agriculture quintile distribution and of the income distribution. All one would have to do is adjust measures λ^A and λ^B so that it captures the designated groups. Now we are ready to get measures required to isolate effects through insurance versus redistribution. The idea is to first get the certainty equivalent consumption for each individual, the inequality is measured by the distribution of certainty equivalent consumption while uncertainty is measured by comparing differences in actual and certainty-equivalent consumption.

Definition 5 *The cost of uncertainty, p_{unc} is defined as:*

$$V(\{(1+p_{unc})C\}_{s=t}^{\infty}) = V(\{\bar{C}\}_{s=t}^{\infty}) \tag{20}$$

The cost of uncertainty can be intuitively thought of as the percent of average consumption agent is willing to give up in order to just get the certainty equivalent average every period (holding inequality fixed).

Definition 6 *The cost of inequality, p_{ine} is defined as:*

$$V(\{(1+p_{ine})\bar{C}\}_{s=t}^{\infty}) = \int V(\{\bar{c}\}_{s=t}^{\infty})d\lambda \tag{21}$$

Now we are ready to define the welfare gains from these various channels.

Definition 7 *The welfare gains from increased levels ω_{lev}:*

$$\omega_{lev} = \frac{C^B}{C^A} - 1 \tag{22}$$

Definition 8 *The welfare gains from lower uncertainty ω_{unc}:*

$$\omega_{unc} = \frac{1-p_{unc}^B}{1-p_{unc}^A} - 1 \tag{23}$$

Definition 9 *The welfare gains from lower inequality ω_{ine}:*

$$\omega_{ine} = \frac{1-p_{ine}^B}{1-p_{ine}^A} - 1 \tag{24}$$

This implies the total welfare gain ω_U approximately approaches:

$$\omega_U = (1+\omega_{lev})(1+\omega_{unc})(1+\omega_{ine}) - 1 \tag{25}$$

C Data Construction

In order to use the data to inform the parameters of my model I need to compute measures from the data with clear analogues in the model. In this section I describe these measures and how I compute them. First, I need measures of income for the three occupations in the model, i.e. agricultural income of landed farmers, income of landless agricultural cultivators, and non-agricultural income.

Agricultural Income of Smallholders For agricultural income, I only compute earnings from non-permanent crops. Fisheries and livestock income will be capital income from fishing equipment and livestock capital. I consider only income from the rain season as in Restuccia and Santaeulalia-Llopis (2017) where Value added in agriculture can be written as:

$$VA_{a,i} = Rev_{c,i} + P_{c,i}(Output^{z}_{c,i} - Sold_{c,i}) - Cost_{c,i} \tag{26}$$

and represents value added from product c for household i. $Rev_{c,i}$ represents household i's revenues from selling crop c, $Output^{z}_{c,i} - Sold_{c,i}$ represents the fraction of production of crop c that household i keeps for its own consumption and $P_{c,i,r}$ is the price received by household i for crop c (sale value of own consumption), which is replaced by the regional price when such price cannot be inferred for household i (households that report production but no sales). In order to compute this price, I proceed as follows:

1. If household i sold crop c, I use reported sales $Rev_{c,i}$ and quantity sold $Q_{c,i}$ and compute $P_{c,i,r} = Rev_{c,i}/Q_{c,i}$
2. Otherwise I attribute the median price of the crop sold by other households in the same region if available, meaning $P_{c,i} = \bar{P}_{c,j}$ where j lives in the same region as i

Finally, for the production of each crop, the household reports costs associated with various inputs and factors. I aggregate costs across inputs

$$Cost_{c,i} = \sum_{v} Cost^{v}_{c,i} \tag{27}$$

where $v \in \{intermediates, labor, capital, land, transportation\}$ represents the costs associated with production. After computing $VA_{c,i}$ for each crop and household, define household farm earnings as the sum of value added across crops

$$VA_i = \sum_c VA_{c,i} \tag{28}$$

which represents the income of farmers that get their land in Malawi through the customary tenure system.

