Long-Run Distortions in Talent Allocation from Land Rights Insecurity

Zichen Deng   Yunyi Jin   Yun Xiao*

November 2, 2025

Preliminary draft, do not cite or circulate

Abstract

This paper examines how insecure land rights generate long-run inefficiencies in talent allocation by distorting educational investments and occupational choices. We develop a Roy model with overlapping generations and endogenous education, in which land insecurity creates barriers to exiting agriculture. When education complements non-agricultural work more than farming, the model predicts that improved land security should induce greater labor reallocation among younger generations who can still adjust their schooling, while misallocation among current workers persists due to past educational decisions. Using a cohort difference-in-differences design that exploits the staggered rollout of China's Rural Land Contracting Law, we find that individuals younger than 15 at the time of the reform were significantly more likely to attend high school and college and to shift from low-skill rural jobs to higher-skill, higher-income urban occupations than slightly older cohorts. These effects are strongest among individuals with lower comparative advantage in farming, suggesting that younger cohorts correct a talent-occupation mismatch by adjusting their education—a correction unavailable to older cohorts among whom the mismatch persists.

Keywords: Misallocation, Land, Human Capital, China, Roy model.

*Deng: School of Economics, University of Amsterdam (email: z.deng@uva.nl); Jin: Department of Economics, Lund University (email: yunyi.jin@nek.lu.se); Xiao: Department of Economics, University of Gothenburg (email: yun.xiao@economics.gu.se). For their helpful comments, we thank Elaine Liu, as well as participants in seminars at the University of Gothenburg and the University of Houston. Xiao acknowledges financial support from the Adlerbertska Research Foundation.

1 Introduction

A remarkable feature of low-income economies is that resource misallocation generates large gaps in output per worker relative to rich countries (Hsieh & Klenow, 2009). While much attention focuses on the misallocation of physical inputs like capital and labor, an equally critical factor is the misallocation of talent across occupations and locations. Mobility frictions such as informational constraints, cultural norms, and institutional barriers frequently prevent workers from sorting into their most productive uses, creating systematic distortions in how human resources are deployed (Munshi & Rosenzweig, 2016; Ngai et al., 2019).

Policy reforms that reduce these mobility frictions do not automatically eliminate misallocation, however.[1] Even when formal obstacles fall, workers may remain trapped in suboptimal positions due to past investment decisions. Among these decisions, educational investments represent a particularly important channel through which earlier mobility frictions perpetuate talent misallocation. When individuals anticipate limited career mobility, they rationally adjust their educational choices, potentially underinvesting in skills and credentials that would be valuable in more productive occupations or locations.[2] Because educational decisions are made early in life with long-term consequences, these distortions can persist even after mobility frictions decline, creating a form of path dependence that limits talent reallocation.

We examine this education-driven talent misallocation through a mobility barrier that exemplifies these dynamics: insecure land rights. In China and many developing countries, rural residents face the risk of losing valuable land-use rights if they migrate for work (Brandt et al., 2002; De Janvry et al., 2015). In China specifically, rural land is collectively owned and periodically reallocated by village authorities, creating uncertainty for those who leave their home villages. Existing research shows this insecurity deters current workers from pursuing non-agricultural employment or urban migration (Ngai et al., 2019; Chari et al., 2021; Adamopoulos et al., 2022), but its effects on forward-looking human capital decisions remain unexplored. Does land rights insecurity distort educational investments by altering expected career paths? And when policy reforms later strengthen land security, can they offset these

[1] For example, Hsieh & Klenow (2009) documents declining misallocation in China (1998–2005) but rising misallocation in India (1987–1994), despite both countries pursuing similar liberalization reforms.

[2] Nakamura et al. (2022) provides compelling evidence of this mechanism, showing that growing up on a wealthy fishing island in Iceland leads to lower educational attainment. When a volcanic eruption forced some households to relocate, their children subsequently obtained more education and achieved higher incomes.

early educational shortfalls, or do the effects persist through cohorts who made irreversible schooling choices under the old regime?

To study how mobility frictions affect education and occupational choices, we exploit the staggered rollout of China's Rural Land Contracting Law (RLCL) as an exogenous shock to land rights security. The law was first implemented in 2003 in two provinces and gradually rolled out across 23 provinces by 2015, covering 77% of the population. The law strengthened land tenure security by prohibiting arbitrary village-wide land reallocations, granting legal farmland leasing rights, and crucially, protecting land-use rights for individuals who migrate temporarily for work or education.

Our empirical analysis proceeds in two complementary parts. We first examine the RLCL's immediate effects on individuals already active in the labor market when the law took effect—the focus of existing research. Building on Chari et al. (2021), we construct a panel dataset of households and villages using the National Fixed Points (NFP) survey to estimate the RLCL's effects on land rental activities, income, and employment among the current workforce. We implement the Extended Two-Way Fixed Effects (ETWFE) estimator (Wooldridge, 2021) to accommodate heterogeneous treatment effects and avoid the negative weights problem inherent in staggered difference-in-differences designs. Consistent with prior studies, we find that the reform increases land rental activities, household income, and non-agricultural employment—effects that align with predictions from classical selection models where workers reallocate toward their comparative advantage (Roy, 1951; Lagakos & Waugh, 2013; Adamopoulos et al., 2022).

However, these immediate responses may understate the reform's full economic impact because individuals already in the labor market face constraints from their fixed educational attainment. Younger cohorts not yet in the workforce can adjust both their educational investments and subsequent occupational choices in response to improved mobility prospects. We formalize this distinction by extending the Roy model to incorporate overlapping generations with endogenous education choice, yielding a key prediction: land security improvements should have disproportionately larger effects on younger individuals, inducing them to invest more heavily in education and subsequently transition into the non-agricultural sector, where educational returns are higher. The difference between the young and the old represents the persistent misallocation among current workers due to past educational decisions.

To test this prediction, we exploit a cohort difference-in-differences (DID) design that compares individuals who reached age 15 before versus after RLCL implementation in their province. Age 15 marks the transition from compulsory to non-compulsory schooling in China, when students decide whether to pursue high school education; beyond this point, the cost of re-entering education is substantially higher, making missed investments difficult to reverse (Lu et al., 2023). The staggered provincial rollout of the RLCL, together with such cohort differences in the flexibility to adjust education, creates the quasi-experimental variation to identify how RLCL affects different cohorts differently. To implement this design, we link RLCL implementation data with the 2015 inter-censal population survey (the "mini-census"), a 1% nationally representative sample that provides the statistical power necessary for precise estimation of cohort-specific reform effects.

We find significant effects of the land reform on educational attainment among younger cohorts as compared to older ones. Being exposed to the RLCL at age 15 or earlier increases high school enrollment by 3 percentage points (8% of sample mean) and college enrollment by 3 percentage points (18% of sample mean). These effects are concentrated among boys, consistent with gender differences in land allocations in rural China. Adult women typically receive land rights through their husbands' families upon marriage rather than retaining rights in their parents' villages, making land rights insecurity less consequential for their educational decisions made at a younger age.

These educational gains translate into more occupational reallocation among the younger cohorts than the older ones. The boys who were exposed to the RLCL at age 15 are 2 percentage points less likely to work in rural low-skilled occupations—including farming and construction—and 1 percentage point more likely to work in urban high-skilled employment.[3] Consistent with improved job matching, we observe a 1% increase in individual income among these younger cohorts, suggesting they are more productive when allocated to positions better suited to their abilities. Our results indicate that land insecurity created systematic misallocation, trapping high-ability rural youth in low-skill employment and preventing them from accessing education that would unlock jobs matching their potential. By removing mobility barriers, the RLCL generates substantially more—and more skill-intensive—labor reallocation among the younger cohorts than observed among current workers, who are

[3] Construction work is classified together with farming because construction workers are predominantly temporary rural migrants who maintain agricultural production during peak seasons, unlike workers in other sectors who typically leave agriculture completely. See Section 3 for details.

constrained to move primarily between low-skill occupations due to their limited educational attainment.

Our extended Roy model distinguishes two types of compliers affected by land security improvements: (A) "security-only switchers" who leave agriculture regardless of whether or not they can adjust their education—and (B) "education-enabled switchers" who only leave agriculture if they are young enough to increase their schooling first. Our cohort DID design identifies both groups when examining educational decisions, but only group B when examining occupational switching out of agriculture. This explains why our estimated occupational effects are roughly two-thirds the magnitude of our educational effects. Our framework rationalizes the common strategy of exploiting key educational windows to estimate the impact of local shocks on human capital formation (Atkin, 2016; Lu et al., 2023). It also explains why estimates for educational attainment and occupational choice can diverge: they pertain to different complier groups. Most importantly, it suggests that focusing solely on farming versus non-farming transitions may substantially underestimate the welfare costs of mobility frictions, since much of the misallocation can happen within the non-agricultural sector.

An importance feature of the Roy model is that individuals choose their occupation based on their comparative advantage. The extended Roy model predicts that the educational responses to improved land security should be stronger among households with low comparative advantage in agriculture, for whom the gains from shifting out of agriculture are the greatest. We test this prediction by exploring heterogeneity across pre-reform marginal productivity of land (MPL) with the NFP data. Consistent with the model's prediction, we find that boys from low-MPL families experience substantially larger increases in high school attendance after the reform; girls from low- and high-MPL families show similar but much smaller (imprecisely estimated) improvements.[4] Furthermore, older cohorts' education does not differ by pre-reform agricultural productivity, ruling out absolute advantage or mean reversion as explanations. Finally, RLCL increases income by a similar magnitude for households with high and low pre-reform MPL, suggesting that income effects alone cannot explain the rise in schooling.

This study relates to several strands of literature. First, our paper contributes to the literature on input misallocation, which emphasizes the inefficient allocation of resources as a critical factor underlying low aggregate TFP levels in developing countries. Building upon the

[4]Because the last available NFP wave with a sufficient sample is 2013, we focus this heterogeneity analysis on high school attendance (the cohort is too young for reliable college or occupational outcomes).

valuable research agenda highlighted by Restuccia & Rogerson (2017), we specifically identify and analyze concrete sources of misallocation within an important developing economy. Previous studies have focused on the effects of trade, migration costs, and the *hukou* system (Tombe & Zhu, 2019; Munshi & Rosenzweig, 2016; Ngai et al., 2019), emphasizing trade-related spatial misallocation of labor through models calibrated to aggregate data moments. In contrast, our analysis investigates specific policy interventions and uniquely highlights the inefficiency in educational investments, occupation choices among children, and their implications on the aggregate economy due to inefficient land institutions. Our findings suggest that land institutions might represent a significant barrier for rural children to pursue education and occupations that match their capabilities. Thus, our study closely relates to recent contributions by Hsieh et al. (2019), Nakamura et al. (2022), Lo Bello & Morchio (2022), and Celik (2023), who study the causes and consequences of talent misallocation across occupations and spaces.

Second, our paper contributes to the extensive literature on land institutions and their economic implications (De Janvry et al., 2015; Chari et al., 2021), emphasizing particularly their effects on labor allocation (Ngai et al., 2018; Adamopoulos et al., 2022, 2024; Gai et al., 2024; Chen et al., 2025). Notably, Adamopoulos et al. (2024) and Gai et al. (2024) specifically model the distinct characteristics of young and old adults, highlighting variations in ability and mobility barriers as factors driving structural transformation by age cohorts. However, these existing analyses typically consider education levels across cohorts as fixed. In contrast, our empirical approach and theoretical model explicitly incorporate the decision of the second generation to invest in education as a response to improved land institutions. Thus, our work bridges the literature on land institutions with that addressing education-driven labor reallocation and structural transformation (Porzio et al., 2022).

The rest of the paper is organized as follows. The next section summarizes the institutional background. Section 3 describes the data, key variables. Section 4 presents a model of occupation choices to guide the empirical analyses, which are detailed in Section 5. Section 6 concludes with policy implications.

2 Institutional Background

Since the Household Responsibility System (HRS) was introduced in the early 1980s, rural households were granted use rights and residual income rights over land. Despite this reform, land remained under collective ownership at the village level rather than being privatized. Village authorities allocated land according to egalitarian principles, with distribution tied to formal village membership through the *hukou* or household registration system.[5] This approach resulted in relatively equal per capita landholdings among households within each village.

Under the initial framework, households received land use rights for 15-year periods. However, the system remained unstable through the late 1990s, as local officials frequently conducted village-wide land reallocations to adjust for demographic shifts such as births, deaths, and migration (Benjamin & Brandt, 2002). The security of land tenure faced additional threats from state expropriation for non-agricultural development projects, including infrastructure construction, urban expansion, and industrial development. When land is expropriated for such projects, land rights become especially valuable as they determine eligibility for compensation, regardless of whether the land was actively used for agricultural production. In the absence of legal protection, village residents may fear that leaving the village for any reason could threaten their rights to this compensation should expropriation occur.

A series of reforms to land policy in China beginning in the late 1990s helped strengthen household property rights but did not privatize farmland. In 1998, the Land Management Law extended the use rights of rural households to 30 years. In 2003, the RLCL was introduced to further strengthen the security of land use rights. The law protected farmers' rights to their land through the 30-year contract period, prohibited village-wide land reallocations, and granted farmers legal rights for leasing out agricultural land. Only when the whole household moves to a city and obtain urban *hukou* there, the village has the right to take the contracted land back.

As discussed by Chari et al. (2021), the implementation of the central law has been delegated to the provincial-level governments. Following many other market reforms that started after 1978, the central government issued general guidelines on the priorities, and local governments were encouraged to implement and to experiment within the guidelines (Xu, 2011). This allows provinces to implement the law at different points in time and to add regulations within the scope of the central law. For example, while the central law only stated that the contracted land cannot

[5]The *hukou* system was established in the 1950s to control migration between the countryside and the cities. The system assigns each individual either agricultural (rural) or nonagricultural (urban) *hukou* status within a specific locality—whether village, town, or city—thereby determining their legal residence, employment opportunities, and access to local public services. Rural-to-urban migration and nonagricultural employment in rural areas were highly restrictively until the early 1980s alongside broader economic reforms. Despite relaxing such restrictions, land rights remained tightly linked to agricultural *hukou* status, while rural migrants continued to be restricted in accessing local public services in urban areas.

be reclaimed by the issuing party during the contract period, province-level implementation might add regulations for specific cases where land rights might be subject to uncertainty, such as temporary migration or marriage.

To illustrate the scope of provincial-level changes, we discuss a few articles from Yunnan province's implementation of the RLCL in 2006. Article 15 states that during the contract period, land cannot be taken back from a household due to changes in the status of its members in the following cases: 1) A family member goes to college, works outside of the village, or joins the military; 2) A person gets married and moves into the spouse's household but has not received new contracted land at the new residence; 3) A woman divorces or becomes widowed and either stays in the original village or has not received new land in the new place of residence.[6] In addition, Articles 18 to 29 outline rules for leasing, transferring leases, and how to address land leasing disputes.

The RLCL is the first law that provides a basis for legal disputes regarding the uncertainty of land rights for those who migrate. In 2005, for example, a graduate student from Jiangxi sued his village committee after being denied land compensation due to having transferred his *hukou* to his university.[7] The court ultimately ruled in his favor on appeal, affirming that leaving the village for education does not strip a person of their collective land rights, especially when they remain dependent on village land income. In another case in 2019, a rural-to-urban migrant in Shandong sued the village committee for illegally reclaiming and reallocating his land without consent during the 30-year contract period after he had transferred his *hukou* to work in the city in 2003.[8] The court ruled in the worker's favor, ordering the village to return the land and compensate him 60,000 yuan, affirming that under the RLCL, land contract rights remain protected despite *hukou* relocation for urban employment. Such cases highlight that migration for education or urban employment might be costly when land rights are insecure, but the RLCL provided legal basis for protecting these rights and reduced the opportunity cost of migration.

Following Chari et al. (2021), we extend the database of the dates of the provincial-level implementation of the RLCL as the relevant time at which the law became effective locally. By the end of 2014, 23 provincial governments had made official announcements about the local implementation of RLCL (as shown in Table A1).

[6] Under existing rural land allocation practices, married women typically lose their land rights in their birth village upon marriage, with entitlements often transferring to their husband's family.

[7] https://zqb.cyol.com/content/2007-01/13/content_1641621.htm

[8] https://thepaper.cn/newsDetail_forward_25862181

3 Data

We link the timing of policy rollout in each province to two key data sources: (1) a 1 percent random subsample of the 2015 Population Survey (the "mini-census") and (2) the National Fixed Point Survey (NFP), a household panel initiated by the Chinese Ministry of Agriculture in 1986.

3.1 2015 Population Survey

We use a random subsample of the representative 1% Population Survey in 2015 (often called the mini-census), collected by the National Bureau of Statistics and containing 2,000,563 individuals. Similar to the earlier population survey conducted in 2005, the 2015 mini-census covers the entire population at their current residence, regardless of whether they hold local household registration (*hukou*). The mini-census provides information on year and month of birth, gender, education, employment status, occupation, area code of current residence, and area code of *hukou* place. The area code allows us to identify the province, prefecture, and county information of a community, which is the smallest administrative unit. We can also identify whether the community is a village or an urban community. The 2015 mini-census contains no information on whether an individual holds an urban or rural *hukou*. To approximate the rural population, we proceed in three steps. First, we exclude Beijing and Shanghai municipalities, which are fully urbanized. Second, we drop county-level administrative units designated as urban districts (*qu*). Finally, following the official urban–rural classification of the National Bureau of Statistics,[9] we retain only individuals whose *hukou* is registered in areas coded as villages, rural centers, town–village junctions, or urban–rural junctions.

