University of Toronto
Department of Economics

Working Paper 762

Land Security and Mobility Frictions

By Tasso Adamopoulos, Loren Brandt, Chaoran Chen, Diego Restuccia and Xiaoyun Wei

October 28, 2023

Land Security and Mobility Frictions[†]

Tasso Adamopoulos
York University

Loren Brandt
University of Toronto

Chaoran Chen
York University

Diego Restuccia
University of Toronto and NBER

Xiaoyun Wei
Shanghai Jiao Tong University

October 2023

Abstract

Frictions that impede the mobility of workers across occupations and space are a prominent feature of developing countries. We disentangle the role of insecure property rights from other labor mobility frictions for the reallocation of labor from agriculture to non-agriculture and from rural to urban areas. We combine rich household and individual-level panel data from China and an equilibrium quantitative framework featuring sorting of workers across locations and occupations. We explicitly model the farming household and the endogenous decisions of who operates the family farm and who potentially migrates, capturing an additional channel of selection within the household. We find that land insecurity has substantial negative effects on agricultural productivity and structural change, raising the share of rural households operating farms by over 40 percentage points and depressing agricultural productivity by more than 20 percent. Comparatively, these quantitative effects are as large as those from all residual labor-mobility frictions. We measure a sharp reduction in overall labor mobility barriers over 2004-2018 in the Chinese economy, all accounted for by improved land security, consistent with reforms covering rural land in China during the period.

JEL classification: O11, O14, O4, E02, Q1.
Keywords: land, labor mobility, agriculture, misallocation, household, productivity, China.

[†]We thank Yan Bai, Lin Ma, Rachel Ngai, Marlon Seror, Kjetil Storesletten, Xiaodong Zhu, and seminar participants at Academia Sinica, Arizona State University, Ashoka University, Boston University, Cagliari Macro Dynamics Workshop, Canadian Macro Study Group at UQAM, EMES in Milan, Melbourne Macro Development Workshop, Midwest Macro Meetings at Michigan State University, NASM in Montreal, Novafrica Development Conference, Peking University, Penn State University, Queen's University, SED in Minneapolis and Madison, Texas A&M University, Trinity College Dublin, University of Nottingham, and University of Toronto. Adamopoulos, Brandt, and Chen gratefully acknowledge the support from the Social Sciences and Humanities Research Council of Canada. Brandt thanks the support from the Noranda Chair for International Trade and Economics. Restuccia gratefully acknowledges the support from the Canada Research Chairs program and the Bank of Canada Fellowship program. Funding from the Research Support Board at the World Bank supported data access and collection. The views expressed herein are not necessarily those of the Bank of Canada and are the authors alone. All errors are our own. Contact: Tasso Adamopoulos, aadamo@yorku.ca; Loren Brandt, loren.brandt@utoronto.ca; Chaoran Chen, chenecon@yorku.ca; Diego Restuccia, diego.restuccia@utoronto.ca; Xiaoyun Wei, xiaoyun_wei@sjtu.edu.cn.

1 Introduction

The movement of workers from agriculture to non-agriculture and from rural to urban locations is fundamental to the process of structural transformation and economic development. In many developing and transition economies, explicit and implicit frictions impede the mobility of workers across occupations and space. These frictions can have important implications for sectoral and aggregate productivity as well as welfare. Disentangling their effects is key to understanding the drivers of productivity and prioritizing policy initiatives.

China is a country with well-documented migration restrictions that have been the focus of extensive research (Chan and Zhang, 1999; Tombe and Zhu, 2019; Caliendo et al., 2019). In addition, there are implicit mobility restrictions tied to insecure property rights over farmland (Giles and Mu, 2018; Ngai et al., 2019; Adamopoulos et al., 2022), frictions that have been identified as important in other developing countries (de Janvry et al., 2015). The risk to farmers of losing their use rights over land if they do not farm it themselves may deter land rentals, work outside agriculture, and migration to urban areas. China offers a unique opportunity to assess the relative importance of alternative mobility frictions that are prevalent in the developing world.

In this paper, we combine rich individual- and household-level panel data from China for 2004-2018 and a structural model to quantify the importance of frictions arising from land insecurity and labor mobility restrictions for agricultural productivity and structural change. For 2004, we find that land insecurity constitutes a substantial barrier to labor mobility that is at least as large in its effect on farm employment and agricultural productivity as all residual labor mobility frictions. We also find that total barriers to labor reallocation in China fall substantially by 2018, all of which is attributed to improved land security as other labor mobility costs actually rise slightly. The importance our analysis attaches to improved land security after the early 2000s in facilitating labor mobility is consistent with institutional reforms affecting property rights in farmland, while efforts to relax other labor mobility barriers have stalled (Chan, 2019).

We develop a quantitative framework of structural change and sectoral selection that explicitly models: (i) the land rights regime in China and its endogenous impact on labor mobility out of agriculture; and (ii) the endogenous household decision of who operates the family farm and who migrates, thereby capturing a novel channel of selection within the household. To account for the rich heterogeneity of individual and household choices in the micro data, our model features families and individuals comprising those families. Each individual can work in agriculture in the village as either a farm operator or a wage worker, or migrate to the cities to work in non-agriculture. Individuals may also choose to work part time in farming and non-agriculture. Household members are heterogeneous with respect to their ability in these occupations, which allows for selection across sectors, and within the family. We allow for idiosyncratic distortions in household and individual choices.

Following the egalitarian nature of the land tenure arrangement in China, we assume that all families are allocated the same use rights over farm land. Farmers can adjust the scale of their operation through land rentals. We capture land insecurity in the model in a tractable way through: (a) the risk that families lose their use rights to land they rent out; and (b) a family-specific perceived income loss associated with the loss of this land. In our setup, land insecurity not only deters farmers from renting out land along the intensive margin, but also deters them from completely abandoning the land in the village out of fear of losing access entirely, thus acting as a de facto barrier to labor mobility out of agriculture.

In addition, individual labor supply choices are subject to residual idiosyncratic occupational labor mobility frictions, a catchall for all other institutions or barriers such as an individual's hukou and explicit migration costs, which may impede the mobility of workers out of agriculture and affect sorting. These individual-level mobility frictions are distinct from the perceived land insecurity at the household level, which emanates directly from the household's participation in the land rental market.

Our framework gives rise to rich patterns of selection, within households and across occupations, which have implications for agricultural and aggregate productivity and structural change. We quantitatively evaluate the role of land insecurity and other labor mobility frictions for these outcomes, exploiting our micro-level data on individuals and households for China between 2004–2018. The data provide detailed information on labor supply, farm production, and incomes by occupation to which we calibrate idiosyncratic abilities and distortions in our model.

In particular, the model allows us to identify separately land insecurity from labor mobility frictions by exploiting the fact that land insecurity is primarily a family-level friction, manifesting through the family decision of whether to operate a farm or not, whereas the labor-mobility barrier operates at the individual level. As a result, land insecurity is pinned down empirically from the share of village families operating farms, while labor mobility barriers are linked to the sectoral income gap between agriculture and non-agriculture. We directly measure the risk of land loss tied to land rental from supplementary survey data we collected in China, expressly for the purpose of measuring land insecurity. We draw on data on the observed historical frequency of village-wide land reallocations and households' responses to the perceived likelihood they lose access to the land in such an event.

We find that in 2004 labor mobility frictions associated with land insecurity are high in levels and large in dispersion and responsible for an inefficiently large number of farms in agricultural production. Moreover, the effects of land insecurity on employment and agricultural labor productivity are at least as large in magnitude to the residual labor-mobility frictions. From the perspective of 2004, introducing secure land rights substantially reduces the share of households operating farms from 74 percent to 28 percent and lowers the share of employment in agriculture among village households from 56 to 46 percent, resulting in an improvement in agricultural productivity of 25 percent. By comparison, while removing residual labor-mobility frictions has a similar effect on the share of labor in agriculture, it reduces the percentage of village households farming to 49 percent and increases agricultural productivity by 16 percent.

Over the period 2004-2018 we find a marked improvement in land security but a slight increase in other labor mobility frictions. Our analysis highlights that much of the increase in labor mobility over the period is tied to improvements in land tenure security, consistent with policy reforms pursued in China since the early 2000s. We also find an important interaction between land insecurity and misallocation within agriculture as the productivity gain from an efficient reallocation of resources within agriculture more than doubles when land is secure. These additional gains arise from improved selection in farming within households and improved selection across sectors.

We extend our main framework along three important dimensions. First, we consider a local non-agricultural sector in the rural area, motivated by the striking regularity in the micro data for China that among those with rural hukou working in non-agriculture, more than half are employed locally, with the rest migrating to the cities to work. Adding this local dimension allows the framework to separate sectoral and spatial reallocation. Second, we also consider age differences among family members, motivated by the empirical regularity that farms are typically operated by older individuals within the household, with younger members more likely to seek employment opportunities outside farming. In 2004, individuals over the age of 45 represented 40.5 percent of the labor force in the countryside, but were operators of 57.8 percent of all farms. Third, we examine regional heterogeneity in the extent of land insecurity and labor mobility barriers, motivated by geographic differences in proximity to urban areas.

Overall, we find that our conclusions about the role of land insecurity and other labor mobility frictions for aggregate outcomes are robust to these extensions. The extended framework offers richer micro-level implications. For those with rural hukou, we find it is much less costly to work part-time in rural non-agriculture than in urban non-agriculture as rural non-agriculture does not involve spatial reallocation. Nevertheless, labor mobility barriers are actually higher for the rural non-agricultural sector. We also find that older workers face substantially higher overall barriers to labor mobility compared to younger workers, which helps explain why they are more likely to operate farms. Our analysis also suggests important village-level differences in labor mobility in China, with peri-urban areas facing less severe land insecurity and lower local labor mobility barriers than more remote areas.

Our paper relates to several strands of literature. In highlighting the role of farming for labor supply choices our work relates to the literature on structural transformation and agriculture (Gollin et al., 2002; Restuccia et al., 2008), and the agricultural productivity gap (Gollin et al., 2014a). The paper is also connected to the agricultural productivity literature emphasizing misallocation (Adamopoulos and Restuccia, 2014), particularly that relating to land market institutions (Chen, 2017; Gottlieb and Grobovšek, 2019; Adamopoulos and Restuccia, 2020; Chari et al., 2021; Adamopoulos et al., 2022), and sectoral selection (Lagakos and Waugh, 2013; Alvarez, 2020; Hamory et al., 2021; Adamopoulos et al., 2022). Equally important is the literature on migration costs and structural transformation (Morten and Oliveira, 2018; Bryan and Morten, 2019; Lagakos et al., 2020; Schoellman, 2020; Hamory et al., 2021).[1] While these papers focus on estimating the magnitude of migration costs using reduced-form or structural approaches, we aim to disentangle the implicit migration cost arising from insecure land rights from other labor mobility barriers. Our paper is also broadly related to the literature on growth and development in the context of China (Song et al., 2011; Brandt et al., 2013, 2020, among others) and the literature on institutions as an obstacle to development in poor countries (Acemoglu et al., 2001, 2005).

The two papers most closely related to ours are Ngai et al. (2019) and Adamopoulos et al. (2022). Ngai et al. (2019) highlight that land insecurity can manifest as a labor mobility barrier out of agriculture in China. The value added of our paper is a quantitative assessment of the effect of land insecurity versus a catchall of other labor mobility barriers in a rich, empirically-grounded model that nests individual and family selection. The rare

[1]The literature has studied several factors that could deter labor mobility, including transportation infrastructure (Asher and Novosad, 2020); land use rights (Field, 2007; de Janvry et al., 2015); rural insurance networks (Munshi and Rosenzweig, 2016); monetary cost and risk (Bryan et al., 2014); administrative registration and services (Chan and Zhang, 1999); housing (Brueckner and Lall, 2015); and regional variation in industrialization (Eckert and Peters, 2023).

individual and household micro-level data we have for China facilitate this investigation, the identification of the two types of barriers, and the quantification of selection. Quantitatively, selection is the largest component of these effects. Adamopoulos et al. (2022) highlight the effect of China's land institutions for misallocation within villages in agriculture and sectoral selection using household-level data for an earlier period. In contrast, we use individual- and household-level data to estimate idiosyncratic mobility barriers and study labor allocation and migration decisions within households, thus allowing us to examine misallocation within households and across space. Relative to the previous literature, we also model differences between young and old adults in their ability and mobility barriers to capture structural transformation by age (Hobijn et al., 2018; Porzio et al., 2022).

The paper proceeds as follows. The next section describes the land and labor mobility institutions in China. Section 3 documents key regularities on China's structural transformation that motivate our framework of analysis. In Section 4, we present the main model and section 5 estimates the model by matching moments from micro and aggregate data for 2004. Section 6 performs quantitative experiments in order to assess the relevance of land security and other mobility frictions and their evolution over time. We also evaluate the interaction of land insecurity and farm-level distortions associated with China's land institution. In section 7, we extend the main framework by adding spatial choices in non-agricultural employment, age differences, and regional village heterogeneity. We conclude in section 8.

2 Institutional Background

We provide institutional context for land and labor mobility policies in China since our framework aims to decompose the costs associated with out-migration of workers from agriculture into those related to land market institutions and other labor mobility frictions.

Land market institutions. With the introduction of the Household Responsibility System in the early 1980s, rural households were granted use rights and residual income rights over land. Farmland was not privatized however, and ownership continued to reside in the collective, i.e. village. The allocation of village land was highly egalitarian, and tied to "membership" in the village through the household registration system. As a result, per capita differences between households in landholdings within a village were small.

Initially, use rights were extended to households for 15 years. Through the late 1990s however local officials often carried out village-wide reallocations of land across households to accommodate demographic changes (Benjamin and Brandt, 2002). Rent-seeking behavior on the part of village leaders was also likely a factor (Kung and Liu, 1997; Brandt et al., 2002). In addition, rural land was occasionally expropriated from households by the state for non-agricultural uses including highway construction, urban development, and industrial parks.

In most villages, land rental was not officially restricted, however rental markets were thin, with less than 3 percent of total land rented out by the mid-1990s (Brandt et al., 2002). "Use-it-or-lose-it rules" were likely partially responsible. Village leaders often regarded rental transactions as a signal of land misallocation, and viewed village land reallocations as an opportunity to redirect farmland to other households. In the case of land expropriation by the state, households who rented land out risked not receiving compensation for the land taken. Since rentals invited dispossession, households were unwilling to rent out their land.

This incomplete set of property rights over land had two important implications. First, land was misallocated within villages, a product of the highly egalitarian distribution of use rights and limited market-based exchange through rental. And second, fear of the loss of land use rights (or compensation in the event of expropriation by the state) prompted households to assign family members to farm the land to protect their rights.

A series of reforms to land policy in China beginning in the late 1990s helped to strengthen household property rights, but did not privatize farmland. In 1998, the Land Management

Law (LML) extended the use rights to rural households for 30 years. In 2003, the Rural Land Contracting Law (RLCL) codified legal rights of rural households for leasing out agricultural land. Village-wide reallocations fell considerably, as we document in Appendix A.5. Estimates suggest that these reforms led to a 7 percent increase in land rentals, and the reallocation of land towards higher productivity households (Chari et al., 2021). In 2013, further reform of land policy allowed farmers to transfer their land use rights to others, including large commercial farms. In 2014, the right to mortgage the use of land was also extended to rural households (Zhou et al., 2021). Finally, in 2018, a formal land titling/certification process was begun that by 2021 was largely complete.

Labor mobility institutions. Labor mobility in China has been heavily influenced by the Household Registration or *Hukou* System (Chan, 2019). Established in the 1950s, the primary purpose of the system was to control migration between the countryside and the cities. Under the hukou system, each individual is assigned agricultural or non-agricultural hukou within a locality, e.g. village, town or city, which determines where they can live and work. It also determines their rights and access to local public services. Major differences exist between rural and urban areas in the level and availability of these services, as well as between cities since larger cities typically confer much better benefits than smaller cities.

Through the late 1970s there were very tight restrictions on migration, as well as on the type of non-agricultural activities that were permitted in the countryside. Restrictions on the type of activity began to be relaxed in the early 1980s, followed by an easing of restrictions on rural hukou holders from moving to the cities to work (de Brauw and Giles, 2018). Rural migrants continued to be restricted however in terms of the public services they could access once they moved. Data from China's population census reveal that migrants' share of non-agricultural employment rose from 6.4 in 1990 to 17.7 percent 2000, and then to 21.0 percent in 2005.

In the early 2000s, new reforms made it easier for migrant workers to obtain working permits in other provinces, with some coastal provinces totally eliminating the requirement (Sun et al., 2011; Kinnan et al., 2018). In 2014, the distinction between agricultural and non-agricultural hukou in the same location was eliminated, thereby entitling, on paper, all residents to the same set of local public services. Chan (2019) argues however that this reform was primarily limited to smaller cities, and that it actually became harder to obtain local hukou in larger coastal cities, which restricted in-migration and even tried to force migrants out by limiting migrants' children access to education. Rising housing prices in the cities likely reinforced the effects of weak enforcement of reforms, and contributed to rising migration costs for rural families.

Data from the Population Census for 2005, 2010 and 2015 reveal a slowing in labor flows, with migrants' share of non-agricultural employment rising to 26.5 percent in 2010, but to only 28.3 percent by 2015. Consistent with these trends, Wu and You (2020) find that the probability of migrants obtaining local hukou also experienced a pronounced reduction over the same period. For individuals that had completed high school, for example, the probability of receiving urban hukou within 5 years fell from 16.3 and 38.6 percent in tier 1 and 2 cities in 2000, respectively, to only 3.9 and 10.5 percent in 2010.

