Reallocating wealth? Insecure property rights and agricultural investment in rural China*

Jessica Leight

October 19, 2012

Abstract

This paper evaluates the impact of village-level land reallocations in China on household economic outcomes. Since land was decollectivized in China in 1983, village leaders have implemented regular forced reallocations of land designed to increase equity and attain other policy goals. I estimate the impact of insecure tenure using the past history of land shifts as an instrument for current tenure insecurity, and find that an increase in the probability of losing the current plot yields a decrease in agricultural inputs and production of around one standard deviation. Though the costs of insecure tenure are high, structural estimates of the varying cost of reallocation across different villages suggest the choice to reallocate does reflect an optimizing process on the part of village officials, who reallocate where the net benefit is larger. However, the observed pattern of reallocations would be optimal only given an objective function for the village leader that places a high weight on equity, and even given this objective function, there is evidence that village leaders may be making some costly mistakes. JEL codes: Q15

1 Introduction

The establishment of clear land rights has long been considered a key milestone in the development of the modern industrialized countries. Because land is the principal asset in a preindustrial economy, the development of an institutional structure that encourages its efficient use is argued to enhance growth substantially (North & Thomas 1973). Conversely, the absence of stable and enforced property rights is widely identified as a major impediment to growth in today's developing countries (De Soto 2000).

Despite this emphasis on the importance of private property rights, however, collectively owned or managed land remains a widespread phenomenon in the developing

*I thank my advisers Abhijit Banerjee, Esther Duflo and Rohini Pande for their guidance throughout this project. Additional thanks are due to Panle Jia and Stephan Ryan for their feedback, and to Timothy Besley, Melissa Dell, Horacio Larreguy, Jennifer Peck, Laura Ralston, Tavneet Suri and participants at the MIT Development and Political Economy and Harvard China Politics Research Workshop for helpful discussions. All errors are my own. Email: jeleight@mit.edu

world. Collective or partly collective land structures continue to be predominant in rural areas in China, in Mexico, and in many parts of sub-Saharan Africa. These forms of land ownership can yield substantial benefits in terms of equity, but they may also generate significant efficiency costs.

In China, the post-Mao period saw the emergence of a hybrid system of landownership, in which formal title to land is held by the village collective and use rights are held by households. Moreover, plots are subject to periodic reallocations between households conducted by village officials every three to five years, thereby generating systematic insecurity in land tenure. These reallocations represent the outcome of a bargaining process between officials and households that weighs the costs and benefits of the associated disruption in property rights. The objective of this paper is to estimate the economic costs of insecure land tenure induced by these periodic reallocations, and to demonstrate that village officials do respond to variation in these costs in shaping the relative security of local property rights.

First, I examine variation in tenure insecurity within a village conditional on a reallocation being conducted. Households that have recently had their land reallocated are less likely to have their land reallocated in a subsequent round, and accordingly the past history of changes in landownership can be employed as an instrument for the probability of loss of the current plot. The results show that the reduction in the probability of losing the current plot as a result of past inclusion in reallocation, a decrease of around 5% on a base probability of 56%, results in an increase in the use of agricultural inputs and in total agricultural output of between .05 and .1 standard deviations, with no evidence of simultaneous substitution out of non-agricultural activities.

Given that this effect of relatively more secure tenure is evident for households that both gained and lost land in the past, any plausible alternative channel for the observed pattern requires that an unobservable shock correlated with reallocation affects both relatively rich and relatively poor households in the same way relative to the mean. The observed pattern of symmetric increased investments by households at both ends of the land-ownership distribution in the year of a reallocation is inconsistent with most obvious sources of omitted variable bias.

Second, while these reduced form estimates capture a uniform effect of reallocation, the observed variation in the propensity to reallocate across space and time suggests that the costs and benefits of reallocation are in fact far from uniform. It is plausible that officials will choose to reallocate in areas where disruption of land tenure is less costly — more specifically, less costly in terms of investment foregone as a result of insecure land rights.

To test this hypothesis, I estimate an agricultural production function that allows for spatial variation in the returns to agricultural inputs. I find that the propensity to reallocate is negatively correlated with returns to lagged inputs, and with the magnitude of the investments lost as a result of the induced insecurity of tenure. Thus despite the fact that reallocations generate large costs, the observed pattern of reallocations does seem to reflect an optimizing process on the part of the village official.

In the final part of the analysis, I postulate a functional form for the objective function that underlies this optimizing process and seek to estimate parameters for this function that would best reproduce the observed pattern of reallocations. The results suggest first, that reallocations are only optimal for a village leader that places an extremely high weight on increased equity relative to the potential output losses induced by a reallocation. Second, even given an objective function that values equity, village leaders are making some potentially costly errors by reallocating where the relative balance of benefits and costs is unfavorable.

To sum up, the evidence from variation in property rights in China suggests that even incremental shifts in the security of land tenure in a context of partly collective land rights can have large economic implications. In addition, variation in the frequency of reallocations is correlated with their costs, with reallocations occurring most frequently where they are least costly. Thus at a local level, property rights adapt to reflect the relative returns to secure property rights in different economic environments.