We use the mini-census to construct the sample for our main analysis, where we examine the effects of the RLCL on education and labor market outcomes. Using information on educational attainment, we generate dummy variables for ever attending high school and ever attending college. For labor market outcomes, we construct four variables. The first is whether the individual is currently employed. The second is whether the individual holds a rural low-skill job, including farming and low-skill construction.[10] The third variable is whether an

[9] https://www.stats.gov.cn/zs/tjws/tjbz/202301/t20230101_1903381.html

[10] We classify construction as rural low-skill employment because examining only farming would miss temporary rural workers who maintain agricultural production during peak seasons rather than leaving agriculture completely. Among occupations available to temporary migrants, construction is the least skill-intensive and offers the greatest flexibility for workers to continue agricultural activities. Unlike manufacturing jobs with rigid production schedules or service positions requiring continuous attendance, construction work's project-based nature allows extended agricultural leave. Data from the 2010 China Family Panel Studies show that 58.4% of construction jobs are in rural areas and 87.3% are held by agricultural *hukou* holders (see Table A2). This dual

individual is employed in a high-skill occupation, defined to include government and party officials, professional and technical personnel (e.g., researchers, engineers, healthcare workers, finance specialists, lawyers, teachers), and service workers in information technology, finance, real estate, technical support, utilities, and healthcare.[11] Finally, because the mini-census does not report income, we predict individual income in 2020 by training a supervised model on data from the 2020 China Family Panel Studies (CFPS)—a large, nationally representative survey—and applying it to the mini-census data.[12]

Summary statistics for the main variables are presented in Table A3. Our sample includes individuals (98,581 males and 92,303 females) who turned age 15 between 2000 and 2010. From 2003 to 2010, 23 provinces implemented the RLCL at the local level. Individuals in China typically make high school enrollment decisions at age 15 and college enrollment decisions at age 18. The youngest cohort in our sample was already 20 years old in 2015, allowing us to observe their high school and college enrollment decisions. When analyzing labor market outcomes, we further exclude individuals who were not yet 23 years old in 2015, ensuring that we do not capture effects from delayed entry into the labor market due to college education, which is usually finished at age 23.

3.2 National Fixed Point Survey

Our analyses are supplemented using the National Fixed Point Survey (NFP), which is a nationally representative panel with rich information on agricultural activities (Chari et al., 2021). The NFP comprises three components: a village-level survey, a household-level survey, and an individual-level survey. The village questionnaire is completed by local officials and provides information on the local socio-economic environment. Households were randomly sampled within each village, and detailed information was collected for each household member. The household survey includes data on agricultural activities, income, and family background, while the individual survey captures demographic and other characteristics of household members.

engagement makes construction work institutionally and economically closer to agricultural production than to other urban employment sectors.

[11] To classify high-skill occupations, we use a separate sample of individuals born between 1971 and 1985 to calculate the share of college graduates by occupation. We define occupations with at least 25% college graduates as high-skill. This data-driven approach aligns closely with the conventional set of white-collar, professional jobs.

[12] Specifically, we estimate gender-specific regressions of log income on occupation, region, education, and age (with interactions and nonlinear terms selected via cross-validation) using employed individuals aged 20–45 in CFPS 2020 ($N = 6{,}217$). Full details on model specification, variable harmonization, and validation are reported in Appendix A.

Our analysis draws on the annual household data waves from 2000 to 2013, as participation rates at the household level dropped substantially after 2013. During this period, the dataset includes over 20,000 households across 417 villages in 31 provinces. For individual-level analysis, we focus on data collected in 2013 due to significant attrition in later years. We also use village-level data for the period 2000–2013 to match the coverage of the household and individual samples. Households that did not cultivate any land prior to the start of the land reform are excluded from our analysis, as they are unlikely to have been affected by the reform.

Summary statistics for the main variables are presented in Table A9. Panel A reports key variables at the household-year level. On average, households cultivate approximately 15.96 mu of land. The average annual income is 27844, measured in real 2002 Renminbi. Land rental activity is defined as whether a household rented in or rented out agricultural land in a given year. Panel B presents key variables at the village-year level, including total population, non-agricultural share, and migration. Panel C reports variables at the individual level, including gender, birth year, and educational levels.

4 A Model of Occupation Choice

To study how insecure land rights distort talent allocation, we build a simple partial equilibrium Roy (1951) model to guide our analysis. In this model, there are two occupations - farming in the rural area and manufacturing in the urban area - that an individual from family i and generation g can choose between. Occupations require occupation-specific skills for workers to be productive. Individuals are endowed with a bivariate skill vector $(z_F^g(i), z_M^g(i))$, where $z_k^g(i) \in (0,1)$ is the productivity of the individual from family i of generation g in occupation k. Each generation consists of a unit mass of individuals distributed across $\mathbb{Z}_F \times \mathbb{Z}_M$.

Individuals live for two periods. In the first period, individuals from generation g (the younger generation) are born as children of parents from generation $g-1$ (the older generation). After the ability is known, they make a binary educational choice $h \in \{0,1\}$ at cost c.[13] In the second period, they are parents and inelastically supply one unit of labor to market work in their chosen occupation, conditional on their education choices made in the first period. This implies that, within a given period, only one generation is actively participating in the labor market. Therefore, we abstract away from the within-household labor allocation, which is a

[13] Intuitively, $h = 0$ corresponds to completing only compulsory education, while $h = 1$ indicates progression to high school and possibly college. In our sample, more than 90% of individuals complete compulsory education while less than 40% attend high school; thus, the key margin of educational choice is high school attendance.

focus in some other related studies (Adamopoulos et al., 2024; Chen et al., 2025).

If individual i chooses to be a farmer, she decides how much land to rent in or out, $\ell^{rent}(i) > -\bar{\ell}(i)$, where $\bar{\ell}(i)$ is her endowment. These choices, however, are subject to the land insecurity risk that farming families face when renting out land. Due to land rights insecurity, individuals that rent out land face an exogenous probability $\eta \in [0,1]$ of losing the rented-out land, in which case there is a individual-specific income loss $\varphi(i) > 0$ per unit of rented-out land. Differences in $\varphi(i)$ across individuals reflect differences in the valuation of the land tied to observable and unobservable characteristics of the land and farming family (Adamopoulos et al., 2022). Individuals that rent in land do not directly face this loss, but bear part of the risk through a higher rental price. This is captured by a risk premium in the rental price, denoted by $\alpha > 0$.[14] As a result, the effective rental price of land is given by:

$$q(\eta, \ell^{\text{rent}}(i)) = \begin{cases} q_0 + \alpha\eta & \text{if } \ell^{\text{rent}}(i) > 0 \quad \text{(renting in)} \\ q_0 - \eta\varphi(i) & \text{if } \ell^{\text{rent}}(i) < 0 \quad \text{(renting out)} \end{cases}$$

The profit maximization problem for farmer i is then given by

$$I_F^g(i) = \max_{\ell^{\text{rent}}(i)} \left\{ p_a z_F^g(i) \left[\bar{\ell}(i) + \ell^{\text{rent}}(i)\right]^\theta - q(\eta, \ell^{\text{rent}}(i)) \cdot \ell^{\text{rent}}(i) \right\}$$

where $p_a > 0$ is the price of agricultural products and $\theta \in (0, 1]$ captures the importance of land relative to labor in agricultural production.

The individual i who chooses not to participate in agricultural production will rent out all the land and work in the manufacturing sector in an urban area.[15] The wage income is given by $w_M z_M^g(i)(1 + \tau h(i))$, where w_M is the wage in the manufacturing sector and $\tau > 0$ captures the complementarity between education and non-agricultural producition.[16] The total income is

$$I_M^g(i) = w_M z_M^g(i)(1 + \tau h(i)) + (q_0 - \eta\varphi(i))\bar{\ell}(i) - ch(i),$$

[14]This structure ensures that land insecurity affects all participants in the land market, not only those who rent out land, in a partial equilibrium model. In a full equilibrium model, the rental price will be endogenously determined by the market-clearing conditions.

[15]The risk is still there, even if the land is not rented out, as long as he pursues a career in an urban area. Therefore, the optimal choice is to rent the land out in this simplified setting.

[16]We assume no complementarity between education and agricultural production. Years of schooling may positively affect agricultural output only through primary and secondary education, but the effect is relatively small in less developed countries (Reimers & Klasen, 2013). Instead, productivity gains in agriculture through learning and skill development are more often realized via institutional reforms, such as technology adoption (Chavas & Nauges, 2020). These mechanisms are not directly relevant in our setting. We thus assume that education beyond compulsory school mainly increases returns outside agriculture.

where c is the cost of education.

Lastly, individuals make their educational choice $h \in (0,1)$, occupational choice $k \in \{F, M\}$, and land rental decision $\ell^{\text{rent}}(i)$ that maximizes their total income in adulthood.

$$I^g(i) = \max_{h(i), l^{rent}(i), k(i)} \left\{ I_F^g(i), I_M^g(i) \right\}$$

Educational choice is time-sensitive and can only be made in the first period. Land rental decision and occupational choice are initially made in period one but can be adjusted in period two. This implies that for a given period, any changes in η will affect the younger generation g and the older generation $g-1$ differently depending on how flexible they are in adjusting their education. We obtain the following comparative statics.

Result 1. *A reduction in η (i.e., improved land security) increases income and urban migration for current workforce who already complete education. In particular:*

- *A reduction in η increases income, i.e., $\frac{\partial I^{g-1}}{\partial \eta} \leq 0$, with strict inequality when $\ell^{rent}(i)$ does not remain zero throughout.*
- *A reduction in η increases the share of individuals working in urban manufacturing, i.e., $\frac{\partial Pr^{g-1}(M)}{\partial \eta} < 0$.*

See Appendix B for the proof of this result.

When land security improves (i.e., η falls), individuals face a lower risk of losing their land if they exit farming, which raises the attractiveness of migrating to urban manufacturing jobs. Households that switch from agriculture to non-agricultural work see income gains, while those remaining in agriculture may also benefit through lower land rental prices. However, for the first generation, additional schooling is no longer feasible, so their opportunities in non-agricultural occupations remain relatively limited.

Result 2. *The younger generation is more sensitive to improvements in land security (i.e., a reduction in η) than the older generation. In particular:*

- *A given reduction in η generates an increase in the share of educated individuals among the younger generation, i.e., $\frac{\partial Pr^g(h=1)}{\partial \eta} < 0$.*
- *A given reduction in η produces a greater increase in the share of individuals working in urban manufacturing among the younger generation, i.e., $\frac{\partial Pr^g(M)}{\partial \eta} < \frac{\partial Pr^{g-1}(M)}{\partial \eta}$.*

See Appendix B for the proof of this result.

Land insecurity affects not only the occupational choices of current workforce but also those of the younger generation, who enter the workforce in the next period. Intuitively, when land becomes more secure (lower η), the younger generation anticipate better returns to migration and urban work. As a result, they are more likely to optimally increase their education. Therefore, the effects of improved land security are stronger for the younger generation, with them showing much larger gains in education and high-skilled urban occupations than the current workforce.

The extended Roy framework identifies two groups of switchers in response to land security improvements: (A) security-only switchers—individuals who would move to non-agriculture when tenure security improves, even holding schooling fixed; and (B) education-enabled switchers—individuals who would otherwise remain in agriculture but, because schooling complements nonagricultural productivity, switch sectors once they can invest in education. Because the older generation's schooling is predetermined at the time of the reform, their occupational response can operate only through channel (A). Comparing them with the younger generation, who adjust schooling post-reform, identifies the reform's effect on education for both groups and its effect on occupational switch only through channel (B). Section B.2.3 in the Appendix provides a more detailed discussion of the switchers.

Result 3. *The younger individuals with lower agricultural productivity are more sensitive to improvements in land security (i.e., a reduction in η) than those with higher agricultural productivity. In particular:*

- *A given reduction in η induces a larger increase in the probability of educational attainment when the individual's agricultural productivity $z_F^g(i)$ is lower, i.e.,*

$$\frac{\partial^2 Pr^g(h = 1|z_F^g(i))}{\partial \eta \, \partial z_F^g(i)} > 0$$

See Appendix B for the proof of this result.

Result 3 highlights how the effects of improved land security vary by individuals' comparative advantage. Individuals with low agricultural productivity are more sensitive to land security improvements because they have a weaker comparative advantage in farming and are more likely to be at the margin between manufacturing and farming. Small improvements in land security can tip them toward manufacturing, where education becomes valuable. High-agricultural-productivity individuals have stronger reasons to remain in farming regardless of moderate improvements in land security. For individuals switching from farming to manufacturing, those with relatively high $z_M^g(i)$ are more likely to choose education. This reflects how land insecurity previously constrained mobility for those better suited to urban high-skill occupations.

5 The Impact of the RLCL on Current and Future Workforce

The empirical work below proceed three steps, testing each of the results outlined in the previous section. First, using NFP data, we estimate the policy's impact on the current workforce directly exposed and confirm consistency with prior studies. Second, using the 2015 mini-Census, we test whether the policy shifts educational attainment and occupational choices more for the future workforce, who are still flexible in adjusting their education, as compared to the current one. Finally, we draw on NFP data to assess the heterogeneous effects of RLCL across households with different agricultural productivity.

5.1 Effects on current workforce

We start with the causal impact of the RLCL reform on outcomes of the current workforce by exploiting its staggered roll-out across provinces. In doing so, we build on the framework of Chari et al. (2021), who focused on land-renting activity, but extend their two-way fixed-effects estimator with recent advances in difference-in-differences methodology (ETWFE, Wooldridge, 2021). This estimator avoids comparisons with already-treated units by saturating the model with fixed effects for all possible combinations of treatment groups (as defined by the timing of treatment) and year, thereby estimating separate average treatment effects on the treated (ATTs) for each group-year combination. Formally, we estimate the following equation for each household h in province p and in year t:

$$Y_{htp} = \sum_{g \in G} \sum_{s=s_0}^{g-1} \theta_{gs}^{pre} D_{hgs} + \sum_{g \in G} \sum_{s=g}^{T} \theta_{gs}^{post} D_{hgs} + \psi_h + \phi_t + \varepsilon_{htp}, \quad (1)$$

where Y_{htp} is the outcome of household h in year t and registered in province p. D_{hgs} is a dummy that takes the value of 1 if the observation is in the treatment group g and 0 if otherwise. G is a set that indicates when the reform started in the province. ψ_h and ϕ_t are household and year fixed effects, respectively.[17] θ_{gs}^{post} represents the average treatment effect that households in treatment group g experience after the RLCL was implemented (ATT(gs)). We use provinces

[17] Given the unbalanced panel structure of our data, we explicitly include household fixed effects, which is identical to the "chained difference-in-differences" framework proposed by Bellégo et al. (2025).

that have never implemented the RLCL by 2015 (the never-treated group) as the control group, which allows us to estimate θ_{gs}^{pre}, that is, the pre-treatment effects for individuals in the treatment group g before the RLCL was implemented.

The coefficient θ_{gs}^{post} captures the group-year specific average treatment effect (ATT): it measures the change in outcome (Y_{hpc}), such as income and occupation, in years s while the RLCL reform was implemented in year g. For example, $\delta_{04,04}$ captures the treatment effect of households in 2004 in Hunan province, where the RLCL was implemented in 2004; $\delta_{04,05}$ captures the effect for households, also in Hunan, but in 2005. It is identified by comparing the trends in outcomes of individuals of group g with the trends in outcomes of individuals in never-treated provinces over the same years. Identification of θ_{gs}^{post} hinges on two assumptions. First, the no anticipation assumption requires that there is no effect of treatment prior to the RLCL. This implies that, on average and conditional on fixed effects, potential outcomes prior to treatment are the same. Second, the conditional common trends assumption requires that in the absence of treatment, there would be no differential time trends in outcome between already-treated households and never-treated controls, after conditioning on unit- and time-invariant covariates.

To look at pre-reform and post-reform trends, we aggregate θ_{gs}^{pre} and θ_{gs}^{post} by event time and simple RLCL level.[18] We examine three key outcomes—land rental activity, household income, and village-level share of non-farming households—and observe that the event-study patterns are strikingly similar across all three panels of Figure 1. We observe a shift in both the magnitude and significance of the coefficients after the implementation of the reform. The pattern of increasing effect sizes supports the idea that land security improves after the implementation of the RLCL. Furthermore, there are no significant trends in the outcome before the implementation.

In Panel (a), we first reproduce Figure 2 in Chari et al. (2021), confirming their original results. More critically, however, it remains unclear whether the reform led to an increase in overall household income. To address this gap, we exploit household-level income information in our NFP dataset. Panel (b) presents an event study of total income, showing a sustained increase of just over 3 percent beginning in the second year post-reform. Notably, the estimated average treatment effect on income closely mirrors that on rental activity.

Because income sources are measured with substantial error, we do not disaggregate effects by source; we leave this valuable exercise for future work. Conceptually, gains in total income can arise from both agricultural and non-agricultural activities. High-productivity households

[18] See Section C.2 in Appendix C for a technical description of how the estimates are aggregated.