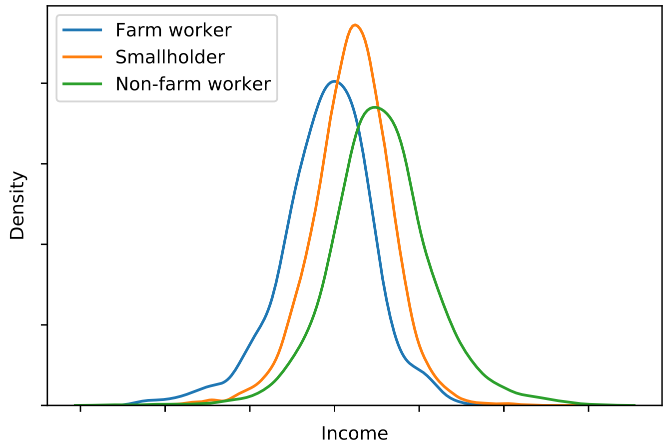
Farm Work Income. There are two ways to compute wages paid to agricultural labor. In the time use side, the only labor income that is tied to agriculture are wages paid in the formal sector. The farms paying these salaries are large and export oriented and highly regulated by the government. These workers are not the ones supplying labor to smallholder farmers who get land through customary tenure. There is another category in the time use portion of the survey that asks for the wage received for informal labor in rural areas. This is the first candidate for w_a. Another way of computing this is by looking at the production side and measuring labor payments made by landed farmers. These can be seen in figure **??** for each one of the years in our panel data. In short this tells us that in 2010, rural wages were reportedly higher than producers reported paying to agricultural labor. This is not surprising as perhaps that year, there were more non-agricultural activities in rural areas that households earned income from - and those paid more than agriculture. In 2013 however, they followed closely one another.
Non-agricultural Income. There are two types of non-agricultural income. Some households run non-agricultural businesses and report in the survey the revenues and costs over the year. Therefore I can write household non-agricultural business income as:

$$VA_{b,i} = Rev_{b,i} - Cost_{b,i} \tag{29}$$

In order to have a comparable measure against other income sources, I compute the non-agricultural business income per hour of household member spent on working at the business. Further, there is also a set of formal workers who earn non-agricultural wages. These are professional workers with education who work in the formal and government sectors of the economy. I also measure their hourly wage.

Summary. In order to assign an occupation to households in our data, I assign them to the occupation (landed agriculture, agricultural labor, non-agriculture) where the household spends most of their time in. After doing this, I can characterize the distribution of income across the three occupations in our economy as seen in figure 11.

Figure 11: Distribution of Sectoral Incomes



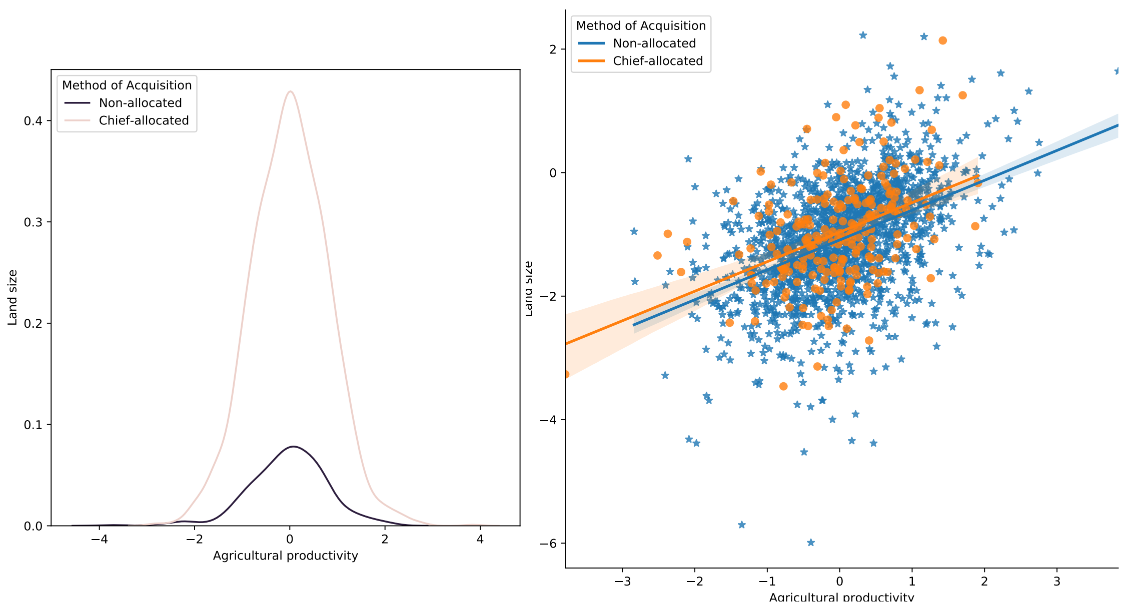


Figure 12: Distribution of Chief Allocated Land

D Computational details

Solving models with continuous and discrete choices presents several challenges and this section details the steps I took to overcome such challenges. In order to deal with non-linearities in borrowing constraints and occupational choices as well as a large state space, I used an adaptive sparse grid method as delineated in Brumm and Scheidegger (2017).