3 Stylized Facts on China's Structural Transformation

The Chinese economy has experienced substantial structural change that has been accompanied by a marked reduction in the share of the labor force engaged in agriculture. Between 1995 and 2018, the share dropped from 48.0 percent to 13.7 percent. This decline reflects two forces: The increase in the percentage of households living and working in the cities, and changes in the employment patterns of households registered in the countryside.

We draw on the nationally-representative rural household survey data collected by the Research Center for the Rural Economy (RCRE) under the Ministry of Agriculture of China to describe salient features of the changes in the countryside that motivate our framework and analysis. For a detailed description of the earlier data see Benjamin et al. (2005). We focus on the period between 2004 and 2018, leveraging the individual-level labor supply data that were added to the survey in 2003. Our unbalanced panel contains information on over 20,000 households per year drawn from 300 villages throughout China. Appendix A provides a detailed description of the data.

Agricultural employment. The first row of Table 1 reports the falling share of China's total labor force employed in agriculture between 2004 and 2018. This decline reflects two forces: The reduction in the share of households registered in rural areas, which fell from 69.1 percent in 2004 to 39.1 percent in 2018, and the reallocation of labor from agriculture to non-agriculture among rural residents.[2] Among those registered in the countryside, the share working in agriculture fell from 69.8 to 44.2 percent. The largest reductions occurred among full-time farm operators and part-time workers, which fell 10.9 and 8.7 percentage points, respectively. Other household members working full time in agriculture fell 6 percentage points. A simple decomposition suggests that the reallocation of labor among rural residents from agriculture to non-agriculture was the source of half of the total reduction in the nationwide share of employment in agriculture.

In 2004, 50.9 percent of rural individuals worked in non-agriculture, of which 30.2 percent were full-time, and 20.7 percent were part-time (see the last two rows of Table 1). Between 2004 and 2018, the percentage of individuals working in non-agriculture increased from 50.9 percent to 67.8 percent of the rural labor force. All of this increase is a product of the growth in full-time employment in non-agriculture, which grew from 30.2 percent to 55.8 percent. The share working part-time fell from 20.7 percent to 12.0 percent.

Average farm size and land rentals. In Table 2 we report for select years average farm size of farming households in the RCRE survey, measured as total cultivated area divided

[2] In each year, the share of nationwide employment in agriculture is the product of the share of households living in rural areas and the share of labor among these households in agriculture.

Table 1: Employment in Agriculture

Variable	2004	2009	2013	2018
Nationwide agricultural employment share (%)	39.1	28.4	20.9	13.7
Share of all households living in rural area (%)	69.1	57.2	46.4	39.1
Share of labor days in agriculture among rural households (%)	56.5	49.6	45.1	35.1
Share of individuals in rural areas involved in agriculture (%)	69.8	62.0	56.7	44.2
Share of farm operators (%)	27.6	23.8	21.9	16.7
Share of full-time workers (%)	21.5	20.3	18.7	15.5
Share of part-time workers (%)	20.7	17.9	16.1	12.0
full-time non-agriculture (%)	30.2	38.0	43.3	55.8

Notes: Nationwide statistics adjusted from the Chinese Statistics Yearbook. All other statistics are calculated from micro-level data from the RCRE survey, as shares out of the total numbers of employed individuals in the survey. For each year the shares of farm operators, full-time and part-time workers sum up to the total share in agriculture, save for rounding.

by the number of farming households. Average farm size was only 0.59 hectares in 2004, and increased by more than fifty percent by 2018, but still remained below a hectare.

Table 2: Average Farm Size and Land Rentals

Variable	2004	2009	2013	2018
Average farm size (ha)	0.59	0.62	0.71	0.91
Share of households renting-in land (%)	17.9	20.0	17.9	18.7
Share of land rented-in (%)	14.4	22.4	18.6	31.8

Notes: Average farm size is the total cultivated land over the total number of operating farms (households) from the RCRE survey data and expressed in hectares (ha). Statistics on land rentals are calculated among farming households in the Survey each year. Data are from authors' calculation from the micro data sample.

Underlying the increase in average farm size is an increase in rentals of farm land. In Table 2 we also report the share of farming households that rent in land (second row), which can be thought of as the extensive margin of renting, and the share of land rented in among farming households, or the intensive margin. Although the share of rural households renting in land increased slightlly, the share of land rented in increased more, from 14.4 percent in 2004 to 31.8 percent by 2018. Hence, the expansion in land rental is primarily driven by the intensive margin and increases in the average amount of land rented in by households in the land rental market. The increase in land market activity is consistent with the alleviation of frictions in land markets through reforms that improved land security in China (see, for example, Chari et al., 2021).

Earlier period 1995-2002. We draw on earlier waves of the RCRE survey to put structural change in the Chinese economy between 2004-2018 into perspective. Prior to 2004, there are signs of structural change however the process was much slower (see Table 3). Between 1995 and 2002, for example, the share of employment in agriculture at the national level fell from 48.0 percent to 42.2 percent, or one-half the annual rate that we observe after. This reflects only modest changes in the share of households living in rural areas, and the allocation of labor by these households to agriculture. Average farm size also remained more or less constant.

Table 3: Structural Transformation before 2004

Variable	1995	2002
Nationwide agricultural employment share (%)	48.0	42.2
Share of households living in rural area (%)	75.4	72.6
Share of labor days in agriculture among rural households (%)	63.6	58.2
Share of individuals in rural areas involved in agriculture (%)	73.6	67.5
Share of full-time workers (%)	53.6	48.8
Share of part-time workers (%)	19.9	18.8
Average farm size (ha)	0.57	0.59

Notes: Nationwide statistics adjusted from the Chinese Statistics Yearbook. Employment statistics are calculated from micro-level data from the RCRE survey for 1995 and 2002, as shares out of the total numbers of employed individuals in the survey. Average farm size is the total amount of farm land over the total number of operating farms (households). The numbers are not directly comparable to those in the Survey over 2004-2018.

These features of China's structural transformation, which are prevalent in other lower income and emerging economies, motivate the framework we develop in the next section.

4 A Model of Frictional Selection

To study the role of insecure property rights and other labor mobility frictions on reallocation and aggregate outcomes, we develop a general equilibrium model of occupational and sectoral choices of heterogeneous individuals subject to idiosyncratic distortions. Our Roy-type model of occupational choice has four key novelties. First, we consider families that make an endogenous decision of who within the family operates the family farm. Second, we introduce family-level insecure property rights over rented farm land. Third, we consider residual labor-mobility barriers to capture all factors other than land insecurity that can impede individual labor mobility. And fourth, in contrast to the standard Roy model, where occupational choices are binary, we allow for part-time employment. We extend this framework in Section 7 to allow for spatial choices in non-agricultural employment, age differences within families, and regional village heterogeneity.

4.1 Description

The population is organized into families living in villages and urban centres. At each date, two goods are produced: an agricultural good (a) and a non-agricultural good (n). In our main specification agricultural production is limited to villages while production of the non-agricultural good occurs in urban centres. Later, we extend the model to include rural non-agricultural production. Village households can work locally in agriculture or migrate to the cities to work in the non-agricultural sector. Urban households work in the non-agricultural sector. We focus on the factors influencing the choice of occupations and locations of individuals and families living in villages.

Preferences and endowments. All households have preferences over the agricultural and non-agricultural goods represented by the non-homothetic utility function:

$$u(c_a, c_n) = \phi \log(c_a - \bar{a}) + (1 - \phi) \log(c_n), \quad \phi \in (0, 1), \tag{1}$$

where c_a is consumption of the agricultural good, $\bar{a}$ is a minimum (subsistence) consumption requirement of the agricultural good, and c_n is consumption of the non-agricultural good. This specification of preferences allows us to capture income effects as a source of structural change. Our results are robust to using alternative specifications, such as the price independent generalized linear (PIGL) preferences (Boppart, 2014), with results reported in Appendix G.

There are two types of households: village and urban. Urban families are homogeneous and make no production decisions. Each is endowed with one unit of labor of average ability $\bar{h}$ and is engaged in non-agricultural production. There is an exogenous mass N_n of urban families.

Village families are at the core of all decisions relating to agricultural production, occupational choice, and migration to the city. Village families can be engaged either in agriculture or non-agriculture. There is a unit measure of village families indexed by i, comprising J individuals per family, indexed by $j = 1, 2, ...J$. Each individual family member is endowed with a pair of abilities (s_{ij}, h_{ij}), where s_{ij} is their ability as a farm operator and h_{ij} is the earning ability in non-agriculture. Families are subject to insecure property rights over rented farmland while individual family members are subject to idiosyncratic labor mobility frictions. Individual-level abilities and idiosyncratic frictions are drawn from a joint multivariate distribution with a cumulative distribution function Φ.

Since there is a unit measure of village families, the total number of individuals in villages is J, and the total amount of farm land is L which we normalize to one without loss of generality. Use rights over farmland are allocated across all village families on an egalitarian basis, with each family endowed with an amount of land $\bar{\ell}$. Individuals are endowed with one unit of time in each period that is supplied inelastically, and can work as a full-time farm operator or as an agricultural or non-agricultural worker on a full-time or part-time basis.

Technologies. The non-agricultural good is produced using a constant returns to scale technology,

$$Y_n = A_n H_n, \tag{2}$$

where A_n is the productivity parameter of the non-agricultural sector and H_n is the total labor input in efficiency units.

The agricultural good is produced by family farms. A farm is a decreasing-returns technology that requires the inputs of a farm operator with managerial ability s_{ij}, cultivated land l_{ij}, and farm labor n_{ij}:

$$y_{ij} = A_a s_{ij} \left(l_{ij}^{\theta} n_{ij}^{1-\theta} \right)^{\gamma},$$

where y_{ij} is agricultural output, A_a is an agricultural technology parameter, $\theta \in (0, 1)$ captures the importance of land relative to labor in production, and $0 < \gamma < 1$ captures returns to scale at the farm-level. Note that farm productivity is determined by the ability of the family member that is chosen to operate the farm s_{ij}, which in turn governs the input choices of the farm.

Land rental market. There is a potentially active market for rentals of farmland that allows a farmer to adjust the scale of their farm operation relative to their endowed use rights $\overline{\ell}$. Their land input is then $\ell_{ij} = \overline{\ell} + \ell_{ij}^{\text{rent}}$, with $\ell_{ij}^{\text{rent}} > 0$ for farmers that rent-in land and $\ell_{ij}^{\text{rent}} < 0$ for those that rent-out land. Let q denote the relative rental price of land in the market.

Other markets. A representative firm operates the non-agricultural technology in a competitive market. We set the non-agricultural good as the numeraire and normalize its price to one. We denote by w_n the competitive price of efficiency units of labor in the non-agricultural sector. Profit maximization implies that the wage equates the marginal product of efficiency units of labor,

$$w_n = A_n.$$

In addition to land, farms also operate in competitive markets of labor and output, subject to frictions that we describe below. We denote by p_a the relative price of the agricultural good, and by w_a the wage rate in agriculture, which can differ from the wage rate in the non-agricultural sector.

4.2 Idiosyncratic Distortions

Families and individuals are subject to idiosyncratic land and labor mobility distortions.

Insecure land rights (family-level). We explicitly model the insecurity in property rights over farmland that families in China face. In principle, land rentals are not prohibited, however families that rent out land risk losing access to this land in the event of an administrative land reallocation. Let η be the probability that households lose the use rights to land they rent out, without compensation. These households suffer a family-specific income loss equal to φ_i per unit of land rented out. Differences in φ_i across families reflect differences in the valuation of the land tied to observable and unobservable characteristics of the land and farming family. Overall, the product $\eta\varphi_i$ captures the family's perceived income loss associated with land insecurity from renting-out land.

Labor mobility frictions (individual-level). Individual labor supply choices outside agriculture are subject to idiosyncratic distortions, modeled as barriers on wage income. An individual j in family i faces a wage-barrier ξ_{ij} if they work in the non-agricultural sector. This barrier is a catchall for all the factors that may impede the reallocation of labor to the non-agricultural sector, exclusive of the land institutions. Note that the land market frictions above are linked directly to a family's participation in the land rental market, while the labor mobility friction is linked directly to migration to the city.

Residual distortions in agriculture Following Adamopoulos et al. (2022), we assume that operating a family farm is subject to an implicit distortion on output τ_{ij} imposed on individual j in family i. This distortion encompasses the residual frictions in agricultural production that cannot be directly captured by the land loss risk described above, and has the property that more productive individuals cannot necessarily access more land as in unfettered land markets. Quantitatively, this distortion captures residual misallocation, over and above that implied by land insecurity, and helps reconcile the observed land allocations in the data.

4.3 Village Families

We describe the occupational choices of village families in two steps. First, we characterize the income of each individual j in family i from working in every possible occupation: farm operator, agricultural worker, and non-agricultural worker. Second, we characterize family i's endogenous allocation of individual members to occupations to maximize total family income.

Farm operator. To operate a farm, each family i must assign one individual to be a full-time farm operator in order to produce the agricultural good and retain their allocated land. The productivity s_{ij} and output distortion τ_{ij} of the family farm are those associated with farm operator j. The operator decides how much land to rent in or out, ℓ_{ij}^{rent}, relative to their endowment $\bar{\ell}$, labor demand, and output supply.

These farm choices however are subject to the land insecurity risk that farming families face when renting out land. Specifically, after agricultural production, there is an exogenous probability η that they lose the rented-out land, in which case there is an associated income loss φ_i that is proportional to the amount of rented-out land. The profit maximization problem for family farm i is then given by,

$$\pi(s_{ij}, \tau_{ij}) = \max_{n_{ij}^d, \ell_{ij}^{\text{rent}}} \left\{ \tau_{ij} p_a A_a s_{ij} \left[\left(\bar{\ell} + \ell_{ij}^{\text{rent}}\right)^{\theta} {n_{ij}^d}^{1-\theta} \right]^{\gamma} - q\ell_{ij}^{\text{rent}} - w_a n_{ij}^d + \eta \varphi_i \ell_{ij}^{\text{rent}} \mathbb{1}[\ell_{ij}^{\text{rent}} < 0]) \right\}, \tag{3}$$

where n_{ij}^{d} is farm labor input and $\mathbb{1}[\ell_{ij}^{\text{rent}} < 0]$ is an indicator variable that takes the value of 1 if the family rents out part of their allocated land. Overall, the last term captures the perceived income loss associated with land insecurity for farms that rent-out land.

Full-time non-agricultural worker. An individual j in family i working full-time in the non-agricultural sector, facing labor mobility barrier ξ_{ij}, earns net income $i^{FN} = w_n h_{ij}(1 - \xi_{ij})$ where the superscript FN denotes "full-time non-agriculture."

Full-time agricultural worker. If an individual j, from family i, works as a full-time agricultural worker, they earn wage income. The skills of agricultural workers are homogeneous across all individuals, and normalized to 1. We assume that family labor and hired labor are perfect substitutes; hence, all agricultural labor is paid at rate w_a. We denote the income from full-time agricultural work as $i_{ij}^{FA} = w_a$, where the superscript FA denotes "full-time agriculture."

Part-time worker. In the data workers often earn wages in both agriculture and non-agriculture. To account for this, we allow individuals to work as wage earners part-time in each sector. We denote by n_{ij} the labor supply of individual j from family i to agriculture. The labor supply of the same individual to non-agriculture is $(1 - c - n_{ij})$, where c denotes a fixed cost associated with working part-time. An individual that works part time in both agriculture and non-agriculture earns income $h_{ij} w_n (1 - \xi_{ij})(1 - c - n_{ij})$ from non-agriculture and $w_a \kappa n_{ij}^{\nu}$ from agriculture, where κ measures the relative efficiency of labor input in the two sectors. The parameter $\nu < 1$ captures decreasing returns of labor supply to the agricultural sector. The concavity allows for incomplete specialization across sectors for part-time workers (Erosa et al., 2022).

Taking the first-order condition of the total income of a part-time worker with respect to labor supply in agriculture n_{ij}, we obtain the optimal agricultural labor supply

$$n_{ij}^* = \min\Big\{\Big(\frac{\nu\kappa w_a}{h_{ij}w_n(1-\xi_{ij})}\Big)^{\frac{1}{1-\nu}}, 1-c\Big\}.$$

The corresponding income is then given by

$$i_{ij}^{PT} = h_{ij}w_n(1-c-n_{ij}^*)(1-\xi_{ij}) + w_a\kappa(n_{ij}^*)^{\nu},$$

where the superscript PT denotes "part-time worker."