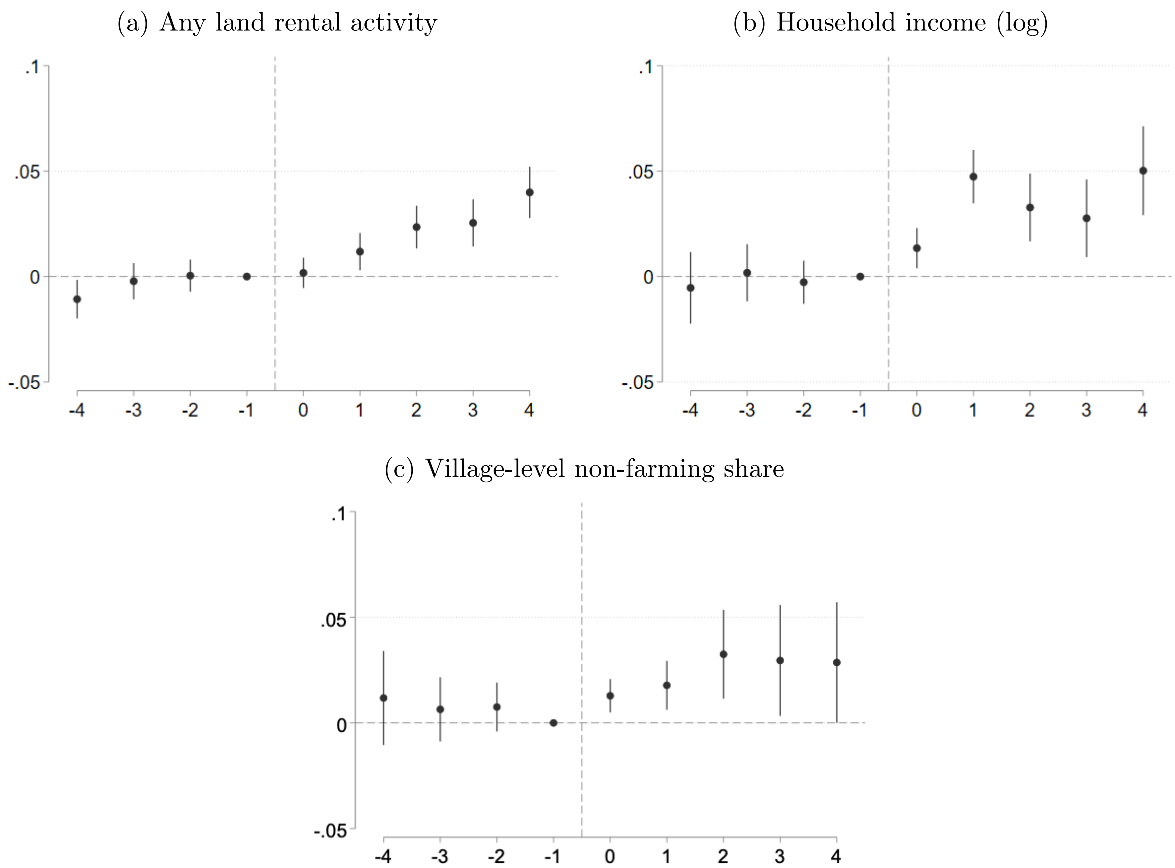


Figure 1: Effects of the RLCL on the first generation

Note: Panels (a) and (b) present the estimated effects of the RLCL on household-level rental activity and income, based on equation (1) estimated at the household level. The outcome variables are defined as an indicator for whether a household engaged in any rental activity in a given year, and the log of annual income, respectively. Panel (c) reports the estimated effects of the RLCL on village-level non-farming household share, based on equation (1) estimated at the village level. The outcome variable is defined as the share of non-farming households in total village households in a given year.

tend to rent in land to expand cultivated acreage, raising farm earnings. By contrast, lower-productivity households often rent out land and shift into non-agricultural employment; for them, higher rental income together with increased urban wages drives the observed rise in total household income.

The theoretical framework in Section 4 implies that land reform should affect the first generation's occupational choices. The ideal test would estimate equation (1) using an individual panel with detailed occupation histories, but the NFP data lack the necessary granularity. We therefore turn to village-level proxy measured by the share of households engaged only in non-agricultural production. Panel (c) of Figure 1 report the corresponding event-study estimates. Consistent with the household-level results, the village-level measure shows a pronounced reallocation of labor out of farming and into non-farm employment.

Following Chari et al. (2021), we also carry out a series of robustness checks by looking at different samples, controlling for confounding shocks, which can be found in Table A10. Overall, we observe an increase in rental activities, household income, and non-agricultural employment showing the real impact of the land reform on the first generation.

5.2 Effects on future workforce

Our primary objective is to assess the effects of the land reform (RLCL) on younger cohorts who are not yet in the labor market, who constitute the future workforce.

5.2.1 Identification

Our identification strategy comes from a key prediction in Section 4 that RLCL should influence education and occupational decisions more strongly among those who could still adjust their educational decisions. In particular, if educational attainment complements non-agricultural productivity, the reform's effects should manifest through altered incentives to pursue high school and college education for the younger cohorts. For older cohorts, they face higher costs to readjust their education, and can only switch occupation conditional on fixed education.

As discussed in Result 2, the land reform increases schooling incentives for two groups: (A) security-only switchers and (B) education-enabled switchers. Because older cohorts' schooling were predetermined at the time of the reform, only type-A switchers among them would switch out of agriculture. By contrast, younger cohorts choose schooling and occupation after the reform. Comparing the younger cohorts with the older ones, therefore, identifies the reform's effect on education for both types and its effect on occupation only for type B, who would only switch occupation if education can be adjusted. Although our theoretical model abstracts to binary educational and occupational choices, the same identification logic applies in richer settings with multiple non-agricultural occupations that vary in their complementarity with education.

We use the age when RLCL was implemented to separate the younger and older cohorts. China mandates nine years of schooling (six years of primary and three years of middle school), and compliance is high. Because job opportunities are scarce for those who do not complete compulsory schooling, the transition to high school is the key educational margin. Children typically start primary school at age six, so the high school enrollment decision is made around age 15 (with some variation). Progression to high school generally requires passing an entrance exam and incurring tuition and foregone earnings. Importantly, while returns to high school alone are relatively low, returns to college education are substantially higher (Li et al., 2012). Students who choose to continue beyond compulsory schooling typically do so with the goal of pursuing higher education, making the high school decision effectively a gateway choice. After this point, returning to school becomes difficult and costly, as students who exit the system face high barriers to re-entry. The age-15 cutoff therefore represents a salient, forward-looking margin through which the RLCL could shift incentives by raising expected returns to continued schooling and access to the college entrance exam.

Formally, we estimate a birth-cohort-based Extended Two-Way Fixed Effects (ETWFE) model that leverages the age-15 breakpoint. We re-index "time" by birth-cohort—defined as the year an individual turns 15—and assign provinces to treatment groups by their RLCL adoption year.[19] Similar to Equation (1), the specification saturates group-by-event-time indicators and includes province and birth-cohort fixed effects to accommodate staggered rollout. Identification comes from comparing adjacent birth-cohorts within provinces that adopt RLCL at different times to the corresponding cohort differences in never-treated counties. In practice, we compare individuals who were younger than 15 at adoption with those just older, across early- and later-adopting provinces, to recover long-run effects on educational and occupational choices. Similarly, the model estimates separate average treatment effects on the treated (ATTs) for each group-cohort combination and is written as follows:

$$Y_{icp} = \sum_{g \in G} \sum_{s=s_0}^{g-1} \theta_{gs}^{pre} D_{igs} + \sum_{g \in G} \sum_{s=g}^{C} \theta_{gs}^{post} D_{igs} + \xi_p + \phi_c + \varepsilon_{icp}, \tag{2}$$

where Y_{icp} is the long-term outcome of individual i who turned 15 in year t and registered in province p. D_{igs} is a dummy that takes the value of 1 if the observation is in the treatment group g and turned 15 in year s and 0 if otherwise. G is a set that indicates at what time the reform started in the province. ξ_p and ϕ_c are province and cohort fixed effects, respectively. θ_{gs}^{post} represents average treatment effect that individuals in treatment group g experiences if turning 15 in year s after the RLCL was implemented (ATT(gs)). We use provinces that have never implemented the RLCL by 2015 (never treated group) at the control group, which allows us to estimate θ_{gs}^{pre}, that is, the pre-treatment effects for individuals in the treatment group g who turn 15 before the RLCL was implemented.

[19] In China, the school year begins in September, and only children who turn seven before September of that year may enroll in primary school. Accordingly, we define a birth cohort as individuals born between September of year t and August of year t+1, who reach age 15 during school year t. The implementation timing of the RLCL is adjusted on the same basis to ensure that treatment status aligns with the schooling calendar relevant for the high school enrollment decision.

The coefficient θ_{igs}^{post} is the group–cohort average treatment effect on the treated (ATT). It captures the causal change in outcome Y_{ipc} (e.g., educational and occupational choices) for individuals who turn 15 in cohort s in provinces that implemented RLCL in year g, relative to the counterfactual without reform. For example, $\theta_{04,04}^{post}$ is the effect for individuals who turned 15 in 2004 in Hunan (RLCL adopted in 2004), whereas $\theta_{04,05}^{post}$ is the effect for Hunan individuals who turned 15 in 2005. Identification comes from comparing cohort-to-cohort changes in outcomes within provinces with adoption year g to the corresponding cohort changes in never-treated provinces over the same years. Similarly, we aggregate θ_{igs}^{pre} and θ_{igs}^{post} by event time and simple RLCL level to look at pre-reform and post-reform trends or understand the magnitude of the impact.

Upon marriage, women often lose their land use rights in their birth village but gain new rights in their husband's village. Land contracts are issued at the household level and signed by the household head, and daughters who marry out are no longer part of their natal household contract (Sargeson, 2012). The RLCL only guarantees that divorced or widowed women who return to their birth villages are entitled to land there. This institutional setup means that land insecurity imposes different mobility constraints by gender: while sons risk losing valuable land rights when they migrate or pursue non-agricultural careers, daughters face fewer such constraints since their future land access depends primarily on marriage rather than maintaining ties to their natal village. Due to these gender differences in how land rights security affect mobility decisions, we conduct our analysis of educational outcomes separately by gender throughout.

5.2.2 Impact on education and occupational choices

In Figure 2, we present the event study estimates calculated using equation (2) for all six outcomes, for males and females separately. The first two outcomes focus on education choices. The last four outcomes focus on their occupation choices, where we further restrict the sample to individuals who turn 15 in or before 2007 to observe their post-education outcomes (at least 23 in 2015). Due to this restriction, the sample size was reduced to 76,610 (71,704) from 98,581 (92,303) for males (females), among which 66,851 (49,555) were working at the time of the survey. Due to the same reason, we are unable to report the post-treatment effects for event times larger than three.

Estimates for positive event times are post-treatment effects, while estimates for negative

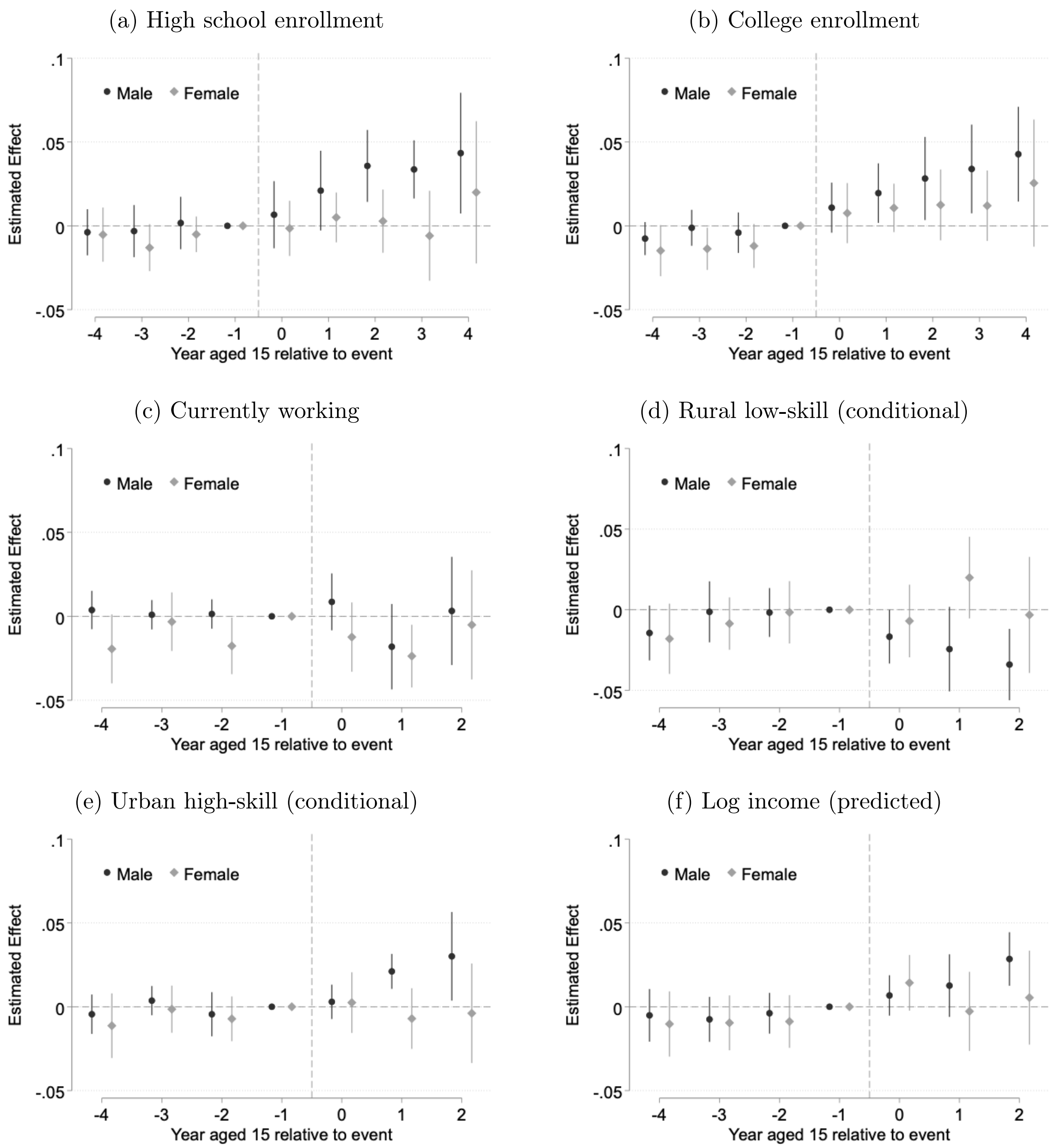


Figure 2: Effects of the RLCL on the younger generation

Note: Figure 2 reports the effects of the RLCL on the outcomes of younger generation, based on the equation (2) for both genders. The sample includes individuals who turned age 15 between 2000 and 2010 and whose *hukou* is registered in a village. Panels (a) and (b) report the effects on high school and college enrollment, respectively. Panels (c)–(f) present the effects on labor market outcomes. Currently working is defined as whether an individual was employed in 2015. Rural low-skilled is defined as whether an individual is employed in farming or construction, while urban high-skilled refers to being employed in high-skilled occupations as defined in Section 3.1. Income is measured using predictions from the 2020 CFPS data.

event times are pre-treatment effects. Four points are worth noting. First, for children who were already 15 when the RLCL was implemented, the cohort trends in all six outcomes are similar in the treated and never-treated provinces. Such parallel pre-treatment trends provide support for the identification assumption of the ETWFE estimator. Second, for males, effect sizes generally rise with event time — except for the outcome variable "currently working" — which is consistent with education and occupational choices adjusting to improved land security after the RLCL. Cohorts who turned 15 in the implementation year show smaller effects, likely because they had less time to respond.[20] Third, we find no effects for females. This is potentially because upon marriage, women often lose their land use rights from their birth village but receive new land use rights from their husband's village. Hence, land use rights when young matter less for women's education and occupation outcomes.

The aggregated ATT estimates reported in columns 1 and 2 of Table A6 indicate that being exposed to the RLCL at or before 15 raises the probability of high school and college enrollment by 0.03. The effects are large, representing 8% of the mean for high school enrollment, 18% of the mean for college enrollment. The similar magnitude of effects across high school and college enrollment is particularly revealing. Given that college admission in China is highly competitive and constrained by quotas, only high-ability students can successfully enter college. The returns to college are also substantially higher than the returns to high school only (Li et al., 2012). The fact that exposure to the RLCL at age 15 increases high school and college enrollment by similar amounts suggests that the affected students are not those who would achieve marginally more education than middle school due to the RLCL. Rather, they are high-ability rural children who would have succeeded academically if given the opportunity. These are students with high returns to education who, absent the RLCL's protections, would have been forced to forgo their educational potential to remain near family land, representing a significant misallocation of talent.

Columns 3-6 of Table A6 show the effects of RLCL at age 15 on occupational outcomes and income. As we explain in Section 4, comparing occupations across two generations detects changes for education-enabled switchers (group (B)), themselves a subset of education compliers (group (A) and (B)). Consistent with this, we find a smaller effect on occupational choice. Among cohorts who turned 15 before the RLCL, the probability of entering a rural low-skill occupation, which include farming and construction jobs, falls by 2 percentage points (6% of the mean), about two-thirds the size of the effect on education, in line with the model's prediction. We also observe that the probability of entering high-skill occupations rises by 1 percentage point (10% of the mean), and income increases by 1%. These results confirm that education-enabled switchers leaving low-skill rural employment are not simply moving to other low-skill positions

[20] The pattern of increasing effect sizes also renders TWFE estimates markedly downward biased; see Table **??** in Appendix A.

Table 1: Effects of the RLCL on education and labor market outcomes

	Education		Labor Market			
	(1) High school	(2) College	(3) Currently working	(4) Rural low-skill	(5) Urban high-skill	(6) Log income (predicted)
Panel A. Males						
RLCL at age 15	0.029***	0.030***	-0.002	-0.022**	0.013***	0.011*
	(0.009)	(0.010)	(0.010)	(0.009)	(0.005)	(0.006)
Mean dep var	0.403	0.170	0.873	0.364	0.125	10.522
Number of clusters	29	29	29	29	29	29
Observations	98581	98581	76610	66851	66851	66851
Panel B. Females						
RLCL at age 15	0.003	0.014	-0.016*	0.003	-0.002	0.007
	(0.009)	(0.009)	(0.008)	(0.009)	(0.008)	(0.007)
Mean dep var	0.382	0.178	0.691	0.344	0.152	9.994
Number of clusters	29	29	29	29	29	29
Observations	92303	92303	71704	49555	49555	49555

Note: Table 1 reports the effects of the RLCL on education and labor market outcomes corresponding to Figure 2, based on the equation (2). The sample includes individuals who turned age 15 between 2000 and 2010 and whose *hukou* is registered in a village. Columns (1) and (2) report the effects on high school and college enrollment, respectively. Columns (3)–(6) present the effects on labor market outcomes. Currently working is defined as whether an individual was employed in 2015. Rural low-skilled is defined as whether an individual is employed in farming or construction, while urban high-skilled refers to being employed in high-skilled occupations as defined in Section 3.1. Income is measured using predictions from the 2020 CFPS data. Panel A reports the results for males. Panel B reports the results for females. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$. Standard errors in parentheses clustered at the province level.

but are accessing high-skill jobs that better match their abilities.