In that paper, their adaptive grid is limited to collocation with linear basis functions. I combine their adaptive grid procedure with a finite element method as in McGrattan (1996). This combination allows me to apply their adaptive grid procedure to higher order basis functions. Higher order basis functions are needed to accurately solve for policies, especially around points that are highly nonlinear. Finally, this allows me to solve for the coefficients on the value function using the Newton-Raphson algorithm. Taken together, this improves upon both speed and accuracy over value function iteration methods. I also use extreme value shocks in order smooth out the discrete choices in the model. Below, I explain each step in detail.

D.1 Extreme value shocks

Nondiffereriabilities in the value function due to discrete choices makes calibration very challenging. I add iid extreme value shocks to the discrete choices in order to smooth out this choice and allow the use of derivatives in the calibration. Once I calibrate, I bring the role of these shocks down until it plays no role. This section describes in detail the implementation of these shocks into the finite element framework I use.

Let $\varepsilon_i^W, \varepsilon_i^V$ for $i \in \{a, n\}$ be extreme value preference shocks with common shape parameter σ_ε We can rewrite the individual choices as:

$$W(b, z, \varepsilon^W) = \max\left\{\pi_a V(b, z|a) + (1-\pi_a) W(b, z|a) + \varepsilon_a^W, W(b, z|n) + \varepsilon_n^W\right\} \tag{30}$$

$$V(b, z, \varepsilon^V) = \max\left\{V^a(b, z|a) + \varepsilon_a^V, W^n(b, z|n) + \varepsilon_n^V\right\} \tag{31}$$

Then the occupation dependent bellman equations can be written as follows:

$$V(b, z|a) = \max_{b' \in \mathcal{B}} u(c, z, b, b', c) + \beta \mathbb{E}_{z'|z}\left[\mathbb{E}_{\varepsilon^V}\left[V(b', z')\right]\right] \tag{32}$$

$$V(b, z|n) = \max_{b' \in \mathcal{B}} u(c, z, b, b', p) + \beta \mathbb{E}_{z'|z}\left[\mathbb{E}_{\varepsilon^W}\left[W(b', z')\right]\right] \tag{33}$$

$$W(b, z|a) = \pi_a \left(\max_{b' \in \mathcal{B}} u(c, z, b, b', c) + \beta \mathbb{E}_{z'|z}\left[\mathbb{E}_{\varepsilon^V}\left[V(b', z')\right]\right]\right) + \tag{34}$$

$$(1-\pi_a)\left(\max_{b' \in \mathcal{B}} u(c, z, b, b', p) + \beta \mathbb{E}_{z'|z}\left[\mathbb{E}_{\varepsilon^W}\left[W(b', z')\right]\right]\right) \tag{35}$$

$$W(b, z|n) = \max_{b' \in \mathcal{B}} u(c, z, b, b', p) + \beta \mathbb{E}_{z'|z}\left[\mathbb{E}_{\varepsilon^W}\left[W(b', z')\right]\right] \tag{36}$$

Instead of approximating the occupation specific value function, I approximate the expected value with respect to extreme value shocks. Once I have an approximation of this, I can back out the occupation specific value functions, and consequently other policies. Hence, define $\mathbb{E}_a\left[W(b,z|a)\right] = \pi_a V^a(b,z) + (1-\pi_a)W^a(b,z)$ I can write the following residual equations

$$R^W(b,z) = \mathbb{E}_{\varepsilon^W}\left[W(b,z)\right] - \left(P_a^W \mathbb{E}_a\left[W(b,z|a)\right] + P_n^W W(b,z|n)\right) \tag{37}$$

$$R^V(b,z) = \mathbb{E}_{\varepsilon^V}\left[V(b,z)\right] - \left(P_a^W V(b,z|a) + P_n^W W(b,z|n)\right) \tag{38}$$