This paper supplements an existing literature that has evaluated the impact of varying regimes of property rights in China. Feder, Lau, Lin & Luo (1992) argue based on a before-and-after analysis that excessive investment in nonproductive assets such as housing is evidence of the negative impact of insecure land tenure. Brandt, Huang, Li & Rozelle (2002) analyze the impact of land tenure by comparing households' private plots, assigned permanently to households in some villages for their personal cultivation, with responsibility land that is subject to reallocations. Similarly, de la Rupelle, Quheng, Shi & Vendryes (2009) use household-level heterogeneity in land rights within a village to identify the impact of reallocations on outmigration, finding that insecure land rights induce temporary, rather than permanent, outmigration in order to ensure claims are retained on land left behind. Both papers make the assumption that plots are exogenously assigned to different contractual types within a village.

Jacoby, Li & Rozelle (2002) analyze the impact of insecure tenure on investment in rural China by using a hazard model to estimate predicted risks of expropriation for different plots held. They find that a higher expropriation risk decreases investment in organic fertilizer, employing the identifying assumption that the hazard of reallocation is exogenous to household characteristics.

There is also a larger literature about the economic impact of property rights that evaluates land reforms in which tenants without formal title are endowed with stronger property rights (Banerjee, Gertler & Ghatak 2002, Besley & Burgess 2000). Goldstein & Udry (2008) analyze property rights in Ghana and conclude that individuals with more secure tenure rights by virtue of their more powerful political positions invest more in maintaining soil fertility. Another set of papers focused on urban land policy in Latin America finds that land titling increases labor supply and investment (Besley 1995, Field 2005, Galiani & Schargrodsky 2010). In the historical literature, Hornbeck (2010) analyzes the impact of the introduction of barbed wire on agricultural productivity in the western U.S., and concludes that the stronger protections of land title afforded by barbed wire led to a significant increase in settlement, land values and crop productivity.

This paper adds to this existing literature while making a number of new contributions. First, I evaluate the impact of insecure tenure on an unusually large set of economic outcomes. Second, I demonstrate a systematic correlation between the frequency of local disruptions to property rights and variation in the costs of those disruptions. Third, I am able to analyze the objective function corresponding to the village leader's choice to reallocate. Given that China has been the site of some of the most far-reaching experiments in property rights over the last fifty years, evidence about both the political economy and the economic consequences of insecure property rights in rural China can potentially be a useful contribution to the ongoing debate about how to structure land rights to maximize rural growth (Deininger & Binswanger 1999)

The remainder of this paper proceeds as follows. Section 2 provides an overview of the relevant institutional background. Sections 3 and 4 provide a conceptual framework for the analysis and describe the data. Section 5 presents the results focusing on intravillage heterogeneity in security of tenure, while Section 6 analyzes cross-sectional variation in the costs of reallocation derived from an estimation of the agricultural production function. Section 7 discusses estimation of the village leader's objective function, and Section 8 concludes.

2 Background

Property rights in China have a long and tumultuous history in the post-1949 era, and the institutional framework that governs rural households remains unusually complex. This section provides a broad overview of the history of property rights during the Communist period, as well as the characteristics of the periodic reallocations that have been a feature of the rural land ownership system since 1983.

2.1 Property rights under the Household Responsibility System

Since 1978, land rights in China have been characterized by a system of collective land tenure in which partial use rights are assigned to the household, a system widely known as the household responsibility system. Prior to this, agricultural production in China between 1962 and 1978 was largely organized around the institution of the production team, a unit consisting of 20 to 30 households that jointly farmed agricultural land and sold the resulting output, distributing the associated income to participating laborers according to a system of workpoints intended to reward labor, skill and political commitment.[1] The overarching imperative of agricultural policy during this Maoist period was to maximize grain production in order to feed urban areas and support industrialization drives, a goal enforced using substantial mandatory production quotas and low procurement prices. By 1978, the cumulative impact of these policies was disastrous, leading to low rural income, land degradation and a severe undersupply of non-grain crops (Walker 1984).

As a result, the new government led by Deng Xiaoping, acceding to power shortly after Mao Zedong's death, introduced major changes in agricultural policy. First, the household was reinstated as the primary unit of agricultural production under a system variously known as household contracting or the household responsibility system. Each household was provided with an allocation of land for its own use, while land title continued to be held by the village. The household also committed to the delivery of a fixed amount of quota grain sold to the state at a preset price, in addition to taxes paid.

Excess production could then either be sold to the state at a higher, above-quota price, or at rural markets (Lin 1992), with the household having full rights over residual, post-quota income. Households were also allowed control over a private plot of land used to cultivate crops other than grain or to raise animals; income from this plot accrued entirely to the household (Walker 1984). The average per capita land endowment was small, less than one fifteenth of a hectare, and a household's endowment generally comprised several fragmented parcels (Wen 1989).