5.2.3 Robustness checks

In this section, we conduct two additional robustness checks that provide further support for the identification assumptions. In addition, we conduct a test on potential sample selection due to *hukou* relocation.

Validation of using age 15—We tested the validity of using age 15 as the exposure age in the analysis. The identification assumption is that age 15 is the cutoff for children to adjust their educational choice since this is the age at which they must decide whether to enroll in high school. Thus, age 15 is defined as the exposure age in the main specification. Exposures at other ages risk misclassification and may lead to biased estimates. In particular, for ages older than 15, children are less likely to adjust their education compared to the scenario when they were at age 15, for several reasons, for example, they are less likely to still be enrolled in school, less likely to return to school, and face higher adjustment costs. At younger ages than 15, children are still subject to compulsory schooling, which restricts their ability to adjust educational choices. Hence, using a higher exposure age would incorrectly assign treatment status to province–cohorts that were not actually treated, thereby attenuating the estimated effect. Likewise, using a younger exposure age would exclude children who were in fact treated from the treatment group while incorrectly including them in the control group. In both cases, the misclassification leads to biased estimates. To test this hypothesis, we replicate the main regression analysis in equation (2) using hypothetical exposure ages ranging from 15 to 18. Table A7 shows the estimated results for the boys sample. The estimates show that the largest positive effects on educational outcomes occur at ages 15, while exposures at older ages greatly attenuate the effects. For labor market outcomes, the strongest effects appear at age 15 (and age 16 only for the effect on urban high skill), whereas exposures at other ages either weaken the effect or even reverse its sign as the estimations misclassify the treatment status.

Control for agricultural tax reform—One common threat to difference-in-differences estimation is the presence of other reforms during a similar period that may also affect the outcomes. During the implementation of the RLCL, all provinces began phasing out the agricultural tax starting in 2004. However, this should not be a concern in our case, as provinces in the sample eliminated the tax in different years, and the aggregate effect should be absorbed by the year fixed effects in the model. Nevertheless, as a robustness check, we control for the agricultural tax rate in the main specification. Specifically, we include exogenously the covariate (Tax_{pt}) in equation (1). The results are reported in Table A10. As anticipated, our previous results are robust to including the tax rate. For the effects on the younger cohorts, we test the equation (2) controlling for the agricultural taxes (Tax_{ps}) in provinces at the year when children turned 15 years old. The results are reported in the table A5.

Placebo on the urban sample—First, we conduct a placebo test using the urban sample, constructed in the same way as our main sample but including only residents who live in cities or towns. This allows us to test whether there are any other province-level policy changes that might have changed the cohort trends in the outcomes of interest. The ATT effects on the urban sample are shown in Table A8, where we find no such evidence. As expected, the RLCL reform has no effect on any outcomes for urban residents.

Hukou relocation—Our analysis sample is defined based on where one's *hukou* is registered. However, the RLCL may also affect *hukou* relocation, especially if the RLCL allows individuals to maintain their land rights even when relocating their *hukou* registration for study or work. To test this possibility, we use the whole sample of individuals from the relevant cohorts and test whether the RLCL changes the probability of having a village *hukou*. Figure A1 presents the event study plots by gender. We find that relative to being exposed to the RLCL at age 16, early exposure does not affect the probability of having a village *hukou*. However, we find some evidence that exposure to the RLCL at age 16 induced some village *hukou* holders, especially males, to relocate their *hukou* to urban areas. Therefore, our main estimates fail to capture the effects of the RLCL on this group of people, who both migrate and change their *hukou* status. Changing one's *hukou* from a village to a city or town usually requires the person to have access to college education or a high-skill urban employment, suggesting that these people should have even higher potential for education, migration, and high-skill employment. If this is the case, our results should provide a lower bound for the true effects of the RLCL.

Results using the NFP data—We use the NFP data for analyzing the effects on household-level outcomes, as well as heterogeneity by agricultural productivity in the next section. To verify that the NFP sample is representative of the census sample, we replicate the main analysis using the NFP sample. As shown in Figure A2, we find very similar results for educational outcomes, except that the estimates are larger in magnitude. We also find some effects on girls' high school enrollment, but no effect on college enrollment.

5.3 Heterogeneity by comparative advantage

An important implication of our simple model is that enhancing land-tenure security will have heterogeneous effects across a range of outcomes—namely, first-generation occupational choices, educational attainment, and second-generation occupational choices—depending on households' comparative advantage in agricultural versus non-agricultural activities. This prediction contrasts with alternative explanations, such as the notion that land reform merely raises household income, which in turn boosts education levels among the younger cohorts.

5.3.1 Measure of comparative advantage

To test the above prediction, we draw on the NFP dataset to construct each family's pre-reform agricultural productivity—measured by the marginal product of land (MPL)—and use it as a proxy for their comparative advantage.[21] In addition to the assumptions laid out in Sections 5.1 and 5.2, our heterogeneity analysis relies on further assumptions. First, the pre-reform MPL we construct must be a valid proxy for each family's true comparative advantage. Second, MPL should not be systematically correlated with other drivers of the increase in the schooling of the

[21] The detailed procedure for calculating MPL follows Chari et al. (2021) and is summarized in Appendix D.

second generation.

To shed light on these concerns, we present two sets of descriptive evidence. The first set shows the relationship between pre-reform MPL and household income before and after the RLCL. The second set shows the relationship between pre-reform MPL and high school attendance for the first and second generations by gender. We use individuals who were at least 16 when the RLCL was implemented to represent the first generation, and individuals who turn age 15 after the RLCL, the second generation. We only explore within-village variation in MPL by normalizing MPL with respect to village mean.

Panels (a) and (b) of Figure 3 display the correlation between log household income and MPL. There is no gradient of income in MPL before and after the RLCL. This suggests that the observed productivity gaps do not simply reflect unobserved differences in land quality or bad luck in production across farms. What is more likely is that low-productivity households invest less in agriculture but more in other activities, given their comparative advantage. Indeed, panels (c) and (d) of Figure 3 show that households with low pre-reform MPL earn a larger fraction of agricultural income and a smaller fraction of wage income before the RLCL. These findings suggest that our MPL measure reflects a comparative advantage in agricultural versus non-agricultural activities.

Figure 4 shows the relationship between high school attendance for the first and second generations by gender. In Panels (a) and (c), we find a very small correlation between educational attainment and agricultural productivity in the first generation for both genders. When looking at educational attainment in the second generation in panels (b) and (d), the correlation changes such that people from families with low MPL have the highest educational attainment, but only if they are boys. Since the second generation turns age 15 after the RLCL, they are able to adjust their educational choice in response to the decrease in land insecurity risk. The boys from low-agricultural-productivity households respond the most. As for girls, the relationship between high school attendance and MPL does not change, which is consistent with the fact that girls' future migration and occupational choices does not directly affect the family's long-term land security.

Overall, these descriptive patterns suggest that boys from low-MPL households, who have a comparative advantage in non-agricultural activities, were constrained in their educational choice before the RLCL. As the RLCL reduces land insecurity risk, these boys are able to re-optimize their educational and occupational choices according to their comparative advantage,

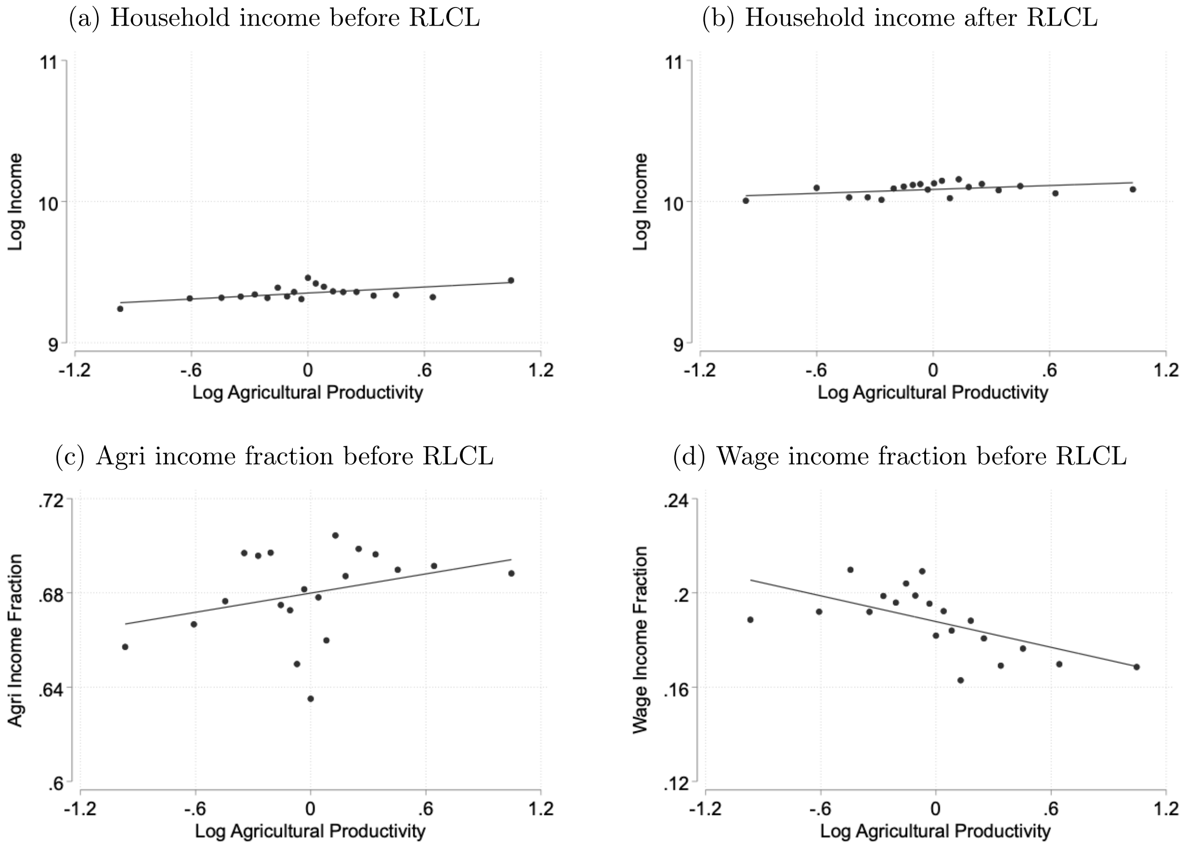


Figure 3: Agricultural productivity and income

Note: This figure presents the relationship between agricultural producitivity and household income before and after RLCL. We focus on households in provinces that implemented RLCL between 2003 and 2006 and use income data in the years of 2000-2002 as years before RLCL, and income data in the years of 2007-2009 as years after RLCL. The y-axis is the average log income and the x-axis is the deviation in log agricultural productivity relative to village mean. The solid line is the estimated linear relationship between log agricultural producitivity and log household income.

leading to an improvement in talent allocation.

5.3.2 Heterogeneous impact by comparative advantage in agriculture

We now conduct the formal analysis using equation (2). Our findings are presented in Table 2, where we split the NFP sample into low- and high-MPL groups within village based on pre-reform agricultural productivity. Columns (1)–(5) report estimates for the first generation from Equation (1).

In Column (1), we confirm the result from Section 5.1: land reform raises household income by about 3% in both productivity groups. Crucially, however, the source of that increase differs. Low-MPL families respond by shifting labor into non-agricultural activities—especially in urban areas—whereas high-MPL families expand their labor supply within agriculture. This divergence is consistent with Chari et al. (2021), who show that high-MPL households tend to rent in

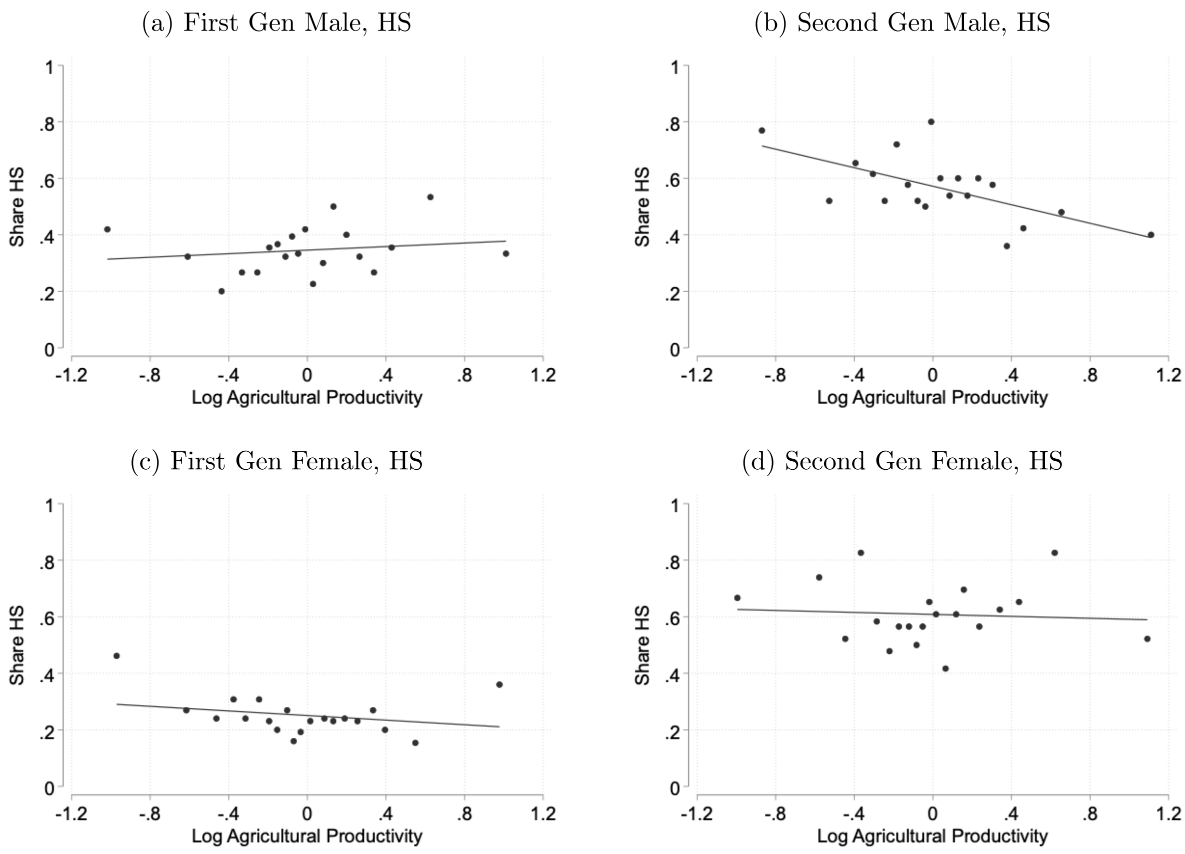


Figure 4: Agricultural productivity and education for the first and second generations

Note: This figure presents the relationship between agricultural productivity and high school and college attendance for two different generations. We focus on individuals living in provinces that implemented RLCL between 2003 and 2006 and use those born between 1985 and 1987 to represent the first generation and those born between 1992 and 1994 as the second generation. The first generation was at least 16 when the RLCL took effect, and the second generation turned 15 after the RLCL. The y-axis is the average high school attendance rate in the upper panel and college attendance rate in the lower panel. The x-axis is the deviation in log agricultural productivity before 2003 relative to the village mean. The solid line is the estimated linear relationship between log agricultural productivity and HS and college attendance.

additional land while low-MPL households are more likely to rent land out.

In Columns (6)–(7), we shift to the second generation using the cohort-based specification in Equation (2). Consistent with our model's prediction, boys from low-MPL families experience a substantially larger gain in educational attainment following land reform. Due to sample limitations, we are unable to explore corresponding heterogeneity in their occupational choices. On the other hand, girls from low- and high-MPL families experience a similar, though insignificant, increase in educational attainment following the reform. The small effects on girls might be driven by the increase in household income that eases their financial constraints.

Table 2: Heterogeneous impacts by household productivity

	First generation					Second generation	
		Labor days per capita					
	(1) Household Income (log)	(2) Agriculture	(3) Non-agri	(4) Non-agri Urban	(5) Non-agri Rural	(6) High School Male	(7) High School Female
Panel A: Low-prodctivity households							
ATT	0.025**	0.196	2.675**	2.086*	0.590**	0.245***	0.073
	(0.012)	(1.170)	(1.141)	(1.114)	(0.300)	(0.056)	(0.071)
Dep. Var. Mean	9.430	79.921	35.154	29.146	6.008	0.433	0.397
Clusters	8825	8792	8792	8792	8792	28	28
Observations	92253	75984	75984	75984	75984	1878	1739
Panel B: High-prodctivity households							
ATT	0.023**	3.192**	-0.234	-0.019	-0.215	0.086*	0.059
	(0.012)	(1.239)	(1.072)	(1.059)	(0.325)	(0.052)	(0.083)
Dep. Var. Mean	9.476	80.741	31.661	25.295	6.367	0.437	0.415
Clusters	8212	8176	8176	8176	8176	28	28
Observations	86589	71358	71358	71358	71358	1803	1565
Diff p-value	0.969	0.064	0.049	0.135	0.075	0.044	0.721

Notes: This table presents the effects of the land reform on the outcomes, conditional on pre-reform household productivity. Column (1) reports the effect of the land reform on household income, estimated using equation (1), with standard errors clustered at the household level. Columns (2) through (5) present the effects of the land reform on labor supply, also estimated using equation (1). The outcome variables are measured as the number of days worked per capita among household members aged 25 and above. Standard errors are clustered at the household level. Means of the outcome variables are calculated based on the values of the sample in 2003. Household productivity is proxied by the marginal product of land (MPL) in 2003. Columns (6) and (7) report the effects on high school and college attainment, respectively, estimated using equation (2). Standard errors are clustered at the province level. Panel A presents results for households with low MPL, while Panel B shows results for households with high MPL.* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

5.3.3 Robustness checks

To make sure our results are not driven by the way of calculating MPL, we implement some robustness checks on the production function estimates. Table A14 shows the analysis using household-crop TFP as a measure of productivity instead of MPL. The results remain qualitatively similar.