where the probabilities follow the usual expression $P_a^V = \frac{\exp(V(b,z|a)/\sigma_\varepsilon)}{\exp(V(b,z|a)/\sigma_\varepsilon)+\exp(V(b,z|n)/\sigma_\varepsilon)}$ [1]. The goal is to approximate $\hat{W}(b,z) = \mathbb{E}_{\varepsilon^W}\left[W(b,z)\right]$ and $\hat{V}(b,z) = \mathbb{E}_{\varepsilon^V}\left[V(b,z)\right]$. In order to numerically approximate this, I use finite element methods. Now consider solving the model over the domain $\mathcal{B}$. In finite element analysis, I can define non-overlapping intervals over $\mathcal{B}$ as $\{[b_0,b_1],...,[b_i,b_i+1],...[b_{k-1},b_k]\}$. Then define lagrange basis functions of order n - $\psi^n(b,z)$ - over each interval i with associated coefficient vector $\boldsymbol{\theta}$. Then the task is to find $\boldsymbol{\theta}$ such that

$$\int_{b_i}^{b_{i+1}} \psi_i(b,z) R^J(b,z;\theta)db = 0, \quad i=0,1,...,k, \quad z \in \mathcal{Z}, \quad J \in \{V,W\} \tag{39}$$

and the approximated value function can be written as:

$$\hat{V}(b,z) = \psi_i(b,z)\theta_i + ... + \psi_{i+n-1}(b,z)\theta_{i+n-1} \quad b \in [b_i, b_{i+1}] \tag{40}$$

So for cubic basis functions, $n=4$. The ability to evaluate the equations in 39 over each element independently makes finite element analysis very amenable to paralellization. I use distributed memory over the permanent states in my model and shared memory to solve the occupation specific value functions at once.

[1] Numerically, this blows up easily, hence we have to make the following transformation in order to make the algorithm more stable. First let $\bar{V}(a,b) = \max\{V(b,z|a), V(b,z|n)\}$. Then we can multiply P_a^V by $\frac{\exp(\bar{V}(a,b)/\sigma_\varepsilon)}{\exp(\bar{V}(a,b)/\sigma_\varepsilon)}$ and get that $P_a^V = \frac{\exp\left((V(b,z|a)-\bar{V}(a,b))/\sigma_\varepsilon\right)}{\exp\left((V(b,z|a)-\bar{V}(a,b))/\sigma_\varepsilon\right)+\exp\left((V(b,z|n)-\bar{V}(a,b))/\sigma_\varepsilon\right)}$

D.2 Solving for Coefficients

Note that solving for the value function requires solving a root finding problem for equations in 39. This implies that having analytical expressions for the jacobian of 39 [2] is an essential step. Unlike the model without extreme value shocks (when the residual equations are linear in $\boldsymbol{\theta}$ like other residual methods like collocation for instance), with extreme value shocks, it takes a bit more work to derive the jacobian since the residual is now nonlinear in the coefficient vector. This is due to the value function appearing in the choice probabilities. This section goes through this process in detail. Let 37 and 38 be the two residual equations that we'll use in order to approximate W and V. Then, these are solely functions of our coefficients and we need the jacobian with respect to coefficients. First, for the current value, it's simple. For any point $b_i \in \mathcal{B}$, if the grid for communal and landless households coincide, then we can write

$$\frac{\partial V(b_i, z)}{\partial \theta_{b_i}} = \frac{\partial W(b_i, z)}{\partial \theta_{b_i}} = \psi(b_i, z)$$

Now note that the maximand on $V^a(b, z), W^a(b, z)$ and $W^n(b, z)$ are $b'_{W,a}, b'_{V,a}$ and $b'_{W,n}$ respectively. Hence, if the order of approximation is k, there will be $k - 1$ non-zero values for the expected bellman for each possible $z' \in \mathcal{Z}$ as well. Hence, at each point $(b, z) \in \mathcal{B} \times \mathcal{Z}$, we have derivatives around optimal decisions in ag and non-ag. This says that in the event that I choose agriculture, there is also some probability that I would have chosen non-agriculture. Either way, if I end up in agriculture, I chose savings as if I was in agriculture.

Finally this implies we are ready to compute all derivatives of the RHS with respect to coefficients. There are two sets of non-zero points in the jacobian for each occupation choice. In agriculture, there is a non-zero point associated with choosing agriculture and another associated with choosing agricultural asset choice. Suppose we have land and want the derivative with respect to unknowns surrounding the agricultural savings choice. Then we can write the derivative as:

$$\frac{\partial RHS_V(b, z)}{\partial \theta_{b',z,z',a}} = P_a^V \beta \pi(z, z') \frac{\partial \hat{V}(b', z')}{\partial \theta_{b',z,z',a}} + \frac{\partial P_a^V}{\partial \theta_{b',z,z',a}} V(b, z|a) + \frac{\partial P_n^V}{\partial \theta_{b',z,z',a}} V(b, z|n) \Rightarrow$$