Using the above notation and characterization, we can summarize effective labor supply to the agricultural sector as:

$$n_{ij} = \begin{cases} 1, \text{ if } i_{ij}^{FA} \geqslant i_{ij}^{FN} \text{ and } i_{ij}^{FA} \geqslant i_{ij}^{PT}; \\ 0, \text{ if } i_{ij}^{FN} \geqslant i_{ij}^{FA} \text{ and } i_{ij}^{FN} \geqslant i_{ij}^{PT}, \\ \kappa(n_{ij}^*)^{\nu}, \text{ if } i_{ij}^{PT} \geqslant i_{ij}^{FA} \text{ and } i_{ij}^{PT} \geqslant i_{ij}^{FN}. \end{cases} \tag{4}$$

Family allocation problem. We now characterize how the household allocates its members across occupations. If individual j is chosen to be the farm operator, total income is given by,

$$I_i^a(\text{operator} = j) = \pi(s_{ij}, \tau_{ij}) + \sum_{k \neq j} \max\{i_{ik}^{FN}, i_{ik}^{FA}, i_{ik}^{PT}\}. \tag{5}$$

Operating in agriculture, the household chooses the farm operator that maximizes total income,

$$I_i^a = \max_{j \in J}\{I_i(\text{operator} = j)\}. \tag{6}$$

If the household chooses not to participate in agricultural production, then no family member is assigned to be a farm operator. The household rents out all of their allocated land $\overline{\ell}$, which is then subject to the risk of income loss. In this case, total household income, accounting for land insecurity, is given by

$$I_i^n = \sum_j \max\{i_{ij}^{FN}, i_{ij}^{FA}, i_{ij}^{PT}\} + (1-\eta)q\bar{\ell} - \eta\varphi_i\bar{\ell}, \tag{7}$$

where $(1-\eta)q\bar{\ell}$ is the rental income without reallocation, and $-\eta\varphi_i\bar{\ell}$ is the perceived income loss associated with land insecurity. The household chooses to operate a farm if and only if $I_i^a \geqslant I_i^n$. Hence, total income of village family i is

$$I_i = \max\{I_i^a, I_i^n\}. \tag{8}$$

Role of land insecurity. Insecure property rights over land affect rural household decisions with respect to farming on both the intensive and extensive margins. On the intensive margin (see equation 3), they discourage farm operators from renting out more land. On the extensive margin (see equations 7 and 8), they deter families from completely abandoning farming, out of fear of losing access to this land. These effects have two critical implications. First, they imply that land insecurity leads to misallocation of land across farms. Due to limited land rentals, the operational scale of farms is more closely tied to a household's administratively allocated use rights. As a result, more productive farmers are constrained in accessing more land. Second, land insecurity acts as a deterrent to migration, resulting in a de-facto barrier to labor mobility out of agriculture.

4.4 Equilibrium

We define an indicator variable D_i for family i, with $D_i = 1$ if family i operates a farm in agriculture and $D_i = 0$ otherwise. An individual is one of the following four occupational types: farm operator, full-time agricultural worker, full-time non-agricultural worker, and part-time agricultural and non-agricultural worker. For individual members, we use the following notation to represent optimal occupation choices. We define indicator variables, with $D_{ij}^O = 1$ if individual j is the farm operator and $D_{ij}^O = 0$ otherwise. We denote by $D_{ij}^{FA} = 1$ if individual j is a full-time agricultural worker, $D_{ij}^{FN} = 1$ if a full-time non-agricultural worker, and $D_{ij}^{PT} = 1$ if a part-time worker.

The net rented-in land among households who operate farms is $\int_i D_i \sum_{j\in i} D_{ij}^O(\ell_{ij}^{\text{rent}})di$, where $D_i D_{ij}^O$ is a product of indicators that household i operates a farm, and individual j in household i is the farm operator. Land rental market clearing requires that,

$$\int_i D_i \sum_{j\in i} D_{ij}^O \ell_{ij}^{\text{rent}} di = \int_i (1 - D_i)\bar{\ell} di, \tag{9}$$

where the right-hand term captures the land rented out by households who do not operate farms.

Agricultural labor supply is $\int_i \sum_{j\in i}(1 - D_i D_{ij}^O) n_{ij} di$, where we sum up labor supplied to agriculture for each household member if they are not farm operators. Agricultural labor market clearing then requires

$$\int_i \sum_{j\in i} (1 - D_i D_{ij}^O) n_{ij} di = \int_i D_i \sum_{j\in i} n_{ij}^d D_{ij}^O di. \tag{10}$$

Non-agricultural labor supply consists of two components: efficiency units of labor supplied by urban households $N_n \bar{h}$, and those supplied by rural households. The rural household labor supply to non-agriculture is the sum of that of full-time non-agricultural workers and that from part-time workers,

$$H_n = N_n \bar{h} + \int_i \sum_{j\in i} (1 - D_i D_{ij}^O) \left[(1 - n_{ij} - c) D_{ij}^{PT} + D_{ij}^{FN}\right] h_{ij} di. \tag{11}$$

Denote the agricultural consumption as c_{ai} for rural household i and c_a^n for urban household. We can then write the agricultural good market clearing condition as,

$$\int_i c_{ai} di + N_n c_a^n = \int_i \sum_{j\in i} D_i D_{ij}^O y_{ij} di, \tag{12}$$

where y_{ij} represents farm output. Similarly, the non-agricultural good market clearing is,

$$\int_i c_{ni} di + N_n c_n^n = A_n H_n. \tag{13}$$

Net revenues from the income loss of land rentals, labor mobility barriers, and farm output distortions are rebated back to households uniformly.

With these equilibrium conditions, we now define the competitive equilibrium of this economy.

Definition. *A competitive equilibrium is a set of prices* $\{w_a, w_n, p_a, q\}$*, allocations of the farm operators* ℓ_{ij}^{rent}*,* n_{ij}^d*,* y_{ij}*,* $\pi_{ij}(s_{ij}, \tau_{ij})$*; village family incomes* I_i*, occupation allocations* D_i*,* D_{ij}^O*,* D_{ij}^{FA}*,* D_{ij}^{FN}*,* D_{ij}^{PT}*, and* n_{ij}*, consumption of village families* (c_{ai}, c_{ni})*, and consumption for the urban households* (c_a^u, c_n^u)*, such that:*

1. *Given prices, farm operators maximize profits* π_{ij} *and* $\bar{\ell} + \ell_{ij}^{rent}$*,* n_{ij}^d *are optimal factor demands.*

2. *Given prices, non-agricultural firms maximize profits.*

3. *Given prices, village families maximize income* I_i *by choosing labor supply* n_{ij} *from equation (4) and occupations* D_i*,* D_{ij}^O*,* D_{ij}^{FA}*,* D_{ij}^{FN}*,* D_{ij}^{PT} *to solve problems (6), (7), and (8).*

4. *Given prices and income* I_i*, village households choose consumption* (c_{ai}, c_{ni}) *to maximize utility in equation (1).*

5. *Given prices, urban households choose consumption* (c_a^n, c_n^n) *to maximize utility subject to income* $w_n \bar{h}$*.*

6. *All markets clear; i.e., prices solve equations (9), (10), (11), (12), and (13).*

5 Estimation

To estimate the model we use two main data sources: (*i*) household- and individual-level survey data from the Research Center for the Rural Economy (RCRE) for the period 2004-2018, which provides detailed individual labor supply and income data, as well as farm-level inputs and outputs; and (*ii*) aggregate data from the Chinese Statistics Yearbook supplemented with revised sectoral employment series from Brandt and Zhu (2010) and Yao and Zhu (2021). We also draw on a supplementary survey undertaken with the RCRE, which covers a more limited number of villages but surveys farmers about their perceptions of land insecurity. A more detailed description of the data and variable definitions are provided in Appendix A.

Our estimation strategy involves two steps. First, we parameterize the distributions for sectoral abilities, land market institutions, and labor mobility frictions of village families. Second, we estimate the parameters of the model to match model moments with empirical moments from the survey and aggregate data. We also discuss our key identification restrictions.

5.1 Parameterization

The parametric distributions that we need to specify for individuals in the model are: farm-operating ability, non-agricultural ability, labor mobility frictions, and residual farm-operating distortions. We also need to specify the distribution for the family-level land insecurity cost.

Ability distributions. Each individual j, in family i, draws a farm-operator ability s_{ij} and a non-agricultural ability h_{ij}. We allow for individual abilities to be correlated across members of the same family. In addition, we allow for abilities to be correlated across sectors.

In particular, the non-agricultural ability of individual j in family i is given by,

$$\log(h_{ij}) = \log(h_i^H) + \log(h_{ij}^I),$$

where h_i^H is a common component for all individuals within family i and h_{ij}^I is individual j's idiosyncratic component. Without taking a stance on the source, the common component can capture, for example, the correlation of innate ability or accumulable skills across family members. The agricultural ability of individual j, in family i, is given by,

$$\log(s_{ij}) = \log(s_i^H) + \log(s_{ij}^I) + \lambda \log(h_{ij}),$$

where s_i^H is the common family component, s_{ij}^I the individual component, and $\lambda \log(h_{ij})$ is the component that is correlated with non-agricultural ability. Note that λ governs the correlation between agricultural and non-agricultural abilities of the same individual. We assume that ability components s_i^H, h_i^H, s_i^I, and h_i^I are all drawn from mean-zero log-normal distributions. The standard deviations of the family components are σ_s, σ_h, respectively, and that of the individual components are $\omega\sigma_s$, $\omega\sigma_h$, respectively, where ω governs the relative importance of individual components versus family components.

Individual-level labor mobility frictions. We assume that labor mobility barriers facing rural individuals working in the non-agricultural sectors are given by,

$$\xi_{ij} = \frac{\exp(\mu_\xi + \varepsilon_{ij}^\xi)}{1 + \exp(\mu_\xi + \varepsilon_{ij}^\xi)}. \tag{14}$$

This parametric form guarantees that labor mobility barriers are bounded between zero and one, where μ_ξ governs the level of barriers and ε_{ij}^ξ their idiosyncratic component, which we assume follows a normal distribution with mean zero and standard deviation σ_ξ.

Residual idiosyncratic farm distortions. Individual j in family i faces implicit distortions to operating a farm, measured by τ_{ij}. We allow for the farm distortion to feature a component that is correlated with farm-operating ability and a component that is uncorrelated, similar to how idiosyncratic wedges are modeled in Bento and Restuccia (2017) and Restuccia (2019),

$$\log(\tau_{ij}) = \zeta \log(s_{ij}) + \varepsilon^{\tau}_{ij},$$

where ζ captures the elasticity of distortions with respect to farming ability s_{ij}, and the uncorrelated component ε^{τ}_{ij} is drawn from a normal distribution with mean zero and standard deviation σ_{τ}.

Family-level land insecurity. In the model, η is the probability that a household loses the use rights to land that it has rented out. Conditional on losing their land, φ_i summarizes the perceived income loss per unit of rented-out land associated with the insecure land rights. We assume that $\log(\varphi_i)$ differs across families and is drawn from a normal distribution with mean μ_{φ} and standard deviation σ_{φ}.

5.2 Identification

There are 23 parameters to be determined in our model, including the parameters of the ability distributions and distortions discussed above. We provide a brief overview of our calibration approach, focusing on the elements that are non-standard and in Appendix B we provide a detailed explanation of how the parameters are assigned values or estimated.

Out of the 23 parameters, 9 are either normalized or set based on a priori information. The remaining 14 parameters are estimated jointly so that model-generated moments for key variables match empirical counterparts. While the parameters are jointly determined, some moments are more relevant for identifying key parameters. For instance, the distribution of employment across sectors and occupations helps identify the subsistence constraint in agriculture $\bar{a}$, the fixed costs to part-time work in non-agriculture (c), and the relative sectoral efficiency parameters (κ, $\bar{h}$).

Ability distributions. The moments of the ability distributions are chosen to match empirical moments of dispersion in sectoral incomes (non-agriculture) and farm TFP (agriculture). In particular, the dispersion of the family component of non-agricultural ability (σ_h) and the relative importance of the individual components (ω) are identified through the standard deviation of non-agricultural wage income and the rank correlation of non-agricultural wage incomes within the family. The dispersion of farm-operating ability (σ_s) is chosen to match the dispersion of farm TFP in the production data, with the parameter capturing the correlation to non-agricultural ability (λ) identified from the correlation of non-agricultural wage income with farm profit within the family. We discuss in Appendix F the robustness of our results to the sectoral correlation of abilities, governed by λ.

Land insecurity. The parameters governing land insecurity are the family-level probability that households renting out land lose the use rights to this land η and the mean μ_φ and standard deviation σ_φ of the family-specific perceived income loss per unit of rented out land, φ_i.

We determine η directly from data drawn from our supplementary survey. Two pieces of information are critical to its estimation. First, the observed historical frequency of village-wide land reallocations and land-takings by the state, events that can trigger land loss. And second, the perceived likelihood that a household loses land they have rented out in the course of either a reallocation or land taking. In the case of land expropriation by the state, the risk for a household that rents out their land is that they are not compensated for the land loss. By law, households are to be compensated on the basis of the agricultural value of the land. Up through 2004, the probability that land was affected by either a reallocation or land takings was 10.2 percent. In addition, the probability a household puts on retaining land they rented out in the event of a land reallocation or land taking was 50 percent. Our estimate of η is the product of these two probabilities, which is 5.1 percent. We provide details of our supplementary survey in Appendix A.5.

We then identify the mean of the family income loss associated with land insecurity μ_φ with the share of village families that operate farms in 2004. Our estimation exploits that in the model, the extent to which families participate in agricultural production as operators depends crucially on the extent of land insecurity. In 2004, the share of village families that operate farms was 74 percent. The dispersion parameter σ_φ governs who these households are. We hence identify the dispersion through its selection implications. If φ_i was the same across families, then families with the highest non-agricultural ability members would choose not to operate farms and move out of agriculture, while families with the lowest non-agricultural ability members would choose to operate farms, implying a counter-factually large wage gap between these two types of families. The dispersion in φ_i across families allows the model to match the wage gap observed in the data between families in agriculture and in non-agriculture. Variation in household size in the data would also be subsumed in the estimated dispersion of φ_i, as we discuss in Appendix E.

The income loss associated with land insecurity in the calibrated model is on average around five times the rural household income. In the data, the income loss can be captured by the discounted present value of the land that a household rents out. A back-of-the-envelope calculation suggests that our calibrated value is empirically plausible. With farming the source of one-third to one-half of total family income in 2004, the agricultural revenue from land represents 13 to 20 percent of total household income. Using subjective discount rates of 3 to 5 percent, and a time horizon of 30 years, the value of land relative to household income is between 2.7 and 6.7, within which our estimate falls.

Labor mobility frictions. The mean of the labor mobility frictions to non-agriculture (μ_ξ) is chosen to match the sectoral gap between farming profit and the non-agricultural wage income. The dispersion (σ_ξ) is identified from selection in the labor supply of the part-time workers: with no dispersion in labor mobility barriers, individuals with higher non-agricultural ability and hence wage income strictly supply more labor to the non-agricultural sector. This sorting pattern is dampened by the dispersion of the labor mobility barriers. We hence choose σ_ξ to match the observed rank correlation between non-agricultural wage income and labor supply for part-time workers.

Residual farm distortions. The individual-level farm distortion τ_{ij} captures residual farm operating frictions related to the land institutions that are not directly related to land insecurity. To identify the parameters of the τ_{ij} distribution we first estimate a summary measure of farm-specific distortions, inferred as a weighted average of actual input deviations from their efficient levels (given farm productivity), what the literature calls revenue productivity or TFPR (Hsieh and Klenow, 2009). Then, following Adamopoulos et al. (2022), the correlation of farm distortions with farmer productivity and the dispersion of their non-systematic component are chosen to match the dispersion of TFPR and, among those who rent in/out land, the correlation of farm TFP and TFPR.

Calibrated parameters. Model parameters do not always uniquely determine model-implied moments. Our identification derives from the sensitivity of the model moments to specific parameters as well the direction of the effect. Appendix B outlines in more detail how model parameters map to model moments, with the estimated parameter values reported in Table B.5. The empirical moments of targeted variables, along with the implied model moments are reported in Table 4. Overall, the estimated model matches the data targets well.

Identifying land insecurity and mobility barriers. A key objective of our analysis is to identify separately the roles of land insecurity and labor mobility barriers. While both frictions lead to over-employment in agriculture and a gap in income between agriculture and non-agriculture, conceptually, the two frictions are distinct, and empirically work through different margins. The income loss associated with land insecurity is a *family-level* friction

Table 4: Targeted Moments, Data and Model

Moments	Data	Model
Employment share among village individuals:		
Full-time non-agriculture	0.302	0.299
Part-time	0.207	0.204
Median fraction of part-time hours in agriculture	0.286	0.286
Rank correlation of wages and part-time hours in nonagr.	0.398	0.397
Share of village households with farm operators	0.737	0.736
Sectoral gap: nonagr. wage versus farming profit	0.081	0.081
Family wage differentials:		
Average nonagr wage, with/without operators	−0.281	−0.282
Wage dispersion among full-time non-agr. workers:		
Standard deviation	0.610	0.608
Within-family correlation	0.558	0.557
Correlation of non-agricultural wage income and farm profit	0.080	0.078
Agricultural production:		
Standard deviation of farm TFP	0.657	0.660
Standard deviation of farm TFPR	0.633	0.631
Rank correlation of farm TFP and TFPR	0.963	0.975
Nominal agr. to non-agr. labor productivity ratio	0.388	0.389

Note: Reporting the 14 moments targeted in the data and the corresponding model estimated moments.

and its effects manifest through the family decision of whether to operate a farm or not. As a result, land insecurity has a first-order effect on the share of village families operating farms. By contrast, the labor mobility barrier is an *individual-level* friction, with its effects manifesting through individual family member decisions of whether to migrate to the city or not. Given that a family needs only one member to operate the farm in order to secure the land, all other members are not bound by the land insecurity friction. Hence, labor mobility barriers are critical for the sectoral income gap between agriculture and non-agriculture. Our identification derives from the sensitivity of the two margins, i.e. the share of families operating farms and the sectoral income gap, to land insecurity and labor mobility barriers.