6 Conclusion

This paper examines how labor mobility frictions due to insecure land rights can distort educational investment decisions and talent allocation. Using China's Rural Land Contracting Law as a natural experiment, we find that providing legal protection for land use rights when rural residents migrate for study and work significantly increases high school and college enrollment among boys by 3 percentage points (8% of the mean for high school and 18% for college). These educational gains translate into higher migration rates by 3 percentage points and increased probability of skilled employment by 1 percentage point. The affected children are high-ability ones who can succeed in school and in the labor market if given the opportunity, but would have been forced to drop out after middle school if the RLCL had not been implemented, representing a significant misallocation of talent.

Our findings carry two important implications. First, the results demonstrate that labor mobility frictions can have profound long-term consequences when they affect educational decisions early in life. While existing studies primarily focus on how such frictions distort occupational choices among individuals already in the labor market, our evidence shows that the most significant welfare costs may arise from deterring human capital investments before labor market entry. Young people anticipating future mobility constraints may rationally underinvest in education, leading to persistent misallocation of talent. Without accounting for distortions in educational investment decisions, existing studies on mobility frictions and labor misallocation may underestimate both the efficiency losses and the impacts on long-term aggregate productivity due to such frictions.

Second, our results help explain the persistent urban-rural gap in high school and college enrollment that characterizes many developing countries. A few studies have examined the role of agricultural land and production in affecting educational investments in rural children. In many cases, rural children face a trade-off between educational advancement and staying in the countryside. For example, Jensen & Miller (2017) finds that when returns to education increase in urban areas, rural parents may strategically underinvest in their children's education to keep the children down on the farm. Our findings suggest that land policies that enhance mobility and reduce the opportunity cost of leaving rural areas—such as legal land rental markets that enable households to maintain income from their land while children migrate, and secure land tenure that protects rights during temporary migration—could help address the distortions in human capital investments that disproportionately affect rural children and break the intergenerational persistence of urban-rural inequality.

References

Adamopoulos, T., Brandt, L., Chen, C., Restuccia, D., & Wei, X. (2024). Land security and mobility frictions. *The Quarterly Journal of Economics*, *139*(3), 1941–1987.

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

Atkin, D. (2016). Endogenous skill acquisition and export manufacturing in Mexico. *American Economic Review*, *106*(8), 2046–85.

Bellégo, C., Benatia, D., & Dortet-Bernadet, V. (2025). The chained difference-in-differences. *Journal of Econometrics*, *248*, 105783.

Benjamin, D., & Brandt, L. (2002). Property rights, labour markets, and efficiency in a transition economy: the case of rural China. *Canadian Journal of Economics/Revue canadienne d'économique*, *35*(4), 689–716.

Borusyak, K., Jaravel, X., & Spiess, J. (2024). Revisiting event-study designs: Robust and efficient estimation. *The Review of Economic Studies*, *91*(6), 3253-3285.

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

Celik, M. A. (2023). Does the cream always rise to the top? the misallocation of talent in innovation. *Journal of Monetary Economics*, *133*, 105–128.

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

Chavas, J.-P., & Nauges, C. (2020). Uncertainty, learning, and technology adoption in agriculture. *Applied Economic Perspectives and Policy*, *42*(1), 42–53.

Chen, T., Gu, J., Ngai, L. R., & Wang, J. (2025, June). *Sowing seeds of mobility: The uneven impact of land reform* (Discussion Paper No. DP20360). CEPR. (Also STEG Working Paper No. 119)

De Janvry, A., Emerick, K., Gonzalez-Navarro, M., & Sadoulet, E. (2015). Delinking land rights from land use: Certification and migration in Mexico. *American Economic Review*, *105*(10), 3125–3149.

Gai, Q., Guo, N., Li, B., Shi, Q., & Zhu, X. (2024). Rural pensions, labor reallocation, and aggregate income: An empirical and quantitative analysis of China. *HKU Jockey Club Enterprise Sustainability Global Research Institute Paper*(2025/002).

Gong, B. (2018). Agricultural reforms and production in China: Changes in provincial production function and productivity in 1978–2015. *Journal of Development Economics*, *132*, 18–31.

Hsieh, C.-T., Hurst, E., Jones, C. I., & Klenow, P. J. (2019). The allocation of talent and us economic growth. *Econometrica*, *87*(5), 1439–1474.

Hsieh, C.-T., & Klenow, P. J. (2009). Misallocation and manufacturing tfp in China and India. *The Quarterly journal of economics*, *124*(4), 1403–1448.

Jensen, R., & Miller, N. H. (2017). *Keepin' 'em down on the farm: Migration and strategic investment in children's schooling* (Working Paper No. 23122). National Bureau of Economic Research.

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

Li, H., Liu, P. W., & Zhang, J. (2012). Estimating returns to education using twins in urban China. *Journal of Development Economics*, *97*(2), 494-504.

Liu, S., Ma, S., Yin, L., & Zhu, J. (2023). Land titling, human capital misallocation, and agricultural productivity in China. *Journal of Development Economics*, *165*, 103165.

Lo Bello, S., & Morchio, I. (2022). Like father, like son: Occupational choice, intergenerational persistence and misallocation. *Quantitative Economics*, *13*(2), 629–679.

Lu, F., Sun, W., & Wu, J. (2023). Special economic zones and human capital investment: 30 years of evidence from China. *American Economic Journal: Economic Policy*, *15*(3), 35–64.

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

Nakamura, E., Sigurdsson, J., & Steinsson, J. (2022). The gift of moving: Intergenerational consequences of a mobility shock. *The Review of Economic Studies*, *89*(3), 1557–1592.

Ngai, L. R., Pissarides, C. A., & Wang, J. (2018, 09). China's mobility barriers and employment allocations. *Journal of the European Economic Association*, *17*(5), 1617-1653. Retrieved from https://doi.org/10.1093/jeea/jvy035 doi: 10.1093/jeea/jvy035

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

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

Reimers, M., & Klasen, S. (2013). Revisiting the role of education for agricultural productivity. *American Journal of Agricultural Economics*, *95*(1), 131–152.

Restuccia, D., & Rogerson, R. (2017). The causes and costs of misallocation. *Journal of Economic Perspectives*, *31*(3), 151–174.

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

Sargeson, S. (2012). Why women own less, and why it matters more in rural China's urban transformation. *China Perspectives*, 35–42. Retrieved from https://journals.openedition.org/chinaperspectives/6025 doi: 10.4000/chinaperspectives.6025

Tombe, T., & Zhu, X. (2019). Trade, migration, and productivity: A quantitative analysis of China. *American Economic Review*, *109*(5), 1843–1872.

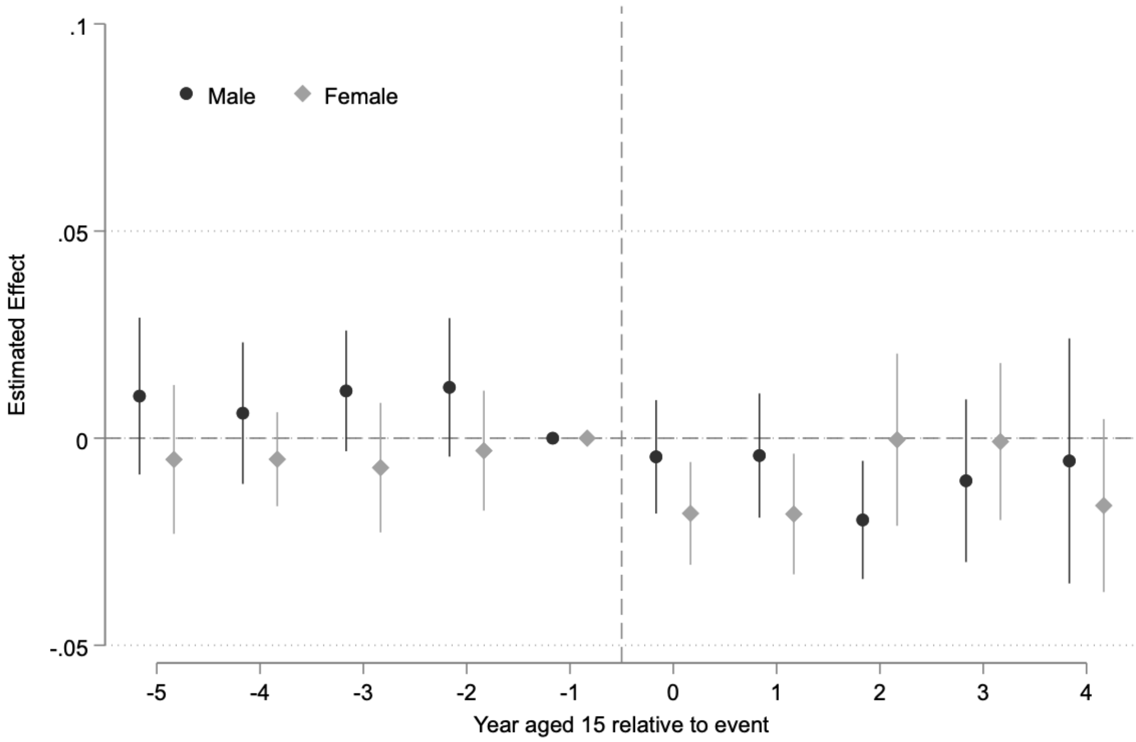
Wooldridge, J. M. (2021). *Two-way fixed effects, the two-way mundlak regression, and difference-in-differences estimators* (SSRN Scholarly Paper). SSRN.

Xu, C. (2011). The fundamental institutions of China's reforms and development. *Journal of Economic Literature*, *49*(4), 1076–1151. doi: 10.1257/jel.49.4.1076

Zhu, S., Xu, X., Ren, X., Sun, T., Oxley, L., Rae, A., & Ma, H. (2016). Modeling technological bias and factor input behavior in China's wheat production sector. *Economic Modelling*, *53*, 245–253.

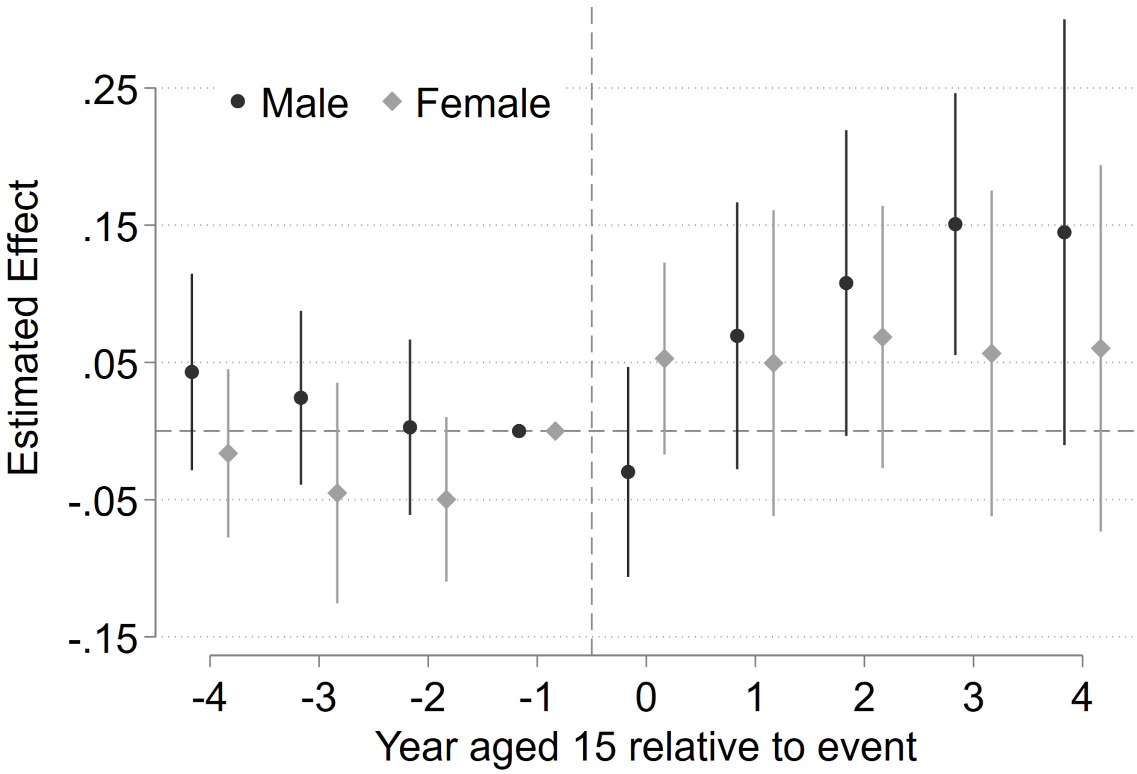
Appendix A Additional Figures and Tables

Figure A1: Effects of the RLCL on *hukou* type



Note: This figure presents the effects of the RLCL on individuals' hukou status in 2015, based on data from the 2015 Census. The hukou type is coded as 1 if an individual holds a village hukou, and 0 if they hold an urban hukou. The sample consists of individuals who turned 15 between 2000 and 2010 and whose hukou is in a village.

Figure A2: The Effects of the RLCL on high school enrollment using NFP



Note: This figure reports the results of replicating the main results on high school enrollment from the 2015 Census by using 2013 NFP for both genders, based on the equation (2) at the individual level.

Table A1: Province-level implementation of the Rural Land Contracting Law

Province	Document Name	Announcement Date	Effective Date
Shanghai	Hu Fu Fa (2003) No. 29	04/25/2003	04/25/2003
Hunan	Hunan Province People's Congress Standing Committee (2004) No. 35	07/30/2004	10/01/2004
Shandong	Shandong Province People's Congress Standing Committee (2004) No. 37	07/30/2004	10/01/2004
Anhui	Anhui Province People's Congress Standing Committee (2005) No. 57	06/17/2005	10/01/2005
Fujian	Fujian Province People's Congress Standing Committee (Min Chang (2005) No. 18)	09/30/2005	11/01/2005
Jiangsu	Jiangsu Province Government Order (2003) No. 21	12/18/2003	02/01/2004
Jilin	Jilin Province People's Congress Standing Committee (2005) No. 29	01/20/2005	03/01/2005
Liaoning	Liaoning Province People's Congress Standing Committee (2005) No. 28	01/28/2005	04/01/2005
Shanxi	Shanxi Province People's Congress Standing Committee (2004) No. 117	09/25/2004	01/01/2005
Tianjin	Jin Zheng Fa (2005) No. 009	02/05/2005	02/05/2005
Xinjiang	Xinjiang People's Congress Standing Committee (2005) No. 24	07/29/2005	10/01/2005
Gansu	Gansu Province People's Congress Standing Committee (Ganzheng Ban Fa (2006) No. 92)	05/29/2006	08/03/2006
Guangxi	Gui Zheng Ban Fa (2006) No. 141	11/14/2006	11/14/2006
Hainan	Hainan Province People's Congress Standing Committee (2006) No. 44	07/28/2006	10/01/2006
Sichuan	Sichuan Province People's Congress Standing Committee (2007) No. 110	11/29/2007	03/01/2008
Yunnan	Yunnan Province People's Congress Standing Committee (2006) No. 41	07/28/2006	09/01/2006
Chongqing	Chongqing Municipality People's Congress Standing Committee (2007) No. 6	04/02/2007	07/01/2007
Jiangxi	Jiangxi Province People's Congress Standing Committee (2007) No. 102	07/27/2007	10/01/2007
Shaanxi	Shaanxi Province People's Congress Standing Committee (2006) No. 59	09/28/2006	01/01/2007
Zhejiang	Zhejiang Province People's Congress Standing Committee (2006) No. 59	09/30/2006	01/01/2007
Inner Mongolia	Inner Mongolia People's Congress Standing Committee (2009) No. 10	07/30/2009	10/01/2009
Qinghai	Qinghai Province People's Congress Standing Committee (2009) No. 15	11/10/2009	03/01/2010
Hebei	Hebei Province Rural Land Contracting Regulation	07/25/2013	11/01/2013
Hubei	Hubei Province Rural Land Contracting Management Regulation	07/27/2012	10/01/2012

Note: Some provinces have not announced any regulation before 2014: Beijing, Heilongjiang, Henan, Guangdong, Guizhou, Tibet, and Ningxia.

Table A2: Rural and Agricultural *Hukou* Share by Occupation Category

Occupation Category	Rural Rate (%)	Agricultural Hukou Rate (%)
Agricultural Workers	86.38	98.70
Construction Workers	58.41	87.32
Low-skilled Service Workers	23.72	51.32
Low-skilled Manufacturing Workers	41.22	63.59
Other Low-skilled Workers	31.42	59.74
High-skilled Workers	17.77	26.86

Notes: Sample restricted to individuals aged 25-50. Rural rate represents the percentage of rural population within each occupation category. Agricultural hukou rate represents the percentage of agricultural household registration within each occupation category. Occupation categories based on 3-digit occupation codes with remapping to align with 2015 census classifications. Data source: CFPS 2010.

Table A3: Summary Statistics for Census Variables

Variables	Count	Mean	Std
Panel A: Boys Sample			
Birth year	98581	1989.39	3.04
Ever attending high School	98581	0.40	0.49
Ever attending college	98581	0.17	0.38
Currently Working	76610	0.87	0.33
Rural low skill occupation	66851	0.36	0.48
Urban high skill occupation	66851	0.12	0.33
Predicted income (log)	66851	10.52	0.35
	Count	Mean	Std
Panel B: Girls Sample			
Birth year	92303	1989.41	3.02
Ever attending high School	92303	0.38	0.49
Ever attending college	92303	0.18	0.38
Currently Working	71704	0.69	0.46
Rural low skill occupation	49555	0.34	0.48
Urban high skill occupation	49555	0.15	0.36
Predicted income (log)	49555	9.99	0.48

Note: This table reports the summary statistics for the Census sample. A low-skill job is defined as farming or low-skill construction. A high-skill job is defined based on the share of college graduates in an occupation, with occupations having at least 25% college graduates classified as high-skill. Income is estimated using a gender-specific regression, as reported in Appendix A. Panel A presents the summary statistics for the boys' sample, and Panel B presents those for the girls' sample.