Since the value function of the landless households is a bit different, we'll also derive

[2] This is the most error prone step in solving for coefficients. In practice, I construct test functions that compute both the numerical and analytical derivatives until they coincide.

them here. Suppose we want the derivatives around the optimal asset choice when the agent gets the land (denoted by f). Let $\tilde{W}^a(b,z) = \pi_a V^a(b,z) + (1-\pi_a)W^a(b,z)$

$$\begin{aligned}
\frac{\partial RHS_W(b,z)}{\partial\theta_{b',z,z',a}} =& P_a^W \pi_a \beta\pi(z,z')\frac{\partial\hat{V}(b',z')}{\partial\theta_{b',z,z',a}} + \frac{\partial P_a^W}{\partial\theta_{b',z,z',a,V}}\pi_a V^a(b,z) + \frac{\partial P_n^V}{\partial\theta_{b',z,z',a,V}}W(b,z|n) \\
\frac{\partial RHS_W(b,z)}{\partial\theta_{b',z,z',a}} =& P_a^W (1-\pi_a)\beta\pi(z,z')\frac{\partial\hat{V}(b',z')}{\partial\theta_{b',z,z',a,W}} + \frac{\partial P_a^W}{\partial\theta_{b',z,z',a,W}}(1-\pi_a)W^a(b,z) + \\
& \frac{\partial P_n^V}{\partial\theta_{b',z,z',a,W}}W(b,z|n)
\end{aligned}$$

and an analogous scheme can be used in order to solve for other derivatives. We can use this to build the weighted residual and analytically solve for the jacobian as described above. This enables us to use a newton in order to solve the individual problem. Now we can move towards getting the partial derivatives from the choice probabilities.

$$\begin{aligned}
\frac{\partial P_a^V}{\partial\theta_{b',a}} =& \frac{\exp\left(V(b,z|a)/\sigma_\varepsilon\right)}{\exp\left(V(b,z|a)/\sigma_\varepsilon\right) + \exp\left(V(b,z|n)/\sigma_\varepsilon\right)}\frac{\partial V(b,z|a)}{\partial\theta_{b',a}}/\sigma_\varepsilon - \\
& \frac{\exp\left(V(b,z|a)/\sigma_\varepsilon\right)\exp\left(V(b,z|a)/\sigma_\varepsilon\right)}{\left(\exp\left(V(b,z|a)/\sigma_\varepsilon\right) + \exp\left(V(b,z|n)/\sigma_\varepsilon\right)\right)^2}\frac{\partial V(b,z|a)}{\partial\theta_{b',a}}/\sigma_\varepsilon \\
=& \frac{\partial V(b,z|a)}{\partial\theta_{b',a}}\left(P_a^V/\sigma_\varepsilon - (P_a^V)^2/\sigma_\varepsilon\right) \\
\frac{\partial P_n^V}{\partial\theta_{b',a}} =& -\frac{\exp\left(V(b,z|n)/\sigma_\varepsilon\right)\exp\left(V(b,z|a)/\sigma_\varepsilon\right)}{\left(\exp\left(V(b,z|a)/\sigma_\varepsilon\right) + \exp\left(V(b,z|n)/\sigma_\varepsilon\right)\right)^2}\frac{\partial V(b,z|a)}{\partial\theta_{b',a}}/\sigma_\varepsilon \\
=& -\frac{\partial V(b,z|a)}{\partial\theta_{b',a}}P_n^V P_a^V/\sigma_\varepsilon
\end{aligned}$$

Further we know that for each z, $\frac{\partial V(b,z|a)}{\partial\theta_{b',a}}$ relates to the future bellman through its local impacts from z' and occupation dependent savings. Consequently, we can expand the expressions above as follows

$$\begin{aligned}
\frac{\partial P_a^V}{\partial\theta_{b',z,z',a}} &= \beta\pi(z,z')\frac{\partial\hat{V}(b',z')}{\partial\theta_{b',z,z',a}}\left(P_a^V/\sigma_\varepsilon - (P_a^V)^2/\sigma_\varepsilon\right) \\
\frac{\partial P_n^V}{\partial\theta_{b'_{V,a,z,z'}}} &= -P_n^V P_a^V \beta\pi(z,z')\frac{\partial\hat{V}(b',z')}{\partial\theta_{b',z,z',a}}/\sigma_\varepsilon
\end{aligned}$$

This complete the derivation of the residual equations 39 and the jacobian with respect to coefficient vector $\boldsymbol{\theta}$ and allows me to solve the individual problem.