To appreciate our identification approach, consider the extreme case without land insecurity. Calibrating the labor mobility barriers to match the sectoral income gap, as is typically done in the literature, the model misses the share of village families operating farms by more than 30 percentage points (42.4 percent versus 73.7 percent in the data). Moreover, the labor mobility barriers required to account for the sectoral income gap are considerably higher: the labor mobility barriers are 79.5 percent on average, compared to 54.0 percent in the baseline calibration with land insecurity. Our explicit modeling of land rental markets and the associated land insecurity unpacks part of the "black box" of labor mobility barriers typically emphasized in the literature. Similarly, in the case without labor mobility barriers and only land insecurity, the model misses the income gap between agriculture and non-agriculture observed in the data. In Table B.6, Appendix B, we more systematically illustrate the sensitivity of the key model moments to changes in each of the parameters associated with the land insecurity and labor mobility frictions. We confirm our identification strategy that the percentage of farm-operating households is particularly sensitive to the level of the land insecurity, while the sectoral income gap is more sensitive to the level of the labor mobility barriers.

Implications for land rental markets. The calibrated model is broadly consistent with the extent of land rentals in the data for China, and the selection of farm households into rentals. In the data, 78 percent of households who operate farms do not participate in the rental market. While not targeted, in the calibrated model we find that 73 percent of farm operators neither rent-in nor rent-out land. The reason for limited participation in rental markets in the model is that with land insecurity, renting out land is associated with an expropriation risk which effectively reduces the return on land that is rented out. Our model endogenously generates the bulk of limited participation in the rental market without resorting to exogenous fixed costs or wedges between reservation prices to rent-in and rent-out. On the intensive margin of rentals, among those who rent in land, our model generates an average rental of 0.98 of average farm size, similar to that in the data (0.89). Our model is also consistent with the data regarding *who* rents in land. In the data, the correlation between farm productivity and an indicator for renting in land is 0.06, a sign of misallocation, which could occur when rental is based on familial relationships rather than productivity.

Without targeting it in the calibration, our model also implies a weak correlation of 0.27.

In Appendix E we also show that if a household member operating the family farm leaves the sample (e.g., passes away), they are replaced with a family member that was previously working in non-agriculture. This is consistent with the model implications that families assign as operators members that do not necessarily have a comparative advantage in agriculture.

6 Quantitative Analysis

We evaluate the quantitative role of land security and labor mobility frictions in 2004, the year of our baseline estimation, and then analyze their evolution over time. We also study the link between land insecurity and misallocation in agriculture.

6.1 Land Security and Labor Mobility Barriers

To assess the importance of land insecurity and labor mobility barriers on outcomes, we perform two quantitative experiments from the benchmark economy. First, we assess the role of land insecurity by setting $\eta\varphi_i = 0$ for all families. Second, we eliminate labor mobility barriers by setting ξ_{ij} to zero for all individuals and families. We examine the quantitative implications of these changes for sector employment, agricultural labor productivity, and other aggregate outcomes.

The results of these counterfactual experiments are presented in Table 5. Eliminating land insecurity ($\eta\varphi_i = 0$) in the second column results in a substantial decline in the fraction of households who operate farms from 73.6 percent to only 28.2 percent. This suggests that more than half of farms in the baseline are "zombie" farms, i.e., low productivity rural households that operate farms simply to avoid losing their allocated land. As a result, average farm size more than doubles with the exit of "zombie" farms. The share of employment in agriculture among village households falls less precipitously from 56.4 percent to 45.6 percent.

Table 5: The Role of Land Security and Labor Mobility Barriers

	Baseline	Land Security	No Labor Barriers
Village families operating farms (%)	73.6	28.2	49.1
Agricultural employment share among villagers (%)	56.4	45.6	46.3
Δ Agricultural output (%)	–	+0.9	−4.9
Δ Agricultural labor productivity (%)	–	+24.7	+15.9
Δ Median log farm operator ability (%)	–	+22.1	+3.9
Δ Non-agricultural output (%)	–	+4.0	+4.6
Δ Real GDP per capita (%)	–	+3.5	+2.7
Within-household selection in farming:			
% of farm operators with highest s_{ij}	61.2	69.6	62.6
Nominal agricultural productivity gap	2.57	2.20	1.39

Notes: Statistics for the baseline model, counterfactual land security ($\eta\varphi_i = 0$ for all i), and counterfactual no labor mobility barriers ($\xi_{ij} = 0$). The first two and the last two statistics are in levels, while all other statistics are displayed as differences compared to the baseline model.

There is improved selection into farming, both within families as well as across families. The ability of the median operator rises by 22.1 percent as the percentage of farming households for whom the farm operator has the highest agricultural ability increases from 61.2 percent in the baseline to 69.6 percent. Overall, agricultural labor productivity rises by 24.7 percent. With labor released from agriculture to non-agriculture, aggregate non-agricultural output increases by 4.0 percent. The total increase in real GDP per capita, calculated using a chain-type quantity index, is 3.5 percent.

Eliminating labor mobility barriers ($\xi_{ij} = 0$) in the third column of Table 5 has similar effects on key moments in the data, however several differences emerge. First, removal of these constraints results in a smaller reduction in the share of households operating farms, but a similar reduction in agricultural employment. This confirms our identification strategy that land insecurity matters especially at the household level, while labor mobility barriers matter more at the individual level. Accompanying these changes are smaller improvements in farm operator ability through selection, which only rises by 3.9 percent. Second, there are much larger reductions in output in farming that are partially offset by larger gains in non-agricultural output. Real GDP per capita rises by 2.7 percent compared to 3.5 percent when we eliminate land insecurity. Overall, labor mobility barriers have similar effects as land insecurity on the share of employment in agriculture and GDP per capita, but weaker effects on selection and agricultural labor productivity.

We also use our framework to shed light on the nominal agricultural productivity gap (APG), defined as the ratio in value added per unit of labor input between agriculture and non-agriculture. It is well-documented that poor countries tend to have a large APG (Gollin et al., 2014b). Our quantitative framework suggests that both land insecurity and labor mobility barriers contribute to the size of the APG. In our baseline economy, the targeted APG is 2.57-fold, but this drops to 2.20 when we eliminate land insecurity and to 1.39 when we eliminate labor mobility barriers. Eliminating labor mobility barriers has a larger effect on the APG despite a smaller effect on agricultural labor productivity. This highlights that closing the nominal labor productivity gap between sectors does not necessarily close cross-country agricultural productivity differences (Restuccia et al., 2008; Gollin et al., 2014a). In Appendix D, we examine the role of within-family selection for real agricultural productivity and the nominal agricultural productivity gap.

6.2 Evolution of Frictions Over Time

We leverage the survey data through 2018 to assess the evolution of land insecurity and labor mobility barriers over time. Our baseline model is calibrated to 2004 moments. We re-calibrate the model to match data moments in 2018, which allows us to back out the land insecurity ($\eta\varphi_i$) and labor mobility (ξ_{ij}) parameters at the end of our period. Details regarding the re-calibration in 2018 are provided in Appendix C. We use the estimated values of land insecurity and labor mobility barriers to perform two separate counterfactual experiments in our baseline economy. First, we change the land insecurity parameters to match their estimated values for 2018. To do this we: (i) reduce the probability that a household renting out their land loses it through expropriation, η, from 5.1 percent in 2004 to 0.8 percent in 2018; and (ii) we reduce the expected income loss associated with land

Table 6: The Evolution of Frictions Over Time

	Baseline	2018 Land Security	2018 Labor Barriers
Village families operating farms (%)	73.6	37.6	83.5
Agricultural employment share among villagers (%)	56.4	47.6	61.8
Δ Agricultural output (%)	–	+1.1	+3.3
Δ Agricultural labor productivity (%)	–	+20.6	−5.7
Δ Median log farm operator ability (%)	–	+14.6	−0.4
Δ Non-agricultural output (%)	–	+3.4	−2.3
Δ Real GDP per capita (%)	–	+3.1	−1.5
Within-household selection in farming:			
% of farm operators with highest s_{ij}	61.2	66.6	63.4
Nominal agricultural productivity gap	2.57	2.32	3.41

Notes: Statistics for the baseline 2004 economy and two counterfactuals with frictions at their 2018 value. We replace the level of land frictions η and μ_φ such that the average $(\eta\varphi_i\bar{\ell}/I_i)$ matches that of the 2018 level (second column) and labor mobility frictions ξ_{ij} to 2018 values (third column). The first two and the last two statistics are in levels, while all other statistics are displayed as differences compared to the baseline 2004 economy.

insecurity ($\eta\varphi_i$) to match the 2018 ratio of 5.8 percent of annual family income (versus 27.9 percent in 2004). Second, maintaining land insecurity parameters at their 2004 levels, we change labor mobility barriers to their 2018 estimated values.

The effects of these experiments are reported in Table 6. When we change the extent of land insecurity alone (second column) from their 2004 to their 2018 level, the percentage of rural households who operate farms falls substantially from 73.6 percent to 37.6 percent. As a result, agricultural labor productivity increases by 20.6 percent. This result suggests that land insecurity in 2018 is much less severe compared to 2004, consistent with a literature emphasizing the role of institutional reforms in China in greatly alleviating frictions in the land market.

Over the period between 2004 and 2018, entire families in the RCRE survey occasionally migrate to urban areas and hence are replaced in the sample. If these households dropped from the sample are less likely to be engaged in farming before migration than newly added households, our calculation could underestimate the decline in the percentage of households who operate farms, and in turn, underestimate the improvement in land security over time.

We examine the robustness of our estimates by also considering a balanced sample of households between 2004 and 2018 and find that our estimates do not differ substantially, and that focusing on the unbalanced panel delivers more conservative estimates for the drop in land insecurity. We provide more details and a discussion of attrition in our sample in Appendix A.6.

When we change labor mobility barriers alone from 2004 to their 2018 values (third column), we find that agricultural employment actually increases and agricultural labor productivity declines slightly. This suggests that labor mobility barriers between agriculture and the non-agricultural sectors have increased over our sample period.

The comparison in the evolution of land and labor frictions has important implications for our understanding of migration costs. Simply looking at data on migration flows, we might infer on the basis of the reallocation of labor from agriculture to non-agriculture over time that migration costs have fallen. But looking at the distinct components of migration costs, we find that during our sample period, the reduction in overall migration costs is due exclusively to a reduction in land insecurity rather than labor mobility barriers, the focus of a large literature. In fact, our estimates suggest that residual labor mobility barriers have increased slightly over time.

6.3 Land Security and Misallocation within Agriculture

We now examine the effect of reforms in the agricultural sector that simultaneously improve land security and eliminate the idiosyncratic farm distortions within agriculture. Recall that idiosyncratic farm distortions τ_{ij} capture residual misallocation within agriculture beyond that captured by land insecurity.

To examine the interaction of land insecurity with idiosyncratic farm distortions, we implement two counterfactual experiments. First, in the baseline model we remove all idiosyncratic farm distortions by setting $\tau_{ij} = 1$ for all individuals and households. Second, in addition to eliminating the idiosyncratic distortions, we remove land insecurity ($\eta\varphi_i = 0$

for all families). Results of these two experiments are reported in columns three and four of Table 7, along with effects of the earlier land security experiment, reproduced in column two. Since in practice agricultural reform is unlikely to eliminate entirely both land insecurity and the farm-level distortions, our results capture an upper bound of these effects.

Consistent with the findings in Adamopoulos et al. (2022), removing idiosyncratic distortions has a large positive effect on agricultural productivity. The striking finding is that land security has much bigger productivity effects in the economy without idiosyncratic farm distortions ($\tau_{ij} = 1$). Comparing columns three and four, implementing land security in an economy without idiosyncratic farm distortions increases agricultural labor productivity by three times, while implementing land security alone increases agricultural labor productivity by 25 percent in the baseline economy with idiosyncratic distortions. This result illustrates an important complementarity between these two frictions.

Table 7: Interaction between Land Security and Misallocation within Agriculture

	Baseline	Land Security	No Idiosyncratic Farm Distortions	+ Land Security
Village families operating farms (%)	73.6	28.2	56.5	2.2
Agr. employment share among villagers (%)	56.4	45.6	36.9	20.2
Δ Agricultural output (%)	–	+0.9	+56.9	+87.7
Δ Agricultural labor productivity (%)	–	+24.7	+140.0	+424.5
Δ Median log farm operator ability (%)	–	+22.1	−1.9	+61.3
Δ Non-agricultural output (%)	–	+4.0	+6.0	+10.6
Δ Real GDP per capita (%)	–	+3.5	+10.6	+17.2
Within-household selection in farming:				
% of farm operators with highest s_{ij}	61.2	69.6	52.3	93.2
Nominal agricultural productivity gap	2.57	2.20	5.70	3.75

Notes: Statistics of the baseline model for three counterfactual experiments: Land security ($\eta\varphi_i = 0$), no idiosyncratic farm distortions ($\tau_{ij} = 1$), and both. Statistics for deviations Δ are from the baseline model.

The reason for the significant quantitative interaction between land security and residual idiosyncratic farm distortions is two-fold. First, on the extensive margin, in an economy without idiosyncratic farm distortions land security encourages the exit of less productive farmers, resulting in a substantial increase of median log farm operator ability of 61.3 percent.

By contrast, in the presence of residual misallocation in agriculture, land security does not feature as strong selection in farming as implied by the more modest increase in median log farm operator ability (22.1 percent). Second, on the intensive margin, land security without idiosyncratic farm distortions enables the reallocation of land from less to more productive farms, whereas with idiosyncratic distortions the land reallocation does not necessary flow to the most productive farms. Also note that eliminating idiosyncratic farm distortions in the presence of land insecurity aggravates the misallocation of farming talent, with only 52.3 percent of the farming households now having the most productive operator. Agricultural reforms that improve land security and resource allocation within agriculture can attain a substantial increase in agricultural productivity due in part to improved selection within households.

7 Extensions

We extend the main framework to incorporate a rural non-agricultural sector and then age differences among family members. We also examine the role of regional heterogeneity by comparing peri-urban and remote villages.

7.1 Rural and Urban Non-Agricultural Sectors

A key feature of China's structural transformation is the significant role played by non-agricultural employment opportunities in the rural areas. In Table 8, we document the changes between 2004 and 2018 in the level and composition of non-agricultural employment among rural households, where we distinguish between individuals working in urban and rural locations. More than half of those leaving agriculture find work in rural non-agriculture, with the percentage slightly higher earlier in our sample period. To capture this feature we extend our model to allow for two non-agricultural sectors: a rural non-agricultural sector (r) in rural townships and an urban non-agricultural sector (u) in urban centres. With this extension, the model can distinguish between spatial reallocation (rural versus urban areas) and sectoral reallocation (agriculture versus non-agriculture).

Table 8: Status and Location of Non-Agricultural Work by Rural Individuals

Variable	2004	2009	2013	2018
Share full-time non-agriculture (%)	30.2	38.0	43.3	51.6
Rural (%)	14.8	17.1	20.6	25.9
Urban (%)	15.4	21.0	22.7	25.7
Share part-time non-agriculture (%)	20.7	17.9	16.1	13.7
Rural (%)	15.9	13.7	12.3	9.8
Urban (%)	4.8	4.2	3.7	3.9

Notes: All statistics are calculated from micro-level data from the RCRE survey, as shares out of the total numbers of employed individuals in the survey. For each year the shares of rural and urban workers sum up to the total share in non-agriculture by full- and part-time status.

The production of the two non-agricultural sectors are summarized by constant returns to scale technologies

$$Y_r = A_r H_r, \quad Y_u = A_u H_u,$$

where A_r and A_u are the productivity parameters and H_r and H_u are the total labor inputs in efficiency units. The outputs from the two non-agricultural sectors are perfect substitutes in consumption. Individuals in villages can choose to work locally in the rural non-agricultural sector, or can migrate to the cities to work in the urban non-agricultural sector. Work in each of the two non-agricultural sectors is subject to potentially different individual-level labor mobility barriers, given by ξ_{ij}^r and ξ_{ij}^u for the rural and urban non-agricultural sectors, respectively. An individual choosing full-time non-agriculture employment works in the urban non-agricultural sector if the wage rate net of mobility barriers exceeds that in the rural non-agricultural sector. Individuals can also choose to work part-time in each non-agricultural sector, with c_r (c_u) capturing the fixed cost associated with part-time farming and rural (urban) non-agricultural work. Note that we do not allow for part-time work between rural and urban non-agricultural sectors in the model as this is extremely rare in the data.

We parameterize the two labor mobility barriers ξ_{ij}^r and ξ_{ij}^u as in equation (14), with means given by μ_r and μ_u and the same dispersion σ_ξ. Extending our model with a rural non-agricultural sector introduces three more parameters to determine (A_r, μ_r, c_r), in addition to (A_u, μ_u, c_u). We normalize $A_r = 1$ and calibrate relative urban productivity A_u to match the split of full-time non-agricultural employment between the rural and urban sectors. We restrict μ_r (μ_u) and c_r (c_u) to match the sectoral income gap between rural (urban) non-agricultural wages and farming profit; and the shares of part-time workers of the rural (urban) non-agricultural sectors. The estimated parameter values are provided in Table C.8, in Appendix C.

We find that it is less costly to work part-time in rural non-agriculture than in urban non-agriculture ($c_r = 0.06$ versus $c_u = 0.11$) as rural non-agriculture does not require individuals to move. At the same time however, we estimate that labor mobility barriers are higher on average for the rural non-agricultural sector ($\mu_r = 1.01$) than for the urban non-agricultural sector ($\mu_u = 0.42$). Although the wage gap between rural non-agriculture and agriculture is larger than that between urban non-agriculture and agriculture, individuals are less likely to move to full-time rural non-agricultural work. The model rationalizes this behavior with a larger mobility barrier to rural non-agriculture.

We perform the same experiments as in our main model, eliminating in turn land insecurity ($\eta\varphi_i = 0$ for all i) and labor mobility barriers ($\xi_{ij}^r = \xi_{ij}^u = 0$). The results of these experiments are summarized in Table 9. Paralleling our baseline results in Table 5, we find that both experiments have a similar effect on the share of labor in agriculture, but other differences emerge. Land security has a larger effect on the fraction of village families who operate farms, selection, and agricultural productivity, while eliminating labor mobility barriers results in larger gains in non-agricultural output.