Table A4: Robustness checks: standard TWFE and ETWFE with not-yet-treated as controls

	Education		Labor Market			
	(1)	(2)	(3)	(4)	(5)	(6)
	High school	College	Currently working	Rural low-skill	Urban high-skill	Log income (predicted)
Panel A.TWFE, Males						
RLCL at age 15	0.014	0.013	-0.006	-0.022***	0.008*	0.011*
	(0.011)	(0.010)	(0.008)	(0.006)	(0.004)	(0.005)
Mean dep var	0.403	0.170	0.873	0.364	0.125	10.522
Number of clusters	29	29	29	29	29	29
Observations	98581	98581	76610	66851	66851	66851
Panel B. TWFE, Females						
RLCL at age 15	0.005	0.016	0.001	0.011	0.002	0.004
	(0.008)	(0.010)	(0.007)	(0.007)	(0.009)	(0.007)
Mean dep var	0.382	0.178	0.691	0.344	0.152	9.994
Number of clusters	29	29	29	29	29	29
Observations	92303	92303	71704	49555	49555	49555
Panel C. ETWFE with not-yet-treated, Males						
RLCL at age 15	0.019*	0.020*	-0.008	-0.023***	0.011**	0.012***
	(0.011)	(0.011)	(0.008)	(0.005)	(0.004)	(0.004)
Mean dep var	0.403	0.170	0.873	0.364	0.125	10.522
Number of clusters	29	29	29	29	29	29
Observations	98581	98581	76610	66851	66851	66851
Panel D. ETWFE with not-yet-treated, Females						
RLCL at age 15	0.007	0.020*	0.000	0.013**	0.002	0.002
	(0.009)	(0.011)	(0.007)	(0.007)	(0.007)	(0.006)
Mean dep var	0.382	0.178	0.691	0.344	0.152	9.994
Number of clusters	29	29	29	29	29	29
Observations	92303	92303	71704	49555	49555	49555

Note: This table reports the results of robustness tests on the main specification. Panel A and Panel B report the results of the two-way fixed effects model for boys and girls respectively. Panels C and D present the results of the extended two-way fixed effects specification, using both the not-yet-treated and never-treated groups as the control group. Standard errors in parentheses clustered at the province level. The sample includes individuals who turned age 15 between 2000 and 2010 and whose *hukou* is registered in a village. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

Table A5: Effects of the RLCL on education and labor market outcomes, controlling for tax reforms

	Education		Labor Market			
	(1) High school	(2) College	(3) Currently working	(4) Rural low-skill	(5) Urban high-skill	(6) Log income (predicted)
Panel A. Males						
RLCL at age 15	0.029***	0.030***	-0.002	-0.020**	0.014***	0.012*
	(0.010)	(0.011)	(0.010)	(0.008)	(0.005)	(0.006)
Mean dep var	0.403	0.171	0.873	0.364	0.125	10.521
Number of clusters	29	29	29	29	29	29
Observations	97810	97810	75839	66223	66223	66223
Panel B. Females						
RLCL at age 15	0.003	0.014	-0.017*	0.004	0.001	0.007
	(0.009)	(0.009)	(0.008)	(0.010)	(0.007)	(0.007)
Mean dep var	0.383	0.179	0.691	0.345	0.152	9.994
Number of clusters	29	29	29	29	29	29
Observations	91619	91619	71020	49108	49108	49108

Note: This table reports the effects of the RLCL on education and labor market outcomes, estimating the equation (2) controlling for the agricultural tax rate. The sample includes individuals who turned age 15 between 2000 and 2010 and whose *hukou* is registered in a village. Rural low-skilled is defined as whether an individual is employed in farming or construction, while urban high-skilled refers to being employed in high-skilled occupations as defined in Section 3.1. Income is measured using predictions from the 2020 CFPS data. Standard errors in parentheses clustered at the province level. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

Table A6: Effects of the RLCL on education and labor market outcomes, older sample

	Education		Labor Market		
	(1) High school	(2) College	(3) Rural low-skill	(4) Urban high-skill	(5) Log income (predicted)
Panel A. Males					
RLCL at age 15	0.028**	0.019**	-0.022**	0.013***	0.011*
	(0.014)	(0.006)	(0.009)	(0.005)	(0.006)
Mean dep var	0.356	0.129	0.364	0.125	10.522
Number of clusters	29	29	29	29	29
Observations	66851	66851	66851	66851	66851
Panel B. Females					
RLCL at age 15	0.004	0.007	0.003	-0.002	0.007
	(0.009)	(0.008)	(0.009)	(0.008)	(0.007)
Mean dep var	0.342	0.146	0.344	0.152	9.994
Number of clusters	29	29	29	29	29
Observations	49555	49555	49555	49555	49555

Note: This table reports the effects of the RLCL on education and labor market outcomes for older sample. The older sample is defined as individuals born between 1951 and 1979 and whose *hukou* is registered in a village. Standard errors in parentheses clustered at the province level. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

Table A7: The Effects of RLCL on the Outcomes at Different Exposure Ages

	(1)	(2)	(3)	(4)
Age	15	16	17	18
Panel A.Effect on High School				
ATT	0.029***	0.012	0.013*	0.011
	(0.009)	(0.010)	(0.008)	(0.009)
Observations	98581	98781	99448	100294
Panel B.Effect on College				
ATT	0.030***	0.020***	0.010**	0.017**
	(0.010)	(0.007)	(0.005)	(0.008)
Observations	98581	98781	99448	100294
Panel C.Effect on Currently Working				
ATT	-0.002	0.007	0.002	-0.011***
	(0.010)	(0.005)	(0.008)	(0.003)
Observations	76610	75812	74522	71340
Panel D.Effect on Rural Low-skill				
ATT	-0.022**	-0.010*	0.003	0.010
	(0.009)	(0.006)	(0.011)	(0.009)
Observations	66851	67098	66639	64317
Panel E.Effect on Urban High-skill				
ATT	0.013***	0.019***	0.001	-0.001
	(0.005)	(0.005)	(0.008)	(0.010)
Observations	66851	67098	66639	64317
Panel F.Effect on Income (Log)				
ATT	0.011*	0.009	0.006	0.006
	(0.006)	(0.006)	(0.010)	(0.009)
Observations	66851	64585	57132	46989

Note: This table reports group–birth–cohort–specific treatment effects of the RLCL reform on outcomes for the boys sample for the exposure ages between 15 and 18, estimated using Equation 2 at the corresponding age cutoffs. Column (1) reports the effect at the exposure age 15. Column (1) presents the effect at exposure age 15, Column (2) at exposure age 16, Column (3) at exposure age 17, and Column (4) at exposure age 18. Standard errors are clustered at the province level. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

Table A8: The Effects of the RLCL on the Outcomes for Urban Sample

	Education		Labor Market			
	(1)	(2)	(3)	(4)	(5)	(6)
	High school	College	Currently working	Rural low-skill	Urban high-skill	Log income (predicted)
Panel A. Males						
RLCL at age 15	0.007	0.020	-0.013	0.000	0.003	-0.011
	(0.018)	(0.018)	(0.009)	(0.008)	(0.010)	(0.009)
Mean dep var	0.735	0.479	0.793	0.053	0.351	10.797
Number of clusters	29	29	29	29	29	29
Observations	45052	45052	34730	27543	27543	27543
Panel B. Females						
RLCL at age 15	0.018*	0.018	-0.003	0.003	0.012	0.010
	(0.010)	(0.011)	(0.014)	(0.007)	(0.014)	(0.009)
Mean dep var	0.727	0.502	0.655	0.063	0.444	10.421
Number of clusters	29	29	29	29	29	29
Observations	46316	46316	36505	23910	23910	23910

Notes: This table reports the placebo test—the effects of the RLCL on the main outcomes for the urban sample. The sample includes individuals who turned age 15 between 2000 and 2010 and who are registered in urban areas. Columns (1) and (2) report the effects on high school and college enrollment, respectively. Columns (3)–(6) present the effects on labor market outcomes. Currently working is defined as whether an individual was employed in 2015. Rural low-skilled is defined as whether an individual is employed in farming or construction, while urban high-skilled refers to being employed in high-skilled occupations as defined in Section 3.1. Income is measured using predictions from the 2020 CFPS data. Panel A reports the results for males. Panel B reports the results for females. Standard errors in parentheses clustered at the province level. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

Table A9: Summary Statistics for NFP Variables

	(1) Count	(2) Mean	(3) Std
Panel A: Variables by Household-year			
Area (mu)	232094	18.07	1007.21
Annual Income	213990	9.92	0.85
Any New Land Renting	232117	0.12	0.33
Panel B: Variables by Village-year			
Total population	4370	1958.42	1448.43
Total number of out-migrants	3760	2.18	0.97
Nonfarm share	4357	0.08	0.15
Panel C: Variables by Individual			
Gender	10218	0.53	0.50
Birth year	10218	1989.47	2.94
High School	9967	0.44	0.50
College	9967	0.32	0.47

Notes: This table provides summary statistics for the NFP samples. Panel A reports the household data. Annual household income is expressed in real 2001 RMB (deflated using the 2001 CPI) and reported in logarithmic form. Panel B reports village-level data. The nonfarm share is defined as the ratio of nonfarm households to farming households within a village. Income from land expropriation refers to the revenue earned by the village through the expropriation of farmland, measured in real 2001 RMB. Panel C reports the information on individual data.

Table A10: Effects of the RLCL on the first generation

	(1) Renting	(2) Income	(3) Non-farm share
Panel A: Full Sample			
ATT	0.029***	0.037***	0.027**
	(0.004)	(0.008)	(0.013)
Observations	221234	202618	4190
Dep Var Mean	0.120	9.889	0.079
Clusters	22747	22285	361
Panel B: Include the agricultural tax rate			
ATT	0.031***	0.040***	0.026**
	(0.004)	(0.008)	(0.012)
Observations	218546	200163	4122
Dep Var Mean	0.120	9.894	0.080
clusters	22403	21952	356
Panel C: Balanced Sample			
ATT	0.038***	0.035***	
	(0.006)	(0.010)	
Observations	128735	118488	
Dep Var Mean	0.116	9.859	
Clusters	10327	10244	

Notes: This table reports the estimation corresponding to Figure 1. Columns (1) and (2) present the estimated effects of the RLCL on household-level rental activity and income, based on equation (1) estimated at the household level. The outcome variables are defined as an indicator for whether a household engaged in any rental activity in a given year, and the log of annual income, respectively. Columns (3) and (4) report the estimated effects of the RLCL on village-level non-farming household share and migration, based on equation (1) estimated at the village level. The outcome variables are defined as the share of non-farming households in total village households in a given year, and the log of the number of migrants in the village in a given year, respectively. Panel A reports the estimation for the full sample households between 2000 and 2013. Panel B reports the results of the estimation using equation (1) controlling for agricultural tax rates. Panel C reports the results of the restricted samples that were present in all periods. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

Table A11: Production Function Coeficients by Crop

	Sowing Area	Labor Days	Machinery Cost	Other Cost
Wheat	0.536***	0.154***	0.083***	0.257***
	(0.021)	(0.014)	(0.007)	(0.019)
Rice	0.554***	0.144***	0.020***	0.238***
	(0.023)	(0.014)	(0.004)	(0.020)
Corn	0.630***	0.179***	0.021***	0.278***
	(0.017)	(0.013)	(0.004)	(0.015)
Soybean	0.549***	0.276***	0.077***	0.145***
	(0.016)	(0.016)	(0.008)	(0.013)
Potato	0.546***	0.275***	0.021**	0.077***
	(0.022)	(0.019)	(0.010)	(0.012)
Other Grains	0.495***	0.211***	0.078***	0.167***
	(0.031)	(0.029)	(0.014)	(0.031)
Cotton	0.672***	-0.080**	0.021**	0.352***
	(0.055)	(0.038)	(0.010)	(0.036)
Oilseed	0.531***	0.276***	0.026***	0.146***
	(0.015)	(0.014)	(0.006)	(0.011)
Sugar	0.336***	0.529***	0.076***	0.235***
	(0.084)	(0.101)	(0.027)	(0.045)
Tobacco Leaf	0.976***	-0.767***	0.095***	0.344***
	(0.093)	(0.085)	(0.022)	(0.074)
Silkworm	0.374***	0.299***	0.083***	0.111***
	(0.022)	(0.015)	(0.007)	(0.007)
Vegetables	0.942***	-0.158***	-0.149***	0.164***
	(0.054)	(0.058)	(0.026)	(0.030)
Other Cash Crops	0.240***	0.405***	0.069***	0.188***
	(0.031)	(0.030)	(0.013)	(0.016)
Fruit	0.733***	-0.265***	-0.178***	0.351***
	(0.096)	(0.093)	(0.043)	(0.039)
Other Orchard	0.235***	0.096**	0.209***	-0.016
	(0.039)	(0.044)	(0.021)	(0.026)
Total	0.527***	0.199***	0.045***	0.190***
	(0.009)	(0.007)	(0.003)	(0.006)
Total Observations	53339			

Notes: This table presents the estimation of the production function for a particular crop based on the equation (D2) using the NFP data before 2003. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

Table A12: Production function coeficients (balanced panel)

	Sowing Area	Labor Days	Machinery Cost	Other Cost
Wheat	0.578^{***}	0.113^{***}	0.010^{**}	0.176^{***}
	(0.036)	(0.012)	(0.004)	(0.021)
Rice	0.606^{***}	0.130^{***}	0.009^{**}	0.136^{***}
	(0.026)	(0.015)	(0.004)	(0.018)
Corn	0.650^{***}	0.121^{***}	0.003	0.148^{***}
	(0.028)	(0.014)	(0.003)	(0.015)
Soybean	0.562^{***}	0.199^{***}	0.016^{***}	0.090^{***}
	(0.020)	(0.016)	(0.006)	(0.010)
Potato	0.505^{***}	0.169^{***}	0.005	0.095^{***}
	(0.020)	(0.016)	(0.007)	(0.011)
Other Grains	0.607^{***}	0.130^{***}	0.013	0.142^{***}
	(0.030)	(0.024)	(0.009)	(0.019)
Cotton	0.838^{***}	0.023	-0.007	0.109^{***}
	(0.035)	(0.025)	(0.007)	(0.027)
Oilseed	0.556^{***}	0.176^{***}	0.009^{**}	0.123^{***}
	(0.021)	(0.014)	(0.004)	(0.011)
Sugar	0.809^{***}	0.127	0.070^{***}	0.075
	(0.115)	(0.094)	(0.022)	(0.069)
Tobacco Leaf	0.525^{***}	0.174^{**}	0.009	0.315^{***}
	(0.112)	(0.068)	(0.015)	(0.096)
Silkworm	0.233^{***}	0.226^{***}	0.039^{*}	0.184^{***}
	(0.077)	(0.059)	(0.020)	(0.054)
Vegetables	0.233^{***}	0.264^{***}	0.048^{***}	0.107^{***}
	(0.034)	(0.015)	(0.006)	(0.007)
Other Cash Crops	0.387^{***}	0.251^{***}	0.070^{***}	0.061^{***}
	(0.029)	(0.026)	(0.016)	(0.012)
Fruit	0.235^{***}	0.418^{***}	0.024^{**}	0.130^{***}
	(0.039)	(0.030)	(0.009)	(0.013)
Other Orchard	0.286^{***}	0.551^{***}	0.046	0.078^{***}
	(0.099)	(0.076)	(0.062)	(0.023)
Total	0.527^{***}	0.169^{***}	0.017^{***}	0.130^{***}
	(0.012)	(0.006)	(0.002)	0.006)
Total Observations	137334			

Notes: This table presents the estimation of the production function for a particular crop using the balanced sample of household-crop combinations in NFP between 2003 and 2013, based on estiamting the equation (D2) and (D3). * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

Table A13: Effects by productivity

	Full Smple		Sub Sample		Low Productivity		High Productivity	
	(1) HS	(2) College	(3) HS	(4) College	(5) HS	(6) HS	(7) HS	(8) HS
Panel A: Boys sample								
RLCL at age 15	0.085**	0.068**	0.149***	0.130***	0.245***	0.265***	0.086*	0.066
	(0.039)	(0.029)	(0.037)	(0.037)	(0.056)	(0.053)	(0.052)	(0.051)
Mean dep var	0.446	0.323	0.435	0.308	0.433	0.434	0.437	0.436
Number of clusters	28	28	28	28	28	28	28	28
Observations	5303	5303	3687	3687	1878	1876	1802	1803
Panel B: Girls sample								
RLCL at age 15	0.056	0.028	0.070	0.052	0.075	0.084	0.056	0.061
	(0.042)	(0.047)	(0.067)	(0.074)	(0.070)	(0.065)	(0.083)	(0.082)
Mean dep var	0.431	0.326	0.406	0.298	0.397	0.400	0.416	0.413
Number of clusters	28	28	28	28	28	28	28	28
Observations	4660	4660	3305	3305	1738	1757	1564	1544

Notes: This table presents the effects of the land reform on the full sample and a subsample of productivity using NFP. Columns (1) and (2) present the results using the NPF full sample, replicating the result from the cohort analysis. Columns (3) and (4) estimate the effects of the land reform on outcomes using the subsample for which we have pre-reform productivity information. Columns (5) and (7) show the estimated effects of the land reform on high school education for low- and high-MPL households in 2003, respectively. Columns (6) and (8) provide the corresponding results when pre-reform MPL from the production function is estimated using data from 2000 to 2010, which serve as a robustness check. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