Table 9: Extended Model with Rural and Urban Non-Agriculture

	Baseline	Land Security	No Labor Barriers
Village families operating farms (%)	73.6	27.7	47.4
Agricultural employment share among villagers (%)	55.3	46.6	46.1
Δ Agricultural output (%)	–	-0.2	-5.6
Δ Agricultural labor productivity (%)	–	+18.4	+13.1
Δ Median log farm operator ability (%)	–	+16.6	+0.4
Δ Non-agricultural output (%)	–	+3.5	+9.8
Δ Real GDP per capita (%)	–	+2.9	+6.6
Within-household selection in farming:			
% of farm operators with highest s_{ij}	54.2	63.3	54.9
Nominal agricultural productivity gap	2.58	2.14	1.33

Notes: Statistics for the extended model with two non-agricultural sectors, counterfactual land security ($\eta\varphi_i = 0$ for all i), and counterfactual no labor mobility barriers ($\xi^r_{ij} = \xi^u_{ij} = 0$). The first two and the last two statistics are in levels, while all other statistics are displayed as differences compared to the baseline model.

7.2 Age Differences within Family Members

Several recent studies have documented the phenomenon of structural transformation by age (Hobijn et al., 2018; Porzio et al., 2022). In the Chinese context, we also find that older individuals are substantially over represented in agriculture, either as farm operators or as full-time agricultural workers, see Tables A.1 and A.2 in Appendix A.4. We further extend our framework to include age differences, in addition to the two non-agricultural sectors described above. This extension allows an examination of the effect of age differences for family, occupational, and sectoral selection and their aggregate implications for productivity and structural change. We perform the same counterfactuals of improving land security and removing labor mobility barriers, and use the framework to disentangle the sources of the over-representation of old adults in agriculture.

We incorporate age demographics in the model by assuming that each individual j in family i can be either old with probability p_o, or young with complementary probability $1 - p_o$. We allow individual-level ability distributions and labor mobility barriers to differ by age group. Specifically, the individual components of non-agricultural and farm-operating abilities, h_{ij}^I and s_{ij}^I, are mean zero for old adults (as in the baseline), but have means μ_h^y and μ_s^y for the young adults. Rural and urban labor mobility barriers of the young and old have the same parametric form as in the baseline equation (14). The means of the labor mobility barriers to rural and urban non-agriculture are μ_r and μ_u for the young adults, but are shifted by μ_r^o and μ_u^o for the old adults.

In the data, we define young individuals as those between 16 and 44 years of age, and old as those with age 45 and above, but our results are robust to alternative reasonable categorizations of age groups. Adding age differences to our framework introduces four more parameters to the estimation $\{\mu_h^y, \mu_s^y, \mu_r^o, \mu_u^o\}$. We use four additional moments in the data to restrict these parameters: the young-old income differentials among rural non-agricultural workers and among farm operators, and the fraction of individuals who are old among full-time rural non-agricultural workers and among full-time agricultural workers or farm operators. All calibrated parameter values are provided in Table C.8, Appendix C.

Compared to old adults, we find that the young have an absolute advantage in both sectors but a comparative advantage in non-agriculture ($\mu_h^y = 0.20$ and $\mu_s^y = 0.15$). In addition, the estimated labor mobility barriers for the old are higher, especially to the urban non-agricultural sector ($\mu_r^o = 0.41$ and $\mu_u^o = 1.10$), which implies that the old are particularly less mobile across space. The relatively lower labor mobility barriers in the urban non-agricultural sector for the young could reflect non-pecuniary benefits for young individuals in the cities, such as life-cycle growth, urban amenities, or social welfare.

We perform the same experiments of providing land security ($\eta\varphi_i = 0$ for all i) and eliminating labor mobility barriers ($\xi_{ij}^r = \xi_{ij}^u = 0$), which we summarize in Table 10. The results are similar to our baseline results in Table 5 and those reported with the two non-agricultural sectors in Table 9. Improving land security has a stronger impact on reducing the fraction of village families who operate farms, and improving selection into farming. However, now those that select into operating farms are disproportionately from the young group. Eliminating the labor mobility barriers has a stronger effect on aggregate output per

Table 10: Extended Model with Age Differences

	Baseline	Land Security	No Labor Barriers
Village families operating farms (%)	73.6	31.9	44.7
Agricultural employment share among villagers (%)	56.7	46.1	43.4
Δ Agricultural output (%)	–	+1.1	−5.6
Δ Agricultural labor productivity (%)	–	+24.4	+23.2
Δ Median log farm operator ability (%)	–	+19.6	+5.9
Δ Non-agricultural output (%)	–	+4.0	+10.3
Δ Real GDP per capita (%)	–	+3.6	+6.8
Within-household selection in farming:			
% of farm operators with highest s_{ij}	57.2	63.2	59.9
Nominal agricultural productivity gap	2.58	2.25	1.20

Notes: Statistics for the extended model with urban and rural non-agriculture and young-old age differences, counterfactual land security ($\eta\varphi_i = 0$ for all i), and counterfactual no labor mobility barriers ($\xi^r_{ij} = \xi^u_{ij} = 0$). The first two and the last two statistics are in levels, while all other statistics are displayed as differences compared to the baseline model.

capita.

We also use our framework to shed light on the fact that the old are much more likely to be involved in farming. Two forces are at work. First, old adults have a comparative advantage in agriculture. Second, labor mobility barriers are more severe for the old than for the young. Both factors predispose families to choose older family members as the farm operator.

Table 11 reports separately agricultural employment shares for young and old individuals, and the split between operators versus agricultural workers, for the (extended) baseline and our counterfactuals. In our baseline, 66.3 percent of old individuals work exclusively in agriculture, either as farm operators or full-time agricultural workers, compared to 38.6 percent of young individuals. Eliminating land insecurity, agricultural employment shares decline among both young and old, with a slightly larger decline among the old. Eliminating labor mobility barriers, the agricultural employment shares also decline substantially among both young and old groups, but especially among the old. This is consistent with our finding that labor mobility barriers are more severe for the old, and key in explaining the young-old gap in agricultural employment.

Table 11: Occupation Differences by Age Groups in the Model

	Baseline	Land Security	No Labor Barriers	Both
Old				
Agricultural employment share (%)	66.3	51.7	37.7	30.8
Operator	31.3	13.2	15.4	8.7
Worker	35.0	38.5	22.3	22.1
Young				
Agricultural employment share (%)	38.6	27.5	30.0	24.4
Operator	19.9	8.9	14.6	9.1
Worker	18.6	18.6	15.4	15.3

Notes: Share of old and young adults among village households in agricultural employment (and between operator and agricultural worker) for the baseline model and counterfactuals.

In the last column of Table 11 we eliminate both land insecurity and labor mobility barriers. In this case, a gap in agricultural employment between young and old of 6.3 percentage points still exists. Out of a total gap of 27.8 percentage points in the calibrated model (and in the data), 20.1 percentage points are the product of labor mobility barriers and 3.7 percentage points due to land insecurity, hence land insecurity accounts for only about 15% of age gap in agricultural employment. However, that land insecurity contributes about 40% to age gap among farm operators. This finding is consistent with the idea that in the face of land insecurity, a family's choice of farm operator may not be based on comparative advantage. Rather, a family could choose an individual with few outside options (and hence low opportunity cost) to operate the family farm. In our estimation this typically means choosing an old individual to be the operator.

7.3 Village Heterogeneity

Rich heterogeneity exists across villages in China. We assess how land security and labor mobility barriers differ across one dimension, proximity to urban centers. On the basis of a village's location indicator provided in the survey, we divide villages into two groups: (a) peri-urban; and (b) remote. For 2004, this classification implies 37 percent of our farm household observations are from peri-urban villages and the remaining from remote villages. Ex-ante, we expect households in the peri-urban areas to have better access to off-farm opportunities, and thus, face lower labor mobility barriers. Land expropriations by the state for non-agricultural purposes are also more prominent in peri-urban areas. However, officials in these villages are also under more careful supervision by higher-level authorities, which may make it easier for village households to assert their rights to compensation in the event of a land taking by the state. In Appendix A.5 we calculate the probability that households renting out land lose the use rights (η) separately for peri-urban and remote villages, and we find that overall the peri-urban areas have slightly lower η.

For each region, we re-calibrate the model. The aggregate target moments remain the same, while micro data moments are calculated using observations from each set of villages. In each re-calibration, we assume that all villages in the economy are either only peri-urban or remote. The calibrated parameter values are reported in Table C.8 in Appendix C. Consistent with our priors, peri-urban villages feature a substantially lower fraction of households operating farms and a lower share of individuals employed in agriculture. Either off-farm opportunities are more appealing to individuals in the peri-urban region, or land insecurity and mobility frictions are less severe. We now take a deeper look at land security and labor mobility barriers for peri-urban and remote regions.

Land insecurity is substantially less severe in the peri-urban than in remote villages. Land income loss relative to annual family income is 38 percent lower in peri-urban compared to remote regions. Providing land security ($\eta\varphi_i = 0$) in the peri-urban area reduces the percentage of village households operating farms by 19.4 percentage points. As expected, the effects are much larger in the remote villages—the percentage of village households operating farms drops by 41.4 percentage points.

The comparison of labor mobility barriers between the two types of villages is slightly more complicated. Labor mobility barriers to the rural non-agricultural sector are substan-

Table 12: Model with Regional Heterogeneity

(a) Peri-urban

	Baseline	Land Security	No Labor Barriers
Village households operating farms (%)	57.2	27.8	33.4
Agricultural employment share among villagers (%)	48.2	40.9	37.3
Δ Agricultural output (%)	–	+1.3	−6.5
Δ Agricultural labor productivity (%)	–	+19.2	+20.7
Δ Median log farm operator ability (%)	–	+21.1	+9.7
Δ Non-agricultural output (%)	–	+2.6	+6.8
Δ Real GDP per capita (%)	–	+2.4	+4.3

(b) Remote

	Baseline	Land Security	No Labor Barriers
Village households operating farms (%)	77.8	36.4	52.2
Agricultural employment share among villagers (%)	60.0	51.5	48.8
Δ Agricultural output (%)	–	+0.4	−5.6
Δ Agricultural labor productivity (%)	–	+17.1	+16.2
Δ Median log farm operator ability (%)	–	+14.2	+2.5
Δ Non-agricultural output (%)	–	+3.6	+9.1
Δ Real GDP per capita (%)	–	+3.1	+5.6

Note: Statistics for the baseline estimated model versus our two counterfactual experiments, land security ($\eta\varphi_i = 0$ for all i) and no labor mobility barriers ($\xi^r_{ij} = \xi^u_{ij} = 0$). Panel (a) shows results for the Peri-urban area and panel (b) for the remote area. The percentage of village households operating farms and the sectoral employment shares among villagers are in levels, while all other statistics are displayed as differences compared to the baseline model.

tially lower in peri-urban compared to remote. This is reflected in the estimated parameters $\mu_r = 0.48$ and $\mu^o_r = 0.36$ for peri-urban compared with $\mu_r = 0.90$ and $\mu^o_r = 0.43$ for remote. Labor mobility barriers to the urban non-agricultural sector are similar between peri-urban and remote areas. Removing all labor mobility barriers has slightly larger effects in remote villages in terms of non-agricultural output and real GDP per capita, but similar effects for the percentage of village households operating farms and the agricultural employment share among village members.

A distinct feature of China's economic growth is the prominence of township and village entreprises (TVE's). These TVE's are often located in towns and peri-urban areas providing off-farm opportunities to village families (Brandt and Zhu, 2010). Since TVE's are part of the rural non-agricultural sector in our model, it is unsurprising that labor mobility barriers to the local non-agricultural sector are lower for village households living in peri-urban areas. The fact that peri-urban and remote village members face similar barriers to the urban non-agricultural sector indicates that the barriers may largely arise from factors other than geographical distances.

There are several reasons why land frictions might be lower in the peri-urban area. First, agriculture is less important in these localities. With more off-farm opportunities and lower labor mobility barriers, there is a smaller share of households and individuals that are farming. This translates into a lower demand for farmland and less lobbying pressure by village households on village leaders to reallocate land. Second, because peri-urban villages are nearer to cities/towns, village leaders in these localities may be under more careful scrutiny of local government, resulting in less uncompensated expropriation.

8 Conclusions

We have developed a unified framework to study jointly the effect of land insecurity and other labor mobility barriers on agricultural productivity, structural change, and economic development. A key ingredient of our model is that we allow for selection of individuals within families, across sectors, and across space. This framework along with the rich heterogeneity in the panel data for China allow us to quantify the role and impact of land insecurity and other labor mobility frictions over time.

We find that in 2004 effective migration costs are high, with land insecurity and all other labor mobility barriers playing an equal role in these costs. By 2018 however, overall migration costs have dropped, with all of the decline accounted for by improved land security. In fact, labor mobility costs unrelated to land institutions increased slightly between 2004-2018. This implies that attributing the substantial increase in migration we observe to a drop in labor mobility barriers is misleading. We also find that farm distortions within agriculture, interacting with selection, amplify the positive effect of land security on productivity.

Our findings have important implications for development policy. Weak property rights over rural land deter both land rentals across rural households and migration out of agriculture. Since 2003, there have been significant reforms in China that have strengthened household property rights in land. Our results are consistent with land policy having been central to China's structural change and agricultural productivity growth over the last two decades. While there have been reforms since the early 2000s to migration policy, particularly the hukou registration system, these changes have not had an impact on out-migration from agriculture, structural change and growth. The implication is that there are effective labor mobility barriers facing rural households that remain high. This could be due to a non-uniform reform or implementation of the migration policy (e.g., it is actually more difficult to move to the coastal mega-cities), or due to other factors that deter living in the cities such as higher housing prices. Our results indicate that there is scope for further agricultural productivity growth and structural change from reductions in other direct and indirect labor mobility barriers. As a result, future policy reforms that reduce effective rural-urban migration costs can have large welfare benefits.

We have focused on the interaction between land insecurity and migration decisions. Many other factors also affect structural transformation and rural-urban migration and could have interesting interactions with insecure land tenure, including international trade (Tombe and Zhu, 2019), capital-labor substitution (Chen, 2020a,b), choices between food crops and cash crops (Adamopoulos and Restuccia, 2020), and local entry barriers of establishments (Brandt et al., 2020). We leave the study of the interaction of these factors with insecure land tenure for future research.

References

Acemoglu, D., Johnson, S., and Robinson, J. A. (2001). The colonial origins of comparative development: An empirical investigation. *American Economic Review*, 91(5):1369–1401.

Acemoglu, D., Johnson, S., and Robinson, J. A. (2005). Institutions as a fundamental cause of long-run growth. *Handbook of Economic Growth*, 1(A):385–472.

Adamopoulos, T., Brandt, L., Leight, J., and Restuccia, D. (2022). Misallocation, selection, and productivity: A quantitative analysis with micro data from China. *Econometrica*, 90(3):1261–1282.

Adamopoulos, T. and Restuccia, D. (2014). The size distribution of farms and international productivity differences. *American Economic Review*, 104(6):1667–1697.

Adamopoulos, T. and Restuccia, D. (2020). Land reform and productivity: A quantitative analysis with micro data. *American Economic Journal: Macroeconomics*, 12(3):1–39.

Alvarez, J. A. (2020). The agricultural wage gap: Evidence from brazilian micro-data. *American Economic Journal: Macroeconomics*, 12(1):153–73.

Alvarez-Cuadrado, F., Amodio, F., and Poschke, M. (2021). Selection and Absolute Advantage in Farming and Entrepreneurship: Microeconomic Evidence and Macroeconomic Implications. *Working Paper*.

Asher, S. and Novosad, P. (2020). Rural roads and local economic development. *American Economic Review*, 110(3):797–823.

Benjamin, D. and Brandt, L. (2002). Property rights, labour markets, and efficiency in a transition economy: The case of rural China. *Canadian Journal of Economics*, 35(4):689–716.

Benjamin, D., Brandt, L., and Giles, J. (2005). The evolution of income inequality in rural China. *Economic Development and Cultural Change*, 53(4):769–824.

Bento, P. and Restuccia, D. (2017). Misallocation, establishment size, and productivity. *American Economic Journal: Macroeconomics*, 9(3):267–303.

Boppart, T. (2014). Structural change and the kaldor facts in a growth model with relative price effects and non-gorman preferences. *Econometrica*, 82(6):2167–2196.

Brandt, L., Huang, J., Li, G., and Rozelle, S. (2002). Land rights in rural China: Facts, fictions and issues. *The China Journal*, (47):67–97.

Brandt, L., Kambourov, G., and Storesletten, K. (2020). Barriers to entry and regional economic growth in China. *Working Paper*.

Brandt, L., Tombe, T., and Zhu, X. (2013). Factor market distortions across time, space and sectors in China. *Review of Economic Dynamics*, 16(1):39–58.

Brandt, L., Whiting, S. H., Zhang, L., and Zhang, T. (2017). Changing property-rights regimes: A study of rural land tenure in China. *The China Quarterly*, 232:1026–1049.

Brandt, L. and Zhu, X. (2010). Accounting for China's growth. *Working Paper*.

Brueckner, J. K. and Lall, S. V. (2015). Cities in developing countries: fueled by rural–urban migration, lacking in tenure security, and short of affordable housing. *Handbook of regional and urban economics*, 5:1399–1455.