Table A14: Effects by household-crop level TFP

	First generation					**Second generation**	
		Labor days per capita					
	(1) Household Income (log)	(2) Agriculture	(3) Non-agri	(4) Non-agri Urban	(5) Non-agri Rural	(6) High School (Male)	(7) High School (Female)
Panel A: Low-prodctivity households							
ATT	0.036***	-0.155	2.877***	2.802***	0.075	0.133***	0.108
	(0.012)	(1.136)	(1.021)	(1.001)	(0.279)	(0.043)	(0.069)
Dep. Var. Mean	9.413	81.410	33.193	27.546	5.647	0.429	0.418
Clusters	8884	8883	8883	8883	8883	28	28
Observations	93919	78835	78835	78835	78835	2304	2031
Panel B: High-prodctivity households							
ATT	0.026**	1.016	-0.220	0.046	-0.266	0.058	0.033
	(0.013)	(1.303)	(1.221)	(1.210)	(0.334)	(0.077)	(0.057)
Dep. Var. Mean	9.477	84.004	31.672	25.675	5.996	0.437	0.395
Clusters	6934	6933	6933	6933	6933	28	28
Observations	74705	62870	62870	62870	62870	1878	1702
Diff p-value	0.533	0.064	0.093	0.472	0.593	0.560	0.515

Notes: This table presents the effects of the land reform, conditional on pre-reform household productivity. Column (1) reports the effect of the land reform on household income, estimated using equation (1), with standard errors clustered at the household level. Columns (2) through (5) present the effects of the land reform on labor supply, also estimated using equation (1). The outcome variables are measured as the number of days worked per capita among household members aged 25 and above. Standard errors are clustered at the household level. Household productivity is proxied by the household-crop TFP in 2003. Means of the outcome variables are calculated based on the sample values in 2003. Columns (6) and (7) report the effects on high school for boys and girls, respectively, estimated using equation (2). Standard errors are clustered at the province level. Panel A presents results for households with low TFP, while Panel B shows results for households with high TFP.* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

Table A15: Effects on Income, Education and Labor Outcomes by Huzhu's education

	First generation					Second generation	
	(1)	(2)	(3)	(4)	(5)	(6)	(7)
	Incomes	Agri	Non-agri	Non-agri urban	Non-agri rural	HS Boys	HS Girls
Panel A: Low Educational Households							
	0.046***	-2.171	0.874	0.005	0.869**	0.062	0.130**
	(0.013)	(1.321)	(1.186)	(1.166)	(0.342)	(0.038)	(0.061)
Dep. Var. Mean	9.389	84.308	33.486	27.480	6.006	0.418	0.372
Clusters	8583	8158	8158	8158	8158	28	27
Observations	86362	70387	70387	70387	70387	1983	1786
Panel B: High Educational Households							
	0.019	1.134	1.285	1.654	-0.369	0.184***	0.032
	(0.012)	(1.108)	(1.141)	(1.115)	(0.337)	(0.061)	(0.083)
Dep. Var. Mean	9.552	75.086	34.483	27.577	6.906	0.494	0.488
Clusters	8192	7811	7811	7811	7811	27	28
Observations	82572	67759	67759	67759	67759	1859	1633

Notes: This table presents the effects of the land reform on the outcomes, conditional on the educational level of the head of the householde. Column (1) reports the effect of the land reform on household income, estimated using equation (1), with standard errors clustered at the household level. Columns (2) through (5) present the effects of the land reform on labor supply, also estimated using equation (1). The outcome variables are measured as the number of days worked per capita among household members aged 25 and above. Standard errors are clustered at the household level. Means of the outcome variables are calculated based on the sample values in 2003. Columns (6) and (7) report the effects on high school for boys and girls, respectively, estimated using equation (2). Panel A presents results for households with low educational level, while Panel B shows results for households with high educational level. Standard errors are clustered at the province level.* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

Appendix B Conceptual Framework: proofs

B.1 Proof of Result 1

B.1.1 Effects on income for the first generation

The first part of Result 1 states that income weakly decreases with η for individuals in generation $g-1$. We prove these by discussing the different scenarios depending on land rental behaviors and occupational choices.

Farmers renting in land ($\ell^{rent} > 0$)—For households with sufficiently high agricultural productivity $z_F^{g-1}(i)$ that optimally rent in land, the effective rental price is $q(\eta, \ell^{\text{rent}}(i)) = q_0 + \alpha\eta$ when $\ell^{\text{rent}}(i) > 0$. The optimization problem becomes:

$$\max_{\ell^{\text{rent}}(i) \geq 0} \left\{ p_a A_a z_F^{g-1}(i)[\bar{\ell}(i) + \ell^{\text{rent}}(i)]^{\theta} - (q_0 + \alpha\eta)\ell^{\text{rent}}(i) \right\}$$

By the envelope theorem:

$$\frac{\partial I_F^{g-1}}{\partial \eta} = \frac{\partial}{\partial \eta}\left[-(q_0 + \alpha\eta)\ell^{\text{rent}*}\right] = -\alpha\ell^{\text{rent}*}$$

Since $\ell^{\text{rent}*} > 0$ for land renters-in and $\alpha > 0$, we have $\frac{\partial I_F^{g-1}}{\partial \eta} < 0$.

Farmers renting out land ($\ell^{rent} < 0$) —For households that optimally rent out land ($\ell^{\text{rent}*} < 0$), the effective rental price is $q(\eta, \ell^{\text{rent}}(i)) = q_0 - \eta\varphi(i)$. The optimization problem becomes:

$$\max_{\ell^{\text{rent}}(i) \leq 0} \left\{ p_a A_a z_F^{g-1}(i)[\bar{\ell}(i) + \ell^{\text{rent}}(i)]^{\theta} - (q_0 - \eta\varphi(i))\ell^{\text{rent}}(i) \right\}$$

By the envelope theorem:

$$\frac{\partial I_F^{g-1}}{\partial \eta} = \frac{\partial}{\partial \eta}\left[-(q_0 - \eta\varphi(i))\ell^{\text{rent}*}\right] = \varphi(i)\ell^{\text{rent}*}$$

Since $\ell^{\text{rent}*} < 0$ for land renters-out and $\varphi(i) > 0$, we have $\frac{\partial I_F^{g-1}}{\partial \eta} < 0$.

Manufacturing Workers —For manufacturing workers, all land is rented out at the base rate q_0, but they face the direct risk of income loss. Their income is:

$$I_M^{g-1}(i) = w_M z_M^{g-1}(i)(1 + \tau\bar{h}) + q_0\bar{\ell}(i) - \eta\varphi(i)\bar{\ell}(i) - c$$

where $\bar{h}$ is the educational choice they made in the first period.

The marginal effect of land insecurity is:

$$\frac{\partial I_M^{g-1}}{\partial \eta} = -\varphi(i)\bar{\ell}(i) < 0$$

By revealed preference, $I^*(i,\eta) = \max\{I_F^{g-1}(i,\eta), I_M^{g-1}(i,h^*,\eta)\}$. We consider three cases:

1. Manufacturing worker: If $I^*(i,\eta) = I_M^{g-1}(i,h^*,\eta)$, then $\frac{\partial I^*}{\partial \eta} = -\varphi(i)\bar{\ell}(i) < 0$.

2. Farming with land rental-in: If $I^*(i,\eta) = I_F^{g-1}(i,\eta)$ with $\ell^{\text{rent}*} > 0$, then $\frac{\partial I^*}{\partial \eta} = -\alpha\ell^{\text{rent}*} < 0$.

3. Farming with land rental-out: If $I^*(i,\eta) = I_F^{g-1}(i,\eta)$ with $\ell^{\text{rent}*} < 0$, then $\frac{\partial I^*}{\partial \eta} = \varphi(i)\ell^{\text{rent}*} < 0$.

In all cases where households actively participate in land rental markets, the income effect is strictly negative. Only when $\ell^{\text{rent}*} = 0$ throughout, we have $\frac{\partial I^*}{\partial \eta} = 0$.[22]

B.1.2 Effects on occupation choice for the first generation

The second part of Result 1 states that the share of individuals choosing manufacturing in generation $g-1$ decreases with land insecurity. We prove this by showing that for any individual at the margin of occupational choice, an increase in η makes manufacturing relatively less attractive and hence decides to switch to farming. Consider an individual that is indifferent between sectors: $I_F^{g-1}(i,\eta) = I_M^{g-1}(i,h^*,\eta)$.

Taking the derivative with respect to η:

$$\frac{\partial I_M^{g-1}}{\partial \eta} - \frac{\partial I_F^{g-1}}{\partial \eta} = -\varphi(i)\bar{\ell}(i) - \varphi(i)\ell^{\text{rent}*} < 0$$

when $-\bar{\ell}(i) < \ell^{\text{rent}*} < 0$ (the marginal individual typically rents out land but cannot rent out more than endowment). As manufacturing income reduces faster than farming income, more

[22]The farmer prefers not to rent out land when the marginal product of their own land is greater than or equal to the effective rental income:

$$p_a A_a z_F^{g-1}(i)\theta[\bar{\ell}(i)]^{\theta-1} \geq q_0 - \eta\varphi(i)$$

This can be rewritten as:

$$z_F^{g-1}(i) \geq \frac{(q_0 - \eta\varphi(i))[\bar{\ell}(i)]^{1-\theta}}{p_a A_a \theta}$$

When $\eta \geq q_0/\varphi(i)$, the effective rental rate becomes non-positive, and farmer i prefers not to rent out land, resulting in $\frac{\partial I^*}{\partial \eta} = 0$.

individuals will switch from manufacturing to farming as η increases. This establishes that $\frac{\partial \mathrm{Pr}^{g-1}(\mathrm{M})}{\partial \eta} < 0$. $\square$

B.2 Proof of Result 2

B.2.1 Cross-generation difference in educational effects

The first part of result 2 states that an increase in η decreases the share of education in generation g but not in $g-1$. Child i in generation g chooses education if

$$\begin{aligned}
I_M^g(i,1,\eta) - I_M^g(i,0,\eta) &= w_M z_M^g(i)(1+\tau) + q_0\bar{\ell}(i) - \eta\varphi(i)\bar{\ell}(i) - c \\
&\quad - [w_M z_M^g(i) + q_0\bar{\ell}(i) - \eta\varphi(i)\bar{\ell}(i)] \\
&= w_M z_M^g(i)\tau - c > 0
\end{aligned}$$

The educational choice is independent of η for a given occupation. Since h does not enter the agricultural income function, the share of educated individuals in generation g is given by:

$$\mathrm{Pr}^g(h=1,\eta) = \mathrm{Pr}^g(h=1|M) \times \mathrm{Pr}^g(M)$$

Taking the derivative with respect to η:

$$\frac{\partial \mathrm{Pr}^g(h=1)}{\partial \eta} = \mathrm{Pr}^g(h=1|M) \times \frac{\partial \mathrm{Pr}^g(M)}{\partial \eta}$$

Since: $\frac{\partial \mathrm{Pr}^g(M)}{\partial \eta} < 0$ (from Result 1), $\frac{\partial \mathrm{Pr}^g(h=1)}{\partial \eta} = \mathrm{Pr}^g(h=1|M) \times \frac{\partial \mathrm{Pr}^g(M)}{\partial \eta} \leq 0$, and $\frac{\partial \mathrm{Pr}^g(h=1)}{\partial \eta} \leq \frac{\partial \mathrm{Pr}^{g-1}(\bar{h}=1)}{\partial \eta} = 0$, with strict inequality if $\mathrm{Pr}^g(h=1|M) \neq 0$.[23]

B.2.2 Cross-generation difference in occupation switching effects

The second part of result 2 states that generation g has a steeper manufacturing share response than generation $g-1$: $\frac{\partial \mathrm{Pr}^g(M)}{\partial \eta} < \frac{\partial \mathrm{Pr}^{g-1}(M)}{\partial \eta} < 0$. We prove this by showing that generation g has additional switching that generation $g-1$ lacks. A key assumption we need is that the distributions of skills (z_F and z_M) are constant across generations.

Common Switching Effect—Both generations experience the same *occupational switching* for individuals with the same skill levels, regardless of their education. Consider individuals in set

[23] This quires that some individuals satisfy $w_M z_M^g(i)\tau - c > 0$ and $I_M^g(i,1,\eta) > I_F^g(i,\eta)$ for the relevant range of η. With sufficient heterogeneity in skills, reasonably high education premium, and non-extreme range of η, we can assume that some individuals will satisfy these conditions.

$\mathcal{A}$ who satisfy the following condition for some η:

$$\mathcal{A} = \left\{ i : \begin{array}{l} I_M^g(i,\eta,1) > I_M^g(i,\eta,0) > I_F^g(i,\eta) \text{ and} \\ w_M z_M^g(i)\tau > c \end{array} \quad \text{for the relevant range of } \eta \right\}$$

When η decreases, those who find manufacturing more profitable than farming even with $h = 0$ would switch to manufacturing no matter which generation they belong to.

Additional Switching in Generation g—Generation g has an additional source of switching through the educational margin. Consider individuals in set $\mathcal{B}$ who satisfy the following condition for some η:

$$\mathcal{B} = \left\{ i : \begin{array}{l} I_M^g(i,\eta,0) \leq I_F^g(i,\eta) \text{ and} \\ I_M^g(i,\eta,1) > I_F^g(i,\eta) \text{ and} \\ w_M z_M^g(i)\tau > c \end{array} \quad \text{for the relevant range of } \eta \right\}$$

For these individuals, education is profitable in manufacturing but farming remains competitive without education. As η decreases (land becomes more secure), individuals in $\mathcal{B}$ will switch from farming to manufacturing with education. The critical switching point occurs when:

$$I_M^g(i,\eta^*,h=1) = I_F^g(i,\eta^*)$$

Let $N_B(\eta)$ denote the number of individuals in set $\mathcal{B}$ who choose manufacturing with education at parameter η:

$$N_A^g(\eta) = |\{i \in \mathcal{B} : I_M^g(i,\eta,h=1) > I_F^g(i,\eta)\}|$$

Because

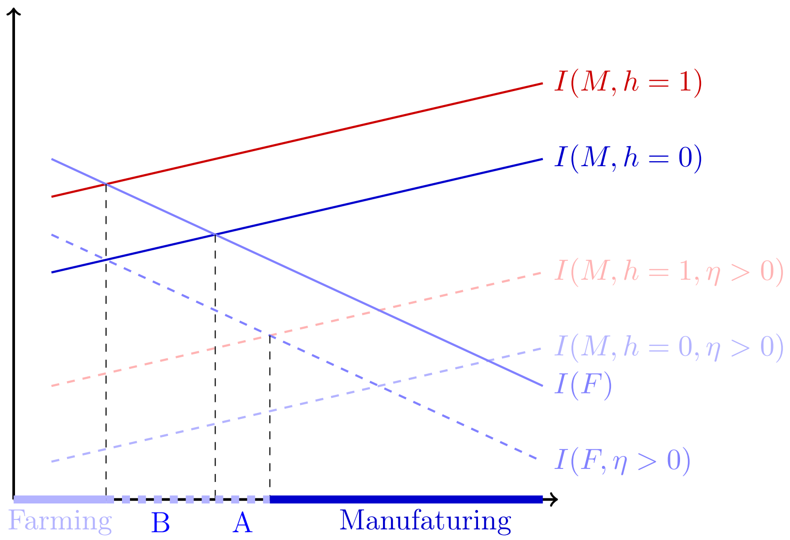
$$\frac{\partial I_M^g(i,\eta,h=1)}{\partial \eta} - \frac{\partial I_F^g(i,\eta)}{\partial \eta} = -\varphi(i)\bar{\ell}(i) - \varphi(i)\ell^{\text{rent}*} < 0$$

then $\frac{\partial N_A^g(\eta)}{\partial \eta} \leq 0$, with strict inequality as long as there are individuals that satisfy the conditions defining set $\mathcal{B}$.

Under the assumption that the distribution of skills is constant across generations, generation g includes all the switching that generation $g-1$ has, plus additional switching from set $\mathcal{B}$:

$$\frac{\partial \Pr^g(M)}{\partial \eta} = \frac{\partial \Pr^{g-1}(M)}{\partial \eta} + \frac{1}{N}\frac{\partial N_S(\eta)}{\partial \eta}$$

Figure B1: Occupational Sorting by Comparative Advantage



Under reasonable parameterization of the model (sufficient heterogeneity in skills and education premium, and intermediate levels of land insecurity), we can assume that set $\mathcal{B}$ is non-empty. Then

$$\frac{\partial \mathrm{Pr}^g(M)}{\partial \eta} < \frac{\partial \mathrm{Pr}^{g-1}(M)}{\partial \eta} < 0$$

□

B.2.3 Who are the switchers?

The proofs above show that there are two groups of switchers who adjust their education, occupation, or both in response to a change in η. To understand who they are, we use Figure B1 to illustrate the economics of the model and the decisions of different individuals. If a worker chooses farming when land is insecure, she earns $I(F, \eta > 0)$ on average. If she chooses manufacturing, she earns $I(M, h = 0, \eta > 0)$ if education is not profitable ($w_M z_M^g(i)\tau < c$), and $I(M, h = 1, \eta > 0)$ if education is affordable ($w_M z_M^g(i)\tau \geq c$). A reform reduces η to zero in period t. This change increases income in both sectors, but manufacturing income rises relatively more, as shown above. As a result, some individuals switch from farming to manufacturing, as illustrated in Figure B1.

The first group of switchers, group A, is defined as

$$\mathcal{A} = \left\{ i : \begin{array}{l} I_M^g(i, \eta, 1) > I_M^g(i, \eta, 0) > I_F^g(i, \eta) \text{ and} \\ w_M z_M^g(i)\tau > c \end{array} \quad \text{for the relevant range of } \eta \right\}$$

People in this group have high skills in manufacturing and low skills in farming. They switch to manufacturing when η is sufficiently low, even without changing their education.