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

Bryan, G. and Morten, M. (2019). The aggregate productivity effects of internal migration: Evidence from Indonesia. *Journal of Political Economy*, 127(5):2229–2268.

Caliendo, L., Dvorkin, M., and Parro, F. (2019). Trade and labor market dynamics: General equilibrium analysis of the china trade shock. *Econometrica*, 87(3):741–835.

Cao, K. H. and Birchenall, J. A. (2013). Agricultural productivity, structural change, and economic growth in post-reform China. *Journal of Development Economics*, 104:165–180.

Chan, K. W. (2019). China's hukou system at 60: Continuity and reform. In *Handbook on urban development in China*. Edward Elgar Publishing.

Chan, K. W. and Zhang, L. (1999). The hukou system and rural-urban migration in China: Processes and changes. *The China Quarterly*, (160):818–855.

Chari, A., Liu, E. M., Wang, S.-Y., and Wang, Y. (2021). Property rights, land misallocation, and agricultural efficiency in China. *The Review of Economic Studies*, 88(4):1831–1862.

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

Chen, C. (2020a). Capital-skill complementarity, sectoral labor productivity, and structural transformation. *Journal of Economic Dynamics and Control*, 116.

Chen, C. (2020b). Technology adoption, capital deepening, and international productivity differences. *Journal of Development Economics*, 143:102388.

Chen, C., Restuccia, D., and Santaeulalia-Llopis, R. (2022). The effects of land markets on resource allocation and agricultural productivity. *Review of Economic Dynamics*, 45:41–54.

Chow, G. C. (1993). Capital formation and economic growth in China. *The Quarterly Journal of Economics*, 108(3):809–842.

de Brauw, A. and Giles, J. (2018). Migrant labor markets and the welfare of rural household in the developing world: Evidence from China. *The World Bank Economic Review*, 32(1):1–18.

de Janvry, A., Emerick, K., Gonzales-Navarro, M., and Sadoulet, E. (2015). Delinking land rights from land use: Certification and migration in Mexico. *American Economic Review*, 105.

Eckert, F. and Peters, M. (2023). Spatial Structural Change. *Working Paper*.

Erosa, A., Fuster, L., Kambourov, G., and Rogerson, R. (2022). Hours, occupations, and gender differences in labor market outcomes. *American Economic Journal: Macroeconomics*, 14(3):543–90.

Field, E. (2007). Entitled to Work: Urban Property Rights and Labor Supply in Peru. *The Quarterly Journal of Economics*, 122(4):1561–1602.

Giles, J. and Mu, R. (2018). Village political economy, land tenure insecurity, and the rural to urban migration decision: Evidence from China. *American Journal of Agricultural Economics*, 100(2):521–544.

Gollin, D., Lagakos, D., and Waugh, M. E. (2014a). Agricultural productivity differences across countries. *American Economic Review: Papers and Proceedings*, 104(5):165–170.

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

Gollin, D., Parente, S. L., and Rogerson, R. (2002). The role of agriculture in development. *American Economic Review: Papers and Proceedings*, 92(2):160–164.

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

Hamory, J., Kleemans, M., Li, N. Y., and Miguel, E. (2021). Reevaluating agricultural productivity gaps with longitudinal microdata. *Journal of the European Economic Association*, 19(3):1522–1555.

Hao, T., Sun, R., Tombe, T., and Zhu, X. (2020). The effect of migration policy on growth, structural change, and regional inequality in China. *Journal of Monetary Economics*, 113:112–134.

Hobijn, B., Schoellman, T., and Q., A. V. (2018). Structural transformation by cohort. *Working Paper*.

Hsieh, C.-T. and Klenow, P. J. (2009). Misallocation and manufacturing TFP in China and India. *Quarterly Journal of Economics*, 124(4):1403–1448.

Kinnan, C., Wang, S.-Y., and Wang, Y. (2018). Access to migration for rural households. *American Economic Journal: Applied Economics*, 10(4):79–119.

Kongsamut, P., Rebelo, S., and Xie, D. (2001). Beyond balanced growth. *Review of Economic Studies*, 68(4):869–882.

Kung, J. K.-s. and Liu, S. (1997). Farmers' preferences regarding ownership and land tenure in post-mao china: unexpected evidence from eight counties. *The China Journal*, (38):33–63.

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

Lagakos, D. and Waugh, M. E. (2013). Selection, agriculture, and cross-country productivity differences. *American Economic Review*, 103(2):948–980.

Midrigan, V. and Xu, D. Y. (2014). Finance and misallocation: Evidence from plant-level data. *American Economic Review*, 104(2):422–458.

Morten, M. and Oliveira, J. (2018). The effects of roads on trade and migration: Evidence from a planned capital city. *NBER Working Paper*, 22158:1–64.

Munshi, K. and Rosenzweig, M. (2016). Networks and misallocation: Insurance, migration, and the rural-urban wage gap. *American Economic Review*, 106(1):46–98.

Ngai, L. R., Pissarides, C. A., and Wang, J. (2019). China's mobility barriers and employment allocations. *Journal of European Economic Association*, 17(5):1617–1653.

Porzio, T., Rossi, F., and Santangelo, G. (2022). The human side of structural transformation. *American Economic Review*, 112(8):2774–2814.

Restuccia, D. (2019). Misallocation and aggregate productivity across time and space. *Canadian Journal of Economics/Revue canadienne d'économique*, 52(1):5–32.

Restuccia, D. and Rogerson, R. (2008). Policy distortions and aggregate productivity with heterogeneous establishments. *Review of Economic Dynamics*, 11(4):707–720.

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

Schoellman, T. (2020). Comment on "migration costs and observational returns to migration in the developing world". *Journal of Monetary Economics*, forthcoming.

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

Sun, W., Bai, C., and Xie, P. (2011). The impact of Hukou reforms on rural labor migration in China. *Economic Research Journal*, 1:28–41.

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

Wu, W. and You, W. (2020). The welfare implications of internal migration restrictions: Evidence from China. *Working Paper*.

Yang, M.-J. (2021). Micro-level misallocation and selection. *American Economic Journal: Macroeconomics*, 13(4):341–68.

Yao, W. and Zhu, X. (2021). Structural change and aggregate employment fluctuations in China. *International Economic Review*, 62(1):65–100.

Zhou, C., Liang, Y., and Fuller, A. (2021). Tracing agricultural land transfer in china: Some legal and policy issues. *Land*, 10(1):58.

Appendix

A Data

Our main source of micro data is the National Fixed Point Survey carried out by the Research Center for Rural Economy (RCRE), Chinese Ministry of Agriculture. In addition to detailed individual labor supply and earnings data by sector and occupation, we have household-level information on farm inputs and outputs in physical quantities and prices. The survey was first conducted in 1986, but it was only in 2003 that individual-level labor supply data were collected. We focus on the period 2004-2018 to exploit the novel individual-level data. Below we describe in detail how we construct our empirical moments.

A.1 Farm-Household Statistics

We follow closely Adamopoulos et al. (2022) in our use of household panel data to construct estimates of farm outputs, inputs, and farm-level total factor productivity (TFP) and output per composite input (TFPR).

Gross output.— We focus on the cropping sector, and exclude sideline agricultural activity (animal husbandry, aquaculture, and forestry). Gross output of each crop is reported in physical quantity. To calculate real gross output at the farm level, we aggregate over all crops using common and constant prices of crops. For each household that sells to the market, we observe the reported quantities and revenues for each crop, which allows us to estimate prices. For each crop, we calculate the median price over all households in 2003. By 2003, quota sales to the government at below market prices were very small, and thus do not distort our measures of real output.

Labor.—Our data record labor input by crop for each family farm. Total labor input is the sum of labor input in all crops supplied by both the household and hired labor, measured in days.

Land.—Land input is calculated as the sum of cultivated area of crops and orchards. Note that we use cultivated area, rather than sown area, given that a plot of land could be used multiple times per year.

Intermediate inputs.—The value of intermediate inputs is calculated as the sum of expenditure on fertilizer, seeds, diesel fuel, pesticides, etc. We observe both quantities and expenditures on each input, and estimate the total real value of intermediate inputs using common and constant prices across all observations in 2003.

Value added.—The real value added of each farm is calculated as the difference between the real value of gross crop output and the real value of intermediate inputs. In addition, we calculate the nominal value added with local and current prices.

Farm productivity.—We separately estimate farm TFP (s) and farm revenue productivity (TFPR). In particular, we define farm TFP and TFPR as,

$$s_i = \frac{y_i}{\left(l_i^{\theta} n_i^{1-\theta}\right)^{\gamma}}, \quad \mathrm{TFPR}_i = \frac{y_i}{l_i^{\theta} n_i^{1-\theta}},$$

where y_i is real farm value added, n_i is farm labor input, and l_i is operated land. To address potential measurement errors and transitory shocks such as rainfall and health shocks, we follow Adamopoulos et al. (2022) and Chen et al. (2022). We first estimate farm productivity s_{it} for each year as the residual from the production function,

$$s_{it} = \frac{y_{it}}{\left(l_{it}^{\theta} n_{it}^{1-\theta}\right)^{\gamma}}.$$

We then estimate the permanent component of s_{it} by extracting household fixed effects from a panel regression,

$$\log s_{it} = \beta_t t + \log \tilde{s}_i + \varepsilon_{it}^{s}.$$

Using the estimated fixed effect farm TFP $\tilde{s}_i$, we reconstruct farm output using the production function:

$$\tilde{y}_{it} = A\tilde{s}_i \left(l_{it}^{\theta} n_{it}^{1-\theta} \right)^{\gamma}.$$

We then use $\tilde{y}_{it}$ to estimate farm TFPR, the dispersion of the log, and its correlation with farm TFP, which we use in our calibration.

A.2 Individual Statistics

Households have unique identifiers that allow us to link them over survey years, however the data do not come with pre-constructed individual identifiers. To construct individual IDs we match individuals within a household using information on age, gender, education and their relationship with the head of the household. This allows us to identify uniquely individuals within each household in 99 percent of the cases. We drop the remaining one percent that we cannot link over time. Our individual-level data set is comprised of all working individuals between the ages of 16 and 65.

Individual labor supply.—We calculate labor supply in days for three sectors: agriculture, rural non-agriculture, and urban non-agriculture. Rural and urban non-agricultural employment are differentiated on the basis of location. Employment in non-agriculture within the same county as the household resides is classified as rural non-agriculture, while employment outside the county is urban non-agriculture.

Individual wage rate.—We calculate the daily wage rate separately for rural and urban non-agriculture by dividing labor income with labor days. For rural non-agricultural employment, labor income is only recorded for labor supply outside the village/town but within county. We assume that the wage rate within a village/town is the same as that outside of village/town but of the same county.

Farming households.—Almost all households in our data have at least some small family plots. To classify farming and non-farming households, we impose minimum thresholds on cultivated area, days supplied to farming, and gross farm revenue. We require the cultivated area to be larger than one mu (1/6th of an acre), labor supply of the farm operator to be more than 60 days, and gross farm revenue to be more than 1000 yuan. These thresholds imply that 73.7 percent of households are farming households in 2004, and 42.4 percent in 2018.

Farm operator.—In our model, we distinguish between farm operators and agricultural workers. In the survey, household members are explicitly asked: "Are you the main decision maker of your family?" We leverage this information to assign an individual within the household as the operator, in addition to requiring all operators to supply at least 60 days to agriculture. If only one individual within the household answers "yes", then this individual is assigned to be the operator. If more than one individuals answer "yes" (or no one does), we select the household member with the highest labor supply to agriculture. In the case of a tie, we select the individual with the lowest number of days supplied to non-agriculture. If there continue to be ties, we select the individual that is male, followed by the male that is oldest. We note that farm operators spend on average 193 labor days (or 86 percent of their labor supply) in agriculture in 2004, compared to only 61 labor days of other household members that work in agriculture, but not as operators.

Occupation of individuals.—We drop from our analysis all individuals that are in school and those that work less than 10 days in any sector. We classify all other individuals into one of the following occupations: (1) farm operator, (2) full-time agricultural worker, (3) full-time rural non-agricultural worker, (4) full-time urban non-agricultural worker, (5) part-time agricultural and rural non-agricultural worker, and (6) part-time agricultural and urban non-agricultural worker. A very small portion of individuals (less than 1 percent) work in both rural and urban non-agriculture; we do not have this occupation in our model and hence we classify them as full-time urban non-agriculture, occupation (4) above, for simplicity. Similarly, a small portion of individuals (around 1 percent) work in all three sectors: agriculture, rural non-agriculture, and urban non-agriculture. We classify this set of individuals as part-time agricultural and urban non-agricultural worker, occupation (6) above.

Age groups.—We classify an individual as old if they are between the ages of 45 and 65, and young if they are between 16 to 44.

Individuals out of the labor force.—There are around 20 percent of individuals who are not working, including those that are in school or retired. Given that in our model, we do not explicitly model education or leisure we drop these individuals in our analysis.

A.3 Aggregate Moments

Employment series.—The statistical yearbook of China's National Bureau of Statistics (NBS) provides a decomposition of employment between rural and urban regions. Following Brandt and Zhu (2010) and Yao and Zhu (2021), agricultural employment is defined as total rural employment minus employment in township and village enterprises (TVE's), and private and family-run enterprises. Employment in TVE's and private and family-run enterprises then corresponds to the rural non-agricultural employment in our model. Urban employment in the data maps into the employment of our urban non-agricultural sector. We denote the economy-wide employment shares as e_a, e_r, and e_u.

Household accounting.—We have three types of households in our model: village households who choose to operate farms and decide between agriculture and non-agriculture, rural non-agricultural (township) households, and urban non-agricultural households, who only work in the rural and urban non-agricultural sector, respectively. Their shares are denoted as m_a, m_r, and m_u. We further denote the agricultural, rural-non-agricultural, and urban non-agricultural employment shares among village households as e_a^v, e_r^v, and e_u^v.

Since only village households work in agriculture, the share of village households is,

$$m_a = e_a/e_a^v.$$

The measures of township and urban households are then given by,

$$m_r = e_r - e_r^v m_a, \quad m_u = e_u - e_u^v m_a.$$

After 2013, the NBS no longer provides measures of TVE employment; hence, we cannot calculate employment shares in the same fashion. We use an alternative strategy to estimate the size of these three types of households, making a linear projection for the relative size of village households (m_a) from 2014 onward. We then calculate the economy-wide agricultural employment share as,

$$e_a = e_a^v m_a.$$

Rural non-agricultural employment e_r is then rural employment from NBS net of agricultural employment e_a, while urban non-agricultural employment e_u is simply the urban employment from NBS. The shares of township and urban households are calculated in the same way,

$$m_r = e_r - e_r^v m_a, \quad m_u = e_u - e_u^v m_a.$$

Main two-sector setup.—In our baseline calibration we focus on two sectors: agriculture and non-agriculture, without differentiating between rural and urban non-agricultural sectors. To map the data into the model, we aggregate the employment shares of the two non-agricultural sectors, e_r and e_u, into a single non-agricultural employment share e_n. In addition, township and urban households are also aggregated and mapped into the non-agricultural (urban) households in the baseline calibration.

A.4 Labor Supply by Age

The distribution of individual labor supply across sectors differs by age group. We divide workers into "young" and "old", using 45 years of age as the cutoff. As reported in Table A.1, in 2004, 58.3 percent of farm operators and 51.0 percent of full-time agricultural workers are old. By 2018, the share of the old in these two occupations increased to 80.7 and 71.2 percent, respectively. These patterns partly reflect the aging of the Chinese population and the overall rise in the number of old. Between 2004 and 2018, the old as a share of the labor force rises from 40.5 to 52.2 percent.

Table A.1: Share of Old Age by Employment Types

Variable	2004	2009	2013	2018
Population (%)	40.5	44.1	48.2	52.2
Agriculture (%)	49.7	57.6	64.8	72.1
Farm operators (%)	58.3	68.3	75.5	80.7
Full-time workers (%)	51.0	57.9	63.7	71.2
Part-time workers (%)	36.9	43.2	51.5	61.2

Notes: All statistics are calculated from the micro-level data (RCRE survey), as shares of old age among all individuals working in each occupational category. The share of old within agriculture is the weighted average of the share of old among farm operators, full-time agricultural workers, and part-time agricultural workers, where the weight is the share of all individuals working in each of these occupations.

In Table A.2, we provide separate breakdowns of labor supply in agriculture among the old (Panel A), and the young (Panel B). Overall, the share of the old engaged in agricultural activities is consistently higher than the share of young. In 2004, 86 percent of the old were engaged in agricultural activities compared to 59 percent of the young. However, the pace of structural change in employment differs significantly between the two age groups. By 2018, the share of old engaged in agriculture dropped by a quarter to 66 percent, while that of the young fell to less than half of its 2004 level, or 28 percent. Among the old working in agriculture in 2004, farm operators constitutes the largest group. Most of the reduction in the share of the old working in agriculture is due to the decline in farm operators; the shares of the old working either full- or part-time in farming fall only modestly. In contrast, in 2004, most of the young are involved in agriculture as either full-time or part-time workers. However, the reduction in the young in farming is distributed far more evenly among farm operators and workers than in the case of the old.

Table A.2: Employment Share by Age Group

Variable	2004	2009	2013	2018
Panel A: Old				
Share of old in agriculture (%)	85.6	81.1	76.1	65.6
of which:				
Farm operators (%)	39.7	36.8	34.3	28.0
Full-time workers (%)	27.1	26.7	24.7	22.7
Part-time workers (%)	18.8	17.6	17.1	14.9
Panel B: Young				
Share of young in agriculture (%)	59.0	46.9	38.6	27.6
of which:				
Farm operators (%)	19.3	13.5	10.4	7.3
Full-time workers (%)	17.7	15.3	13.1	10.0
Part-time workers (%)	22.0	18.2	15.1	10.3

Notes: All statistics are calculated from the micro-level data (RCRE survey), as shares out of the total numbers of old (Panel A) and young (Panel B) employed individuals in the survey. For each year the shares of farm operators, full-time workers, and part-time workers sum up to the total share in agriculture for that age group, save for rounding.