The first group of switchers, group B, is defined as

$$\mathcal{B} = \left\{ i : \begin{array}{l} I_M^g(i,\eta,0) \leq I_F^g(i,\eta) \text{ and} \\ I_M^g(i,\eta,1) > I_F^g(i,\eta) \text{ and} \\ w_M z_M^g(i)\tau > c \end{array} \quad \text{for the relevant range of } \eta \right\}$$

Individuals in this group have medium-level skills in both sectors. They switch to manufacturing only when combined with education; without education, they remain in farming.

Groups A and B thus comprise the switchers, who suffer from talent misallocation due to land rights insecurity. Their decisions differ by generation in response to the change in η:

$$A : \begin{cases} F \to M{+}E & \text{if generation } g, \\ F \to M & \text{if generation } \textit{g-1} \end{cases}$$

$$B : \begin{cases} F \to M{+}E & \text{if generation } g, \\ \text{Stay in } F & \text{if generation } \textit{g-1} \end{cases}$$

The generational difference implies that the effect of a decrease in η on education can be identified by comparing younger cohorts (generation g) with older cohorts (generation $g-1$). Younger cohorts in both groups are treated while still in their first period and can adjust their education; older cohorts cannot. Thus, the education effect reflects the responses of both groups A and B. Using the same framework, we can also identify the effect of a decrease in η on occupation, but this comes only from group B. Because group A would switch to manufacturing regardless of education, only group B switches sectors when they are allowed to change education too. Consequently, the framework identifies a larger effect on education than on occupation, since some older individuals in group A would still move to manufacturing even without education. We therefore refer to group A as pure-security switchers and group B as education-enabled switchers.

B.3 Proof of Result 3

We prove that families with lower agricultural productivity are more responsive to land security improvements in their children's education decisions. We assume that child i imperfectly inherits skills from the parent, i.e. $z_k^g(i) = \pi z_k^{g-1}(i) + (1-\pi)\,\varepsilon_n^g(i), \pi \in [0,1]$.

Child i with agricultural productivity $z_F^{g-1}(i)$ chooses education if manufacturing with education yields higher utility than the best alternative:

$$h^*(i) = 1 \text{ if } I_M^g(i,1,\eta,z_F^{g-1}(i)) > \max\{I_F^g(i,\eta,z_F^{g-1}(i)), I_M^g(i,0,\eta,z_F^{g-1}(i))\}$$

Define the critical level of land insecurity where individual i is indifferent between manufacturing with education and farming:

$$\eta_i^* \text{ such that } I_M^g(i,1,\eta_i^*,z_F^{g-1}(i)) = I_F^g(i,\eta_i^*,z_F^{g-1}(i))$$

For $\eta < \eta_i^*$, individual i chooses manufacturing with education (assuming $w_M z_M^g(i)\tau > c$).[24] For $\eta > \eta_i^*$, individual i chooses farming.

Taking the total differential of the indifference condition:

$$\frac{\partial I_M^g(i,1,\eta)}{\partial \eta} d\eta + \frac{\partial I_M^g(i,1,\eta)}{\partial z_F^{g-1}(i)} dz_F^{g-1}(i) = \frac{\partial I_F^g(i,\eta)}{\partial \eta} d\eta + \frac{\partial I_F^g(i,\eta)}{\partial z_F^{g-1}(i)} dz_F^{g-1}(i)$$

Since manufacturing income doesn't depend on agricultural productivity: $\frac{\partial I_M^g(i,1,\eta)}{\partial z_F^{g-1}(i)} = 0$.

Rearranging:

$$\frac{d\eta_i^*}{dz_F^{g-1}(i)} = \frac{\frac{\partial I_F^g(i,\eta)}{\partial z_F^{g-1}(i)}}{\frac{\partial I_M^g(i,1,\eta)}{\partial \eta} - \frac{\partial I_F^g(i,\eta)}{\partial \eta}}$$

Since $\frac{\partial I_F^g(i,\eta)}{\partial z_F^{g-1}(i)} > 0$ (higher productivity increases agricultural income) and $\frac{\partial I_M^g(i,1,\eta)}{\partial \eta} - \frac{\partial I_F^g(i,\eta)}{\partial \eta} < 0$ (manufacturing becomes relatively less attractive as η increases), we have:

$$\frac{d\eta_i^*}{dz_F^{g-1}(i)} < 0$$

This means higher agricultural productivity lowers the threshold η_i^*, making individuals less likely to choose manufacturing and education.

Consider a reduction in land insecurity from η_{high} to η_{low} where $\eta_{\text{high}} > \eta_{\text{low}}$. Consider two

[24] When $w_M z_M^g(i)\tau \leq c$, education does not respond to a change in η.

types of families with high and low agricultural productivity (high-prod and low-prod). From the results above, we have

$$\eta^*(\text{high-prod}) < \eta^*(\text{low-prod})$$

The change in education probability for each type is:

$$\Delta\Pr(h = 1|\text{low-prod}) = \Pr(h = 1|\text{low-prod}, \eta_{\text{low}}) - \Pr(h = 1|\text{low-prod}, \eta_{\text{high}})$$

$$\Delta\Pr(h = 1|\text{high-prod}) = \Pr(h = 1|\text{high-prod}, \eta_{\text{low}}) - \Pr(h = 1|\text{high-prod}, \eta_{\text{high}})$$

We consider different scenarios based on the relative positions of η_{high}, η_{low}, and the switching thresholds:

Case 1: $\eta_{\text{high}} > \eta_{\text{low}} > \eta^*(\text{low-prod}) > \eta^*(\text{high-prod})$

Both family types choose farming throughout, so

$$\Delta\Pr(h = 1|\text{low-prod}) = \Delta\Pr(h = 1|\text{high-prod}) = 0$$

We can rule out this case by assuming sufficient heterogeneity in agricultural productivity and reasonable range of η.

Case 2: $\eta_{\text{high}} > \eta^*(\text{low-prod}) > \eta_{\text{low}} > \eta^*(\text{high-prod})$

Low-productivity families switch from farming to manufacturing with education, while high-productivity families remain in farming:

$$\Delta\Pr(h = 1|\text{low-prod}) > 0$$
$$\Delta\Pr(h = 1|\text{high-prod}) = 0$$

Case 3: $\eta_{\text{high}} > \eta^*(\text{low-prod}) > \eta^*(\text{high-prod}) > \eta_{\text{low}}$

Both family types switch from farming to manufacturing+education, but low-productivity families were closer to the margin:

$$\Delta\Pr(h = 1|\text{low-prod}) > 0$$
$$\Delta\Pr(h = 1|\text{high-prod}) > 0$$
$$\Delta\Pr(h = 1|\text{low-prod}) > \Delta\Pr(h = 1|\text{high-prod})$$

The last inequality holds because low-productivity families have a higher switching threshold, making them more responsive to changes in land security.

In all cases where there is any response to land security improvements, low-productivity families respond more strongly:

$$\Delta P(h = 1|\text{low-prod}) > \Delta P(h = 1|\text{high-prod})$$

For small changes in η, this discrete analysis implies:

$$\left|\frac{\partial \mathrm{Pr}^g(h = 1|\text{low-prod})}{\partial \eta}\right| > \left|\frac{\partial \mathrm{Pr}^g(h = 1|\text{high-prod})}{\partial \eta}\right|$$

For small productivity differences, this further implies

$$\frac{\partial^2 \mathrm{Pr}^g(h = 1|z_F^{g-1}(i))}{\partial \eta \, \partial z_F^{g-1}(i)} > 0$$

$\square$

Appendix C Extended Two-way Fixed Effect Model

With homogeneous treatment effects, a common approach to estimating the average treatment effect on the treated (ATT) is to estimate a two-way fixed-effects (TWFE) model:

$$Y_{ipt} = \beta PostRLCL_{pt} + \xi_p + \phi_t + \varepsilon_{ipt}, \tag{C1}$$

where Y_{ipt} outcome of individual i who turned 15 in year t and whose *hukou* is registered in province p. μ_p are province of birth fixed effects, α_t are cohort fixed effects, and $PostRLCL_{ipt}$ is an indicator variable for individual i who turned age 15 after the RLCL was implemented in the province. Under a common trend assumption, β can be interpreted as the causal effect of being exposed to RLCL at age 15, the critical age when deciding whether or not to attend high school.

C.1 ETWFE Model

However, treatment heterogeneity may occur both in the group and time dimensions. For example, effects might be larger in provinces that adopted the reform earlier, as they faced greater pre-reform uncertainty about land rights, or smaller for children who are closer to 15 when the reform started, since they had already made preliminary educational plans. Recent literature on difference-in-differences with staggered treatment adoption shows that in the presence of such treatment heterogeneity, the estimates from the TWFE estimator in equation (C1) may be biased due to 'forbidden comparisons', i.e., the use of already-treated units as controls for later treated units (Borusyak et al., 2024). A range of heterogeneity-robust estimators has been proposed in recent literature to address this issue.

We follow Wooldridge (2021) in extending the traditional two-way fixed effects estimation. Specifically, for the first-generation analysis, we allow for treatment effect heterogeneity across cohorts and over time by interacting cohort dummies of the treated units with period dummies, while using households in the never-treated provinces as the control group. In this way, we obtain group- or cohort-specific average treatment effects, which capture the change in outcomes responding to the implementation of the RLCL reform in a given year for each household in our sample.

Wooldridge (2021)'s baseline model assumes balanced panel data, where controlling for group fixed effects effectively accounts for individual fixed effects. However, in cases of unbalanced panels—such as ours—this equivalence breaks down. To accommodate this, we explicitly control for household fixed effects. Our approach is identical to the "chained difference-in-differences" framework proposed by Bellégo et al. (2025), while retaining the flexibility of Wooldridge (2021).

For the second-generation analysis, we use a birth-cohort-based extension of the two-way fixed effects estimator to examine the effect of the RLCL reform on individual outcomes. We define the birth cohort as the calendar year in which children turn 15 years old and assign treatment status based on the year of RLCL implementation in their province. Following a similar approach to the previous estimation, we interact group-by-event-time dummies with birth-cohort fixed effects and province fixed effects. In this way, we obtain group–birth-cohort-specific average treatment effects of the RLCL reform.

C.2 Estimation Target

The ETWFE estimator estimates treatment effects at a very granular level. However, interest often lies in more aggregated quantities rather than the directly estimated cohort-time-specific treatment effects from equation 1. In this paper, we are interested in evaluating how much change in education attainment and occupational choices, on average, in all province pairs and post-reform years, as a result of the introduction of RLCL. Therefore, our main estimation target is a single estimate, which is defined by the weighted sum of the estimated cohort-time-specific treatment effects as

$$\bar{\hat{\delta}} = \sum_{g=q}^{T} \sum_{s=g}^{T} \frac{N_{gs}}{N_D} \hat{\delta}_{gs},$$

where we assign equal weight to all posttreatment observations, which correspond to the number of observations of cohort g in period s, N_{gs}, relative to the total number of treated observations, $N_D = \sum_{g=q}^{T} \sum_{s=g}^{T} N_{gs}$. Similarly, the standard errors for the ETWFE estimator are computed as a (weighted) linear combination of the cohort-year-specific effects, taking the covariance between the coefficients into account.

Another estimation target is an event-study estimand, for which we compute event-time-specific treatment effects by averaging over the cohort dimension as follows:

$$\bar{\hat{\delta}}_{\cdot s} = \sum_{g=q}^{s} \frac{N_{gs}}{N_{\cdot s}} \hat{\delta}_{gs},$$

where $N_{\cdot s} = \sum_{g=q}^{s} N_{gs}$ is the total number of treated observations in period s. Event-time-specific treatment effects are akin to results from a dynamic TWFE (or event-study) specification without its shortcomings, based on the heterogeneity-robust innovations.

As discussed above, caution needs to be exercised when comparing positional changes due to differences in weights and in the composition of groups/ periods, which can complicate their interpretation. For example, early-treated cohorts have more observations many years after the treatment onset than late-treated cohorts. Consequently, in the presence of treatment effect heterogeneity across cohorts, changes in event-time-specific treatment effects are the result of changes in event time as well as the higher share of early-treated cohorts relative to late-treated cohorts.

For birth-cohort-based ETWFE, the main equation 2, we re-index "time" by birth-cohort, defined as the calendar year an individual turns 15, and assign provinces to treatment groups by their RLCL adoption year. Accordingly, we adjust the formulas for these two estimation targets.

Appendix D Marginal Product of Land (MPL)

We also use NFP data to construct a household-year level measure of the marginal product of land (MPL) across all crops farmed by a household in a year, as an indicator of pre-reform productivity. The MPL is calculated as follows:

$$MPL^{C}_{h,t} = \sum_{c \in C_{ht}} \omega_{cht} \left(\alpha_c \cdot \frac{Y_{cht}}{L_{cht}} \right) \tag{D1}$$

where C denotes the types of crops, h denotes households, and t denotes the year. ω_{cht} denotes the weight where we take the mean of these values for each household–year pair. This measurement enables us to rank households by how efficiently they utilize land at the margin in 2003, prior to the implementation of the land reform.

To estimate α_c, we use household-crop-year level data from 1986 to 2003, excluding 1994 due to missing information. We assume a Cobb-Douglas agricultural production function for each crop. It requires land, labor, machines, and other expenses as input. We measure agricultural yield, sowing area, household labor days for sowing, machinery costs, and other input costs—such as fertilizer expenses—at the household-crop-year level. These variables are incorporated into the household's production function. Then we jointly estimate a logarithm function for each crop following Chari et al. (2021):

$$y_{hcvt} = \alpha_c log L_{hcvt} + \beta_c log N_{hcvt} + \gamma_c log K_{hcvt} + \sigma_c log O_{hcvt} + e_{hcvt} \tag{D2}$$

where y_{hcvt} denotes log physical output of household h growing croptype c in village v in year t. L_{hcvt}, N_{hcvt}, K_{hcvt}, and O_{hcvt} represent the farming area, labor days, machinery cost, and all other input costs, respectively. We absorb ϕ_{hcvt} by controlling for year, household, and crop fixed effect[25]. Our estimation is 0.52 on average, which is slightly larger than the literature, for example, Zhu et al. (2016) and Gong (2018), and similar to the estimation from Liu et al. (2023). Table A11 represents the results of the estimation for each crop.

In addition to the pre-reform MPL, we construct a second measure of household-level productivity as a robustness check using NFP data. Specifically, we estimate household–crop level total factor productivity (TFP) using data from 1986 to 2010, which enables us to further decompose TFP when estimating equation (D2). In particular, we decompose ϕ_{hcvt} as follows:

[25]We relax the assumption for the production function estimation due to limited observations that we have only using pre-reform data.

$$\phi_{hcvt} = \phi_{hc} + \phi_{ht} + \phi_{cvt} + e_{hcvt} \tag{D3}$$

Where ϕ_{hc} represents the household-crop level productivity that we use as fixed ability regarding crops of each household, ϕ_{ht} denotes the household-year fixed effect capturing time-varying shocks given to the household, and ϕ_{cvt} denotes a time-varying component that is common to all households in a village farming crop c. Table A14 represents the result of the effects of the land reform on the outcomes, conditional on this household-crop level total factor productivity ϕ_{hc}. This measure captures a household's average fixed ability to farm a given crop.

We conduct several robustness checks on our production function estimations for the two measures. First, we test whether the household–crop-level TFP estimates are sensitive to assumptions regarding the unobserved component e_{hcvt} in Equation (D3). Specifically, we allow for the possibility that farmers observe production outcomes from previous years and adjust their input choices accordingly. To address this, we instrument inputs with their lagged values in addition to absorbing the fixed effects in Equation (D3). The estimated coefficients are generally consistent with previous results, although less precisely estimated, as the elasticity of land is 0.60, greatly higher than both our earlier estimates using only the fixed effects model and the values reported in the literature on land elasticity in China.

Second, we change the production function form to test the robustness of the Cobb-Douglas function form assumption. Specifically, we test a translog production function that is an approximation of the CES function by a second-order Taylor polynomial. We estimate the following equation:

$$\begin{aligned}
y_{hcvt} = {} & \alpha_0 + \alpha_L \log L_{hcvt} + \alpha_N \log N_{hcvt} + \alpha_K \log K_{hcvt} + \alpha_O \log O_{hcvt} \\
& + \tfrac{1}{2}\beta_{LL}(\log L_{hcvt})^2 + \tfrac{1}{2}\beta_{NN}(\log N_{hcvt})^2 + \tfrac{1}{2}\beta_{KK}(\log K_{hcvt})^2 + \tfrac{1}{2}\beta_{OO}(\log O_{hcvt})^2 \\
& + \beta_{LN}(\log L_{hcvt})(\log N_{hcvt}) + \beta_{LK}(\log L_{hcvt})(\log K_{hcvt}) + \beta_{LO}(\log L_{hcvt})(\log O_{hcvt}) \\
& + \beta_{NK}(\log N_{hcvt})(\log K_{hcvt}) + \beta_{NO}(\log N_{hcvt})(\log O_{hcvt}) + \beta_{KO}(\log K_{hcvt})(\log O_{hcvt}) \\
& + e_{hcvt}
\end{aligned} \tag{D4}$$

This specification allows for a non-linear relationship between input choices and output in the production function. The R-squared of the translog model is 0.9696, which is very similar to the fit of the Cobb–Douglas model (0.9640). Our estimation yields an elasticity of land of 0.53 in the pre-reform years, which is slightly but not meaningfully larger than the earlier estimate. The correlation between the previous TFP residual and the current TFP residual is 0.9638, indicating a very strong consistency.

Third, we examine the sensitivity of household-level selection caused by the entry and exit of households in the unbalanced NFP data. To address this concern, we estimate the Cobb–Douglas production function using a balanced sample of household–crop combinations from 1986 to 2010. The estimated coefficients are very similar to those obtained previously, which helps us rule out the possibility of significant impacts from unobserved productivity shocks across years. Table A12 reports the production function estimates for the balanced household–crop pairs.