A.5 Information on Land Insecurity

Households that rent out their land are at risk of losing their land through two mechanisms: village land reallocations and land expropriations by the state for non-agricultural use. In the case of a village reallocation, land is taken back from households by the village, and redistributed anew to village households. In this case, the household risks losing access to the value of the land associated with farming it. In the case of a land expropriation, at risk is the compensation households are entitled to in the event of a land taking. For farmland, compensation was tied to the returns to land in agriculture.

The law governing the Household Responsibility System provided secure use rights over cultivated land for 15 years, nonetheless, village officials often reallocated village land among households before the term expired (Benjamin and Brandt, 2002). In the late 1990s and early 2000s, the State Council began to promote restrictions on land reallocations through the 1998 Land Management Law (LML) and the 2002 Rural Land Contracting Law (RLCL). Under the RLCL, household land-use rights were also extended an additional 30 years.

Table A.3: Summary of Frequency of Reallocations and Takings

(a) Land Reallocations

Period	Number	Number per year	Probability
Survey, 2004:			
1991–1999	140	15.6	13.0%
2000–2003	15	3.8	3.1%
1991–2003	155	11.9	9.9%
Survey, 2018:			
2003–2017	16	1.1	0.9%

(b) Land Takings

Period	Number	Number per year	Probability	Land (Ha)	Households
Survey, 2004:					
1991–2003	123	9.5	7.9%	581.6	11,076
Survey, 2018:					
2003–2017	123	8.2	7.3%	1,433.0	12,881

Notes: Probability is calculated as the number of events per year divided by the number of villages (120) in our sample. For land takings, not all land/households in a village are involved. We hence report the measure of land and households involved in land taking events.

To capture changes in land tenure security resulting from reallocations and land takings, we draw on two separate supplementary surveys we have undertaken with RCRE. The first was in 2004, covering the period between 1991 and 2003, and the second in 2018, covering the period from 2003–2018. Both surveys covered 120 villages. In Table A.3, we provide summary information on the number, frequency, and thus the probability of a reallocation, and the same for land takings in a village. Between 1991-2003, a total of 155 reallocations were carried out, or a reallocation in a village every 10.9 years. This implies a probability of a reallocation in a village in any given year of 9.9 percent.[3] By contrast, the 2018 survey reveals that only 16 reallocations were undertaken between 2003 and 2018, implying a probability of a reallocation in a village in any given year of less than one percent, or a reallocation every 100 years in a village. Consistent with the aims of the LML and RLCL, and data from other surveys (Brandt et al., 2017), the two surveys suggest a marked decline in the likelihood of

[3]There are some differences between villages. In 31/120 villages, there were no reallocations; in 59/120 there was a single reallocation and in 30/120 there were 2 or more.

land reallocations. Moreover, after 2003, a majority of the administrative reallocations were tied to land-takings and a redistribution of remaining land among households.

We report the same information with respect to land takings in the bottom half of Table A.3. Based on the 2004 survey, we observe 123 land-takings between 1991-2003, which implies a land-taking in a village every 12.7 years. A typical land-taking over this period covered 4.7 hectares of land and affected 90 households. The 2018 survey suggests a similar frequency in land-takings, however the amount of land that was involved was almost two and a half times larger. There was also a modest increase in the number of households affected.

An important message from the summary is that up through 2003, reallocations and land-takings both contributed to the risk of land loss. After 2003, the risks fell with improved land tenure security associated with the reduction in land reallocations. Nevertheless, perceptions of these risks by households likely adjust slowly.

A piece of information that is critical to our assessment of property rights is households' perceptions of land security if they rent land out. Households were asked the likelihood that they retained their land at either end of the contract period or within the contract period if they rented out under the following cases: (1) family members were working in the village; (2) family members were working outside the village, but older household members and children are still in the village; and (3) the entire family was working and living in the city.

In Table A.4, we summarize this information at the provincial level for the 10 provinces for which we have information. Over all households, 42 (63) percent believed they would lose their land if they rented it out (let it lie fallow) and the entire family moved to the city. The percentage falls to 30 percent if they have family members still living and working in the village. Slightly corresponding lower probabilities are reported if we focus on rentals within the contract period. Note the substantial differences between regions, with property rights much more secure in Zhejiang and Guangdong compared to the other provinces.

We asked the same set of questions to village leaders. Without exception, and with only

Table A.4: Likelihood that a Household Retains Land

	(a) At the end of contract period					
	Land is fallowed			Land is rented out		
	Family in village		Family	Family in Village		Family
	Working	Non-working	outside	Working	Non-working	outside
Shanxi	39.9	47.7	21.7	58.3	61.9	53.3
Jilin	44.1	36.8	25.7	84.4	86.0	64.1
Jiangsu	34.1	33.2	21.6	63.0	63.0	53.5
Zhejiang	71.4	69.4	68.7	80.7	80.7	80.7
Anhui	57.2	55.8	36.8	85.2	85.7	64.7
Henan	40.3	34.8	31.2	41.1	38.6	28.0
Hunan	63.2	63.7	53.9	90.2	90.2	79.4
Guangdong	61.4	59.9	58.6	71.1	73.2	77.8
Sichuan	47.3	46.6	23.7	69.2	69.2	49.0
Gansu	43.5	54.4	24.3	63.3	64.5	29.0
Average	50.3	50.2	36.6	70.7	71.3	58.0

	(b) Before the end of contract period					
	Land is fallowed			Land is rented out		
	Family in village		Family	Family in village		Family
	Working	Non-working	outside	Working	Non-working	outside
Shanxi	45.3	55.4	28.5	61.7	64.0	60.0
Jilin	51.3	53.7	45.2	87.2	87.2	85.2
Jiangsu	61.2	60.3	47.2	89.0	88.3	87.0
Zhejiang	97.5	96.3	92.7	100.0	100.0	100.0
Anhui	70.9	66.2	50.9	92.9	92.7	91.0
Henan	50.6	45.5	40.8	54.1	46.7	38.3
Hunan	83.2	89.8	79.5	98.5	97.0	97.2
Guangdong	69.1	71.2	74.1	80.4	82.5	95.0
Sichuan	69.0	68.6	48.2	89.3	89.9	87.4
Gansu	53.0	43.0	31.5	57.5	70.3	65.0
Average	65.1	65.0	53.8	70.7	81.8	80.6

Notes: The table reports the likelihood that a household retains land if it is fallowed or rented out. Numbers are in percentages.

minor differences between provinces, village leaders reported that property rights were much more secure, with the risk of losing land in the case of rental being very small.

Information from this supplementary survey helps us determine η in our baseline calibration. Recall that η represents the probability that a household loses land they rent out. Conceptually, this probability can be represented by $p_1 \times p_2$, where p_1 is the probability of either a land reallocation or land-taking in a village, and p_2 denotes the probability that a village household actually loses their land in the course of either of these events, and is not compensated for the loss.

We assign p_2, the probability of retaining rented out land conditional on a land reallocation or land taking, to be 0.5 in our baseline calibration, consistent with statistics reported in Table A.4. We highlight that the table reports the perceptions surveyed in 2018, while the perceived likelihood of losing land in 2004 should be higher. In that regard, our choice of the probability of 0.5 is conservative in our baseline calibration of 2004.

We next estimate p_1. For land reallocations, typically all of the land in a village is affected. Between 1991–2003, a total of 155 reallocations occurred in our sample of 120 villages, implying an annual probability of a land reallocation of 9.9 percent (155 events divided by 120 villages divided by 13 years). For land takings, over the period between 1991–2003, a total of 123 land taking events occurred, implying a probability of 7.9 percent. In the case of land takings, not all the land in a village is affected; on average, we find that approximately 4 percent of cultivated land is expropriated. We assign $p_1 = p_1[\text{reallocation}] + p_1[\text{takings}] = 0.099 + 0.079 \times 0.04 = 0.102$. The parameter η is hence $p_1 \times p_2 = 0.051$ in 2004.

Similarly, we can also determine p_1 for 2018. As we described above, only 16 land reallocations were undertaken between 2003 and 2018, implying $p_1[\text{reallocation}] = 0.009$. Regarding land takings, a total of 123 land taking events occurred, implying a probability of 7.3 percent. On average, we find that 9.4 percent of cultivated land is expropriated. We hence assign $p_1[\text{takings}] = 0.073 \times 0.094 = 0.007$. The parameter η is then $p1 \times p2 = (0.009 + 0.007) \times 0.5 = 0.008$ for the year 2018. We note that it is possible that η is under-estimated in 2018. For instance, even with the decline in the frequency of land reallocations, it may take time for farmers to update their beliefs regarding land tenure security. Moreover, for land takings, only a small amount of land is affected in each village, but there may be uncertainty over which plots may be affected; hence, the perceived insecurity could also be higher. We note however that, even if η were under-estimated, what matters for perceived land insecurity in our model is the product $\eta\varphi_i$, which is anchored by the share of farm operating households in villages.

For the year 2004, we also separately calculate p_1 for remote and peri-urban areas. The implied probability of land reallocation is 10.2 percent for the remote areas and 8.6 percent for the peri-urban areas. The probability of land takings is 7.1 percent for the remote area and 15.4 percent for the peri-urban area. This suggests that land reallocation is much more common in remote areas, while land takings can be relatively more important in peri-urban areas. The implied p_1 is 10.5 percent in the remote areas and 9.2 percent in the peri-urban areas.

A.6 Attrition

We briefly explain how attrition in our panel data might affect the estimation of the model. We start by noting that in the baseline calibration and identification of land insecurity and labor mobility barriers (Section 5), we estimate the model to the moments of the micro data at the beginning of our sample period, the year 2004. The main quantitative results in Section 6.1 that disentangle the role of land insecurity and other labor mobility frictions make use of the baseline estimated model from the year 2004. As a result, for Section 5 and Section 6.1 we do not explicitly use the panel dimension of the data where attrition would show up.

Attrition is potentially relevant for our over-time experiment in Section 6.2, where we calibrate the economy to 2018 to back out land insecurity and labor mobility barriers at the end of our sample period. While the economy-wide labor share in agriculture is drawn from aggregate data, and is thus immune to the extent of attrition, the share of village families that operate farms is drawn from our survey data in 2018. If attrition biases the estimate of the share of farm operators in 2018, this could potentially bias our end-of-period estimate of land insecurity.

To examine the robustness of our estimate, we consider what the share of farm operators is in a balanced panel of households in 2004 and 2018. We find that it declines from 78.0 percent to 34.4 percent, a -43.6 percentage points drop. Recall that in the unbalanced panel the share declines from 73.7 percent to 42.4 percent, a -31.3 percentage points drop. Hence, the share of farming households declines by more and the level in 2018 is lower in the balanced panel, suggesting that land security would have improved even more than in our baseline estimate.

We conclude that attrition is not substantially altering the implications for land insecurity over time, and if anything our estimates appear conservative. It is possible that attrition could still matter by changing the composition of the pool of households and individuals. However, this is also unlikely given that statisticians at the RCRE replace the households that leave the survey with households that have similar characteristics.

We have also investigated the extent of attrition over time in our data set. In our sample, household attrition is about 3 percent per year. We find that much of the attrition (about 2 percent) is related to the exit of entire villages from the survey. Further, the households that exit as well as those that enter do not differ systematically in their characteristics as one would expect by the survey design, which aims to maintain the representativeness of the sample. This is echoed in the fact that sample statistics do not differ substantially between the whole sample and the balanced sample. See also the related discussion in Benjamin et al. (2005).

B Baseline Estimation

Given the parameterization of abilities and distortions, there are a total of 23 parameters to be determined in our model: 4 parameters of the ability distribution (σ_s^H, σ_h^H, ω, and λ); 3 parameters of land insecurity (η, μ_φ and σ_φ); 2 parameters of the labor mobility barrier (μ_ξ and σ_ξ); 2 parameters of idiosyncratic farm distortions (ζ and σ_τ); 2 parameters of utility (ϕ and $\bar{a}$); 3 parameters of part-time labor supply (κ, ν, and c); 5 parameters of technology and productivity (γ, θ, A_a, A_n, and $\bar{h}$); and 2 parameters of endowments (J and N_n).

Out of these 23 parameters, 9 are either normalized or assigned values based on external data. Parameters A_a and A_n only affect the units in which output is measured, and are normalized to one. In the model, ϕ determines the share of employment in agriculture when subsistence consumption is asymptotically negligible in the case of positive productivity growth. In advanced countries, the share of employment in agriculture is below 2 percent; thus, we conservatively set $\phi = 0.02$. We set the shares of labor and land in agricultural income, based on a priori information from the literature, to 0.6 and 0.4. This implies $\gamma = 0.75$ and $\theta = 0.533$.[4] We set the curvature parameter of part-time labor supply ν to 0.6. We choose $J = 3$ to reflect the fact that the median village household has three working members in our sample. For the measures of urban households, we choose $N_n = 1.339$ to reflect non-agricultural employment that does not arise from village households. Appendix A.3 explains in detail how we determine this number in the data. Note that the value of N_n is relatively unimportant for our analysis since urban families make no production decisions in our model. We choose $\eta = 0.051$ to reflect the annual probability of 5.1 percent that a household loses land they rent out in the event of a land reallocation or expropriation, as we described in detail in Appendix A.5.

There are 14 remaining parameters that are jointly determined by comparing model and

[4] The parameter γ determines the span-of-control or extent of decreasing returns to scale, and 0.75, falls in the ballpark of values used in the literature (Restuccia and Rogerson, 2008; Hsieh and Klenow, 2009; Midrigan and Xu, 2014; Yang, 2021). We target a land income share of 0.4, consistent with aggregate estimates for China (Chow, 1993; Cao and Birchenall, 2013).

data moments. Below, we discuss their identification. All the data moments are calculated for the year 2004, except those related to farm TFP and residual farm distortions for which we use the entire panel to estimate household fixed effects. Table B.5 displays the estimated parameter values. We now discuss the empirical moments used in our estimation in more detail.

Employment and value-added shares.—We use the following four moments for employment. Among rural residents, the share of individuals working full-time in agriculture, either as agricultural workers or as farm operators, is 49.1 percent, which helps determine the value of the subsistence level of consumption $\bar{a}$ as in standard models of structural transformation. The share of village members that work part-time is 20.7 percent, which helps determine the value of the fixed cost of part-time work, c. The average number of labor days supplied to agriculture among part-time workers, normalized by total labor days, is 0.286, which is informative of the relative efficiency of labor supply in agriculture (κ). The average ability of urban households ($\bar{h}$) affects the value-added share of the non-agricultural sector. Hence we choose $\bar{h}$ such that, given the sectoral employment shares, the ratio of the average product of labor between agriculture and non-agriculture is 0.388.

Labor income.—For non-agricultural ability, we need to estimate two parameters: σ_h^H, which governs the family components of non-agricultural ability, and ω, which governs the relative importance between individual and family components of abilities. We use two moments: first, the standard deviation of log wage income among full-time non-agricultural workers, which is 0.610; and second, for those households who have two full-time non-agricultural workers, we compute the rank correlation between their wage income, which is 0.558. The correlation between the agricultural and non-agricultural ability, λ, is chosen to match the within-family correlation of non-agricultural wage income and farm profit of 0.080. Our estimate of λ is 0.32 which implies a rank correlation of abilities across sectors of 0.36. This estimate is similar to that in Lagakos and Waugh (2013) and we show in Appendix F that our results are robust to a range of values for this parameter. Recent

Table B.5: Model Parameters and Values

Parameter	Value	Description
Productivity:		
A_a	1	TFP of the agricultural sector (normalized)
A_n	1	TFP of the non-agricultural sector (normalized)
$\bar{h}$	3.670	Average ability urban households
Technologies:		
γ	0.75	Span-of-control in the agricultural sector
θ	0.533	Land income share in the agricultural sector
Labor Supply:		
ν	0.60	Curvature of labor supply in agriculture
κ	1.040	Relative efficiency of labor supply in agriculture
c	0.080	Time cost of part-time working in non-agriculture
Preferences:		
$\bar{a}$	0.219	Subsistence consumption of agricultural good
ϕ	0.02	Long-run agricultural employment share
Endowments:		
J	3	Number of individuals per village household
N_n	1.339	Measure of urban households
Ability Distribution:		
λ	0.319	Correlation between two-dimensional abilities
σ_s^H	0.577	Household component of agricultural ability
σ_h^H	0.733	Household component of non-agricultural ability
ω	0.523	Individual to household ability components
Distortions:		
ζ	-0.860	Correlated component of residual farm distortions
σ_τ	0.261	Random component of residual farm distortions
μ_ξ	0.374	Mean labor mobility barrier for rural non-agriculture
σ_ξ	0.627	Dispersion of labor mobility barrier
η	0.051	Probability household loses land they rent out
μ_φ	2.304	Mean cost of land loss if not farming
σ_φ	0.407	Dispersion land loss cost

Notes: List of parameters and calibrated values. A set of 9 parameters (A, A_n, γ, θ, ν, ϕ, J, N_n, and η) are either normalized or directly assigned values from outside evidence. The remaining 14 parameters are jointly determined by comparing model moments and targeted data moments.

studies have emphasized the link between absolute and comparative advantage as important for occupational choice (Alvarez, 2020; Alvarez-Cuadrado et al., 2021), however, we note that in our framework, as in the data, occupational choice is also dictated by distortions at the individual and household levels. In this context, given frictions, moments on sectoral incomes are more informative on occupational choices.

Agricultural production.—We use three moments on agricultural production. The dispersion of farm TFP, 0.657, is informative of agricultural ability dispersion, and pins down σ_s^H. We also use the dispersion of farm TFPR, which is 0.633, and the correlation between farm TFP and farm TFPR among those who rent in land, which is 0.968. These two moments jointly determine the parameters of idiosyncratic residual agricultural distortions, σ_τ and ζ. In order to estimate more robust measures of farm TFP and farm distortions from the data, we exploit the panel dimension of the data following Adamopoulos et al. (2022) by estimating household fixed effects that are less susceptible to measurement error.

Land insecurity.—As described earlier, the probability that a household loses the use rights to land they rent out, or η, is calculated using information from our supplementary survey directly. We use two moments to determine the level and dispersion of φ_i, which captures the income loss per unit of land the household loses. First, among village households, 73.7 percent operate farms which helps determine μ_φ. Second, the average wage income in the non-agricultural sector for families not operating a farm is 28.1 percent higher than for families with farm operators, which helps determine σ_φ. To illustrate why the wage ratio moment is informative, consider the following two cases. If $\sigma_\varphi = 0$, i.e., $\varphi_i = \mu_\varphi$ for every family, families with higher non-agricultural ability are more likely to surrender farm land and specialize in non-agriculture, resulting in a large average non-agricultural wage ratio between farming and non-farming families. If σ_φ is infinitely large, the decision of surrendering land depends mostly on the realization of φ_i rather than ability, generating a small average wage ratio between the two groups of families. We hence choose σ_φ to match the observed wage gap between these two groups of families.

Labor mobility barriers.—To identify the level of the labor mobility barriers μ_ξ, we use the sectoral earnings gap between non-agricultural wages and farm profit, which is 0.081. The dispersion σ_ξ is identified from the labor supply of the part-time workers: If there was no dispersion in the labor mobility barriers, individuals with higher non-agricultural ability and hence wage rates would strictly supply more labor into the non-agricultural sector. This sorting pattern is dampened by the dispersion of the labor mobility barriers. We hence choose σ_ξ to match an observed rank correlation of 0.398 between non-agricultural wage rates and labor supply.

Identifying land insecurity versus labor mobility barriers.—We assess the sensitivity of model moments regarding the parameter values associated with these two frictions. Specifically, we increase each parameter at a time by one percent of its calibrated value, keeping all other parameters unchanged, and report how the model moments change in Table B.6.

Table B.6: Identification of Land Insecurity and Labor Mobility Barriers

Moments	μ_φ	μ_ξ	σ_φ	σ_ξ
Village households with farm operators (%)	+0.28	+0.09	−0.08	−0.03
Sectoral gap: non-agr. wage vs. farming profit	+0.25	+0.29	−0.09	−0.23
Family wage diff., with/without operators	+0.20	−0.01	−0.16	−0.01
Rank corr., part-time non-agr. wages and hours	−0.02	−0.16	−0.09	−0.27

Notes: Percentage changes of model moments associated with changes in parameter values. Each column represents an increase of one parameter by one percent of its calibrated value, keeping all other parameters unchanged. For μ_φ and μ_ξ, which govern the means of lognormal distribution and hence are already in logs, we add one percentage point to their levels.

If we increase the level of the land insecurity friction (μ_φ) or that of the labor mobility barriers (μ_ξ), then the fraction of village households operating farms increases, and the gap between non-agricultural wages and farming profits also increases. The elasticities, however, differ. The fraction of village households operating farms is much more sensitive to the level of land insecurity (μ_φ), while the labor mobility barriers (μ_ξ) mainly affect the sectoral income gap. Similarly, the non-agricultural wage difference between two groups of families—those who operate farms and those who do not—is most informative in identifying the heterogeneity of land insecurity (σ_φ), while the correlation between labor supply and wage among part-time workers is especially informative in identifying the heterogeneity of labor mobility barriers (σ_ξ). This is consistent with our identification strategy in the baseline estimation.

C Calibration of Over-Time and Extensions

Over-time calibration. In Section 6.2, to run the over-time experiment, we re-calibrate our baseline model to the end of our sample period, 2018. Table C.7 reports the calibrated parameter values for 2018. As outlined in Section 6.2, we follow the same estimation strategy as for 2004.

Extensions. In Section 7, we present three extensions of our baseline framework: the addition of a rural non-agricultural sector; the addition of age differences across household members; and rural versus peri-urban areas. The calibration of the additional parameters under each extension are discussed respectively in Sections 7.1, 7.2, 7.3. Table C.8 provides the calibrated parameters for each extension. Column (1) is the baseline model with rural and urban non-agricultural sectors, as described in Section 7.1. Column (2) adds age differences (young and old) within family members as described in Section 7.2. Columns (3) and (4) report parameter values of the extended model for the peri-urban and remote regions as described in Section 7.3.

Table C.7: Calibrated Parameter Values for 2004 and 2018

Parameters	2004	2018
Productivity:		
A_a	1	1
A_n	1	1
$\bar{h}$	3.670	0.415
Technologies:		
γ	0.75	0.75
θ	0.533	0.533
Labor Supply:		
ν	0.60	0.60
κ	1.040	0.856
c	0.080	0.075
Preferences:		
$\bar{a}$	0.219	0.084
ϕ	0.02	0.02
Endowments:		
J	3	3
N_n	1.339	4.669
Ability Distribution:		
λ	0.319	0.543
σ_s^H	0.577	0.631
σ_h^H	0.733	0.581
ω	0.523	0.745
Distortions:		
ζ	-0.860	-0.910
σ_τ	0.261	0.238
μ_ξ	0.374	1.264
σ_ξ	0.627	0.200
η	0.051	0.008
μ_φ	2.304	2.900
σ_φ	0.407	1.582

Notes: List of parameters and calibrated values. A set of 9 parameters (A, A_n, γ, θ, ν, ϕ, J, N_n, and η) are either normalized or directly assigned values from outside evidence. The remaining 14 parameters are jointly determined by comparing model moments and targeted data moments.

Table C.8: Parameter Values for Alternative Calibrations

Parameters	(1) Baseline + Rural non-ag.	(2) (1) + age diff.	(3) (2) for peri-urban	(4) (2) for remote
Productivity:				
A	1	1	1	1
A_r	1	1	1	1
A_u	0.702	0.741	0.794	0.792
$\bar{h}$	3.883	3.806	2.566	3.876
Technologies:				
γ	0.75	0.75	0.75	0.75
θ	0.533	0.533	0.533	0.533
Labor Supply:				
ν	0.60	0.60	0.60	0.60
κ	1.036	1.037	1.012	1.036
c^r	0.059	0.061	0.062	0.057
c^u	0.110	0.107	0.077	0.110
Preferences:				
$\bar{a}$	0.189	0.233	0.214	0.236
ϕ	0.02	0.02	0.02	0.02
Endowments:				
J	3	3	3	3
N_r	0.314	0.314	0.314	0.314
N_u	1.025	1.025	1.025	1.025
p_o	0.405	0.405	0.417	0.402
Ability Distribution:				
λ	0.465	0.379	0.345	0.367
σ_s^H	0.554	0.577	0.535	0.571
σ_h^H	0.733	0.732	0.828	0.708
ω	0.468	0.488	0.490	0.528
μ_s^y	—	0.147	0.117	0.170
μ_h^y	—	0.197	0.257	0.248
Distortions:				
ζ	-0.897	-0.870	-0.818	-0.907
σ_τ	0.286	0.220	0.204	0.223
μ_r	1.014	0.844	0.481	0.901
μ_u	0.416	0.241	0.358	0.290
μ_r^o	—	0.409	0.361	0.429
μ_u^o	—	1.101	0.854	1.342
σ_ξ	0.840	0.607	0.422	0.592
η	0.051	0.051	0.046	0.052
μ_φ	2.348	2.048	1.539	2.304
σ_φ	0.471	0.396	0.251	0.528

Notes: Parameter values for alternative calibrations. Column (1) is the baseline model extended to urban and rural non-agricultural sectors. Column (2) adds to the extended model age differences among family members. Columns (3) and (4) are parameter values for extended model of peri-urban and remote regions.

D Role of Within-Family Selection

We emphasize that an important distinction between our framework and the previous literature on sectoral selection is the explicit consideration of individuals and households. This aspect differs from canonical household models such as Lagakos and Waugh (2013) and Adamopoulos et al. (2022). In our framework, households choose whether to operate a farm and which household member to assign as the farm operator. As a result, the productivity of the family farm is endogenous, and depends on the within-family selection of the household. This is relevant in our context because land insecurity affects household's decisions. The canonical household models, however, do not allow for this within-family selection and hence farm-level productivity is exogenous.

To quantify the importance of this channel in our framework, we perform the following experiment. In the baseline economy, we implement land security by setting $\eta\varphi_i = 0$, but restrict rural households operating farms to keep the same operator as in the baseline economy with land insecurity. The idea of this experiment is to shut down the within-family selection channel of who operates the farm. Note that the experiment still allows for a rural household to choose not to operate a farm, which is the standard selection channel highlighted in Lagakos and Waugh (2013) and Adamopoulos et al. (2022). We also allow for other members to change their occupations.

In this experiment, land security increases agricultural labor productivity by 20.5 percent. Recall that in Section 6.1, land security increases agricultural labor productivity by 24.7 percent. As a result, roughly 16 percent $(1 - \log(1.205)/\log(1.247))$ of the productivity gain associated with land security is accounted for by improvement in within-family selection.

In Section 6.1, land security reduces the nominal agricultural productivity gap from 2.57-fold to 2.22-fold. Without the within-family selection, land security reduces the agricultural productivity gap from 2.57-fold to 2.18-fold, a similar change to that with within-family selection. This suggests that, while within-family selection matters for the level of real agricultural productivity, it does not play a substantial role for the agricultural productivity gap. The reason is that the price of the agricultural good adjusts to roughly cancel out the effects of within-family selection on real agricultural labor productivity, leaving its effects on the nominal gap negligible. This experiment also highlights that improving real agricultural labor productivity does not necessarily close the nominal labor productivity gap between sectors.

E Other Model Implications

Our analysis emphasizes that land insecurity impedes labor mobility across sectors, occupations, and family members. We describe two implications that we validate in the data.

Heterogeneity in household size and land security. Land insecurity in our framework implies that households need to assign a member to be the farm operator in order to secure the land-use rights. If a household has more members, then it is easier for this household to allocate a member to be the operator. As a result, we might expect that large households are less affected by land insecurity than small households, ceteris paribus. We indeed find empirical support for this implication. We estimate a probit specification with a dummy indicating if an individual works in the non-agricultural sector on the left-hand-side and the number of household members on the right-hand-side, controlling for individual gender, education, and age, as well as land endowment per capita of the household. We find that the coefficient on the number of household members is positive and significant, indicating that individuals of larger households are significantly more likely to work in the non-agricultural sector. We find similar results if we estimate the effect of the number of siblings or the number of members in the labor force, instead of the number of household members.

While our model assumes a constant household size, it does allow for heterogeneity in the cost of land insecurity φ_i. The fact that larger households are affected less by land insecurity is captured in our estimation as a lower land insecurity cost φ_i for these households. In this context, heterogeneity in household size is one of the sources of heterogeneity in φ_i across households in our estimates.

Farm operator assignment when old operator exits. Land insecurity in our framework implies that a household needs a farm operator to secure their land-use rights. If an old operator exits, we would expect that another household member whose comparative advantage is outside farming switches to become the new farm operator.

We investigate this implication in our panel data. We observe farm operators that exit our sample (likely because they die) and analyze these households over time. In 2004, 36 percent of households whose old operator exits, continue operating farms and a younger household member becomes the farm operator. Among these new young farm operators, more than half of them were previously involved in non-agricultural employment, an indication that their comparative advantage was likely not in agriculture.

In addition, we also find evidence that, over time, fewer households continue operating a farm following the exit of an old farm operator, while the amount of rented-out land increases. This is consistent with our results of an improvement in land security during our sample period.

F Sectoral Correlation of Abilities

Recall that parameter λ governs the correlation between agricultural and non-agricultural abilities (s_{ij} and h_{ij}). In our baseline calibration, we choose this parameter to match the observed correlation (0.080) between a family's farming profit and the average non-agricultural wage of its members. The estimated λ is 0.32 which implies that the rank correlation between s_{ij} and h_{ij} is around 0.36, similar to that in Lagakos and Waugh (2013). To assess the importance of this parameter for the implications of land insecurity and other mobility barriers, we conduct robustness analysis by considering values of λ below and above than our baseline estimate. In particular, we set: (1) λ to zero, such that s_{ij} and h_{ij} are uncorrelated; and (2) $\lambda = 0.46$, so that the rank correlation of sectoral abilities is 0.50.

Table F.9: Robustness on the Correlation of Sectoral Abilities

Moments	Village families operating farms (%)	Δ Agricultural labor productivity (%)	Δ GDP per capita (%)
Baseline calibration ($\lambda = 0.32$):			
Benchmark	73.6	–	–
Land security	28.2	+24.7	+3.5
No labor barriers	49.1	+15.9	+2.7
No correlation ($\lambda = 0$):			
Benchmark	73.2	–	–
Land security	28.6	+14.7	+2.5
No labor barriers	48.9	+16.0	+2.7
High correlation ($\lambda = 0.46$):			
Benchmark	73.6	–	–
Land security	27.8	+30.8	+3.9
No labor barriers	49.4	+15.9	+2.5

Note: Model implications for the baseline calibration and alternative calibrations where we set the rank correlation between s_{ij} and h_{ij} to zero (the "no correlation" panel) and to 0.5 (the "high correlation" panel). In each case, we implement the land security and no labor mobility barriers experiments. We report the percentage of village households operating farms, the changes in agricultural labor productivity, and the changes in GDP per capita.

Table F.9 reports the implications of the model for the baseline calibration ($\lambda = 0.32$) and the alternative calibrations where we set the rank correlation between s_{ij} and h_{ij} to zero (the "no correlation" panel) and to 0.5 (the "high correlation" panel). Under each we run the same main experiments. With land security, the percentage of households operating farms decreases from around 73 percent to around 28 percent in all three calibrations. Land security also increases agricultural labor productivity and GDP per capita in all three calibrations, with the magnitude being slightly higher when λ is high. Similarly, when we eliminate labor mobility barriers, the implications are remarkably similar in the three calibrations. We hence conclude that our main results are not particularly sensitive to the correlation of sectoral abilities.

G Alternative Preferences

In our baseline model, we use Stone-Geary preferences, which are common in the literature on structural transformation (Kongsamut et al., 2001; Restuccia et al., 2008). We show that our results are similar if we instead use alternative more flexible preferences such as the Price-Independent Generalized Linear (PIGL) preferences. Following Boppart (2014) and Hao et al. (2020), we specify preferences as:

$$v(e, p) = \frac{1}{\varepsilon_{\text{PIGL}}} e^{\varepsilon_{\text{PIGL}}} - \frac{\nu_{\text{PIGL}}}{\gamma_{\text{PIGL}}} p^{\gamma_{\text{PIGL}}},$$

where e is total expenditure and p is the price of the agricultural good.

We re-calibrate the model with PIGL preferences to the same set of data moments as in our baseline specification. Specifically, parameters $(\phi, \bar{a})$ are no longer relevant, but we have three new preference parameters: $(\varepsilon_{\text{PIGL}}, \nu_{\text{PIGL}}, \gamma_{\text{PIGL}})$, where $\varepsilon_{\text{PIGL}}$ and γ_{PIGL} determine the income and price elasticities of the agricultural good. We follow Hao et al. (2020), who also use the PIGL preferences in the Chinese context, and set $\varepsilon_{\text{PIGL}} = 0.7$ and $\gamma_{\text{PIGL}} = 0.3$. We choose the level shifter ν_{PIGL} to match the share of full-time agricultural employment among rural village members, a moment which determined $\bar{a}$ in our baseline model.

Table G.10 displays the quantitative results using PIGL preferences. Quantitatively, the results are similar to our baseline specification in Table 5. The most noticeable differences are that the baseline model with Stone-Geary preferences implies slightly smaller productivity gain associated with the land security and no labor mobility barriers experiments. In this regard, Stone-Geary preferences provide a more conservative assessment of land insecurity and labor mobility barriers. The reason why the two preference specifications have similar effects is that the income effect in our experiments is relatively small.

Table G.10: Model Results under PIGL Preferences

	Baseline	Land Security	No Labor Barriers
Village families operating farms (%)	73.5	26.6	49.2
Agricultural employment share among villagers (%)	55.8	44.0	38.7
Δ Agricultural output (%)	–	+4.7	−16.3
Δ Agricultural labor productivity (%)	–	+32.8	+20.7
Δ Median log farm operator ability (%)	–	+28.9	−3.7
Δ Non-agricultural output (%)	–	+4.3	+8.0
Δ Real GDP per capita (%)	–	+4.4	+3.5
Within-household selection in farming:			
% of farm operators with highest s_{ij}	62.1	72.6	56.3
Nominal agricultural productivity gap	2.59	2.12	1.75

Notes: Statistics for the baseline model (with PIGL preferences), counterfactual land security ($\eta\varphi_i = 0$ for all i), and counterfactual no labor mobility barriers ($\xi_{ij} = 0$). The first two and the last two statistics are in levels, while all other statistics are displayed as differences compared to the baseline model.