At the same time, major adjustments were made to the state's system of agricultural targets and agricultural procurement. Prices for government procurement of most agricultural goods, previously so low that they often did not cover costs, were raised substantially. In addition, a previously elaborate system of targets for sown area, inputs, production and yield for a variety of agricultural productions was simplified to government procurement targets for key agricultural goods only (Lin 1992).

These changes were implemented in a piecemeal fashion between 1979 and 1983, be-

[1] The team farming system was itself a retreat from the much larger agricultural communes formed during the Great Leap Forward between 1958 and 1962, in which land and labor were collectivized in communes of 6,000-8,000 households (Chinn 1979).

ginning with a few isolated provincial or local experiments, and subsequently spreading widely to a point of almost total decollectivization by the end of 1983 (Unger 1985). The establishment of the household responsibility system led to a substantial increase in the growth rate of agricultural output, which had been only 2% annually over the previous 25 years. Between 1978 and 1984, agricultural output increased nearly 8% annually. One analysis estimated that roughly half of this growth was due to increased use of inputs, particularly fertilizer, and half to the assignment of land use rights to households (Lin 1992).

2.2 Land reallocations

However, property rights under the household responsibility system remained crucially incomplete, principally because land was subject to periodic land reallocations. The stated aim of these reallocations was to promote equity in land ownership, and to adjust landholdings in response to changes in household size.[2] However, the policy clearly created an opportunity ripe for rent-seeking by local officials (either local government officials or Party leaders, known as cadres).

Accordingly, the literature has observed that "it is not uncommon that a few village cadres or officials choose to conduct readjustments simply in order to exert their influence and authority for other dubious purposes" (Keliang, Prosterman, Jianping, Ping, Reidinger & Yiwen 2007). Another analysis noted that the threat of reallocation was frequently used as a carrot and stick to ensure compliance with other administrative goals (family planning targets, grain quotas, corvee labor obligations, and taxes) relevant to local leaders' continuity in their positions and opportunities for promotion. Leaders employed the threat of land reallocation to induce households to comply with other policy goals and minimize their enforcement costs, or to punish households for an absence of compliance (Rozelle & Li 1998).

At the same time, reallocations required considerable investment of time on the part of village leaders, entailing "countless discussions and negotiations among village cadres and the involved households pertaining to the new land assignment exercise" (Kung 2000, Brandt et al. 2002). To cite a specific example, a survey in July-August of 1999 found that a third of villages that had decided to carry out a reallocation, in accordance with a land law passed the previous August, had still not implemented it (Schwarzwalder, Prosterman, Jianping, Riedinger & Ping 2002). Though reallocations

[2]Given that variation in the number of children is limited by virtue of the One Child Policy, the relevant changes in household size are normally driven by marriage of adult children. Daughters typically exit the household, while daughters-in-law will arrive. Other changes might be driven by migration, death, or changes in extended family structure.

normally occurred at the end of the year during the fallow winter period, the lapse in time required for implementation introduced scope for strategic behavior, for example hastening the marriage of sons (or delaying the marriage of daughters) in order to maximize the number of family members in the household when its required allotment of land was determined (Unger 2005).

A larger literature in political science has assembled descriptive evidence about the frequency and nature of land reallocations over time. Brandt et al. (2002) find that there is a negative correlation between the frequency of reallocations and the number of plots per household, as well as the total number of households in the village. Kung (2000) uses a separate survey of land reallocations and notes that reallocations decline in frequency when terrain is more rugged or hilly, and when landholdings are more scattered or fragmented. Unger (2005) also documents the negative relationship between topography and reallocations, arguing that more topographically variable locations face higher transaction costs both in drawing boundaries and in designating parcels presumed to be of comparable value that can be used as the basis of swaps. The paper also finds a negative relationship between the frequency of reallocations and the availability of off-farm income-earning opportunities.

The central government has made periodic and largely unsuccessful attempts to regulate or rein in reallocations. By the 1990s, national policymakers became increasingly concerned that insecure tenure was the primary reason for a decline in agricultural growth rates relative to the early years of the Household Responsibility System. As a result, a (nonbinding) policy directive was issued in 1993 establishing a fixed term of land tenure equal to thirty years. This policy was then embodied in formal law in 1998, requiring that land be contracted to households for 30 years. Readjustments during this period were still allowed, but needed to be approved by two thirds of village members; villages were also allowed to conduct a reallocation immediately after the introduction of the new policy.[3] The law also mandated the issuance of written contracts or certificates to farmers.

Despite the seeming boldness of this reform, subsequent survey evidence indicated that its implementation was mixed at best. A majority of farmers continued to express low confidence in their tenure security and believed subsequent readjustments were inevitable (Schwarzwalder et al. 2002). A later law in 2002 outlawed reallocations completely except in extreme cases and spelled out the right to lease, exchange and carry out other land transactions, excluding sale and mortgage. This reform is, however, beyond the chronological scope of this analysis (Keliang et al. 2007), which will focus on the impact

[3] A survey in 1999 reported in Schwarzwalder et al. (2002) inquired about whether villages had decided to conduct such a reallocation and whether it had taken place, the source of the previously cited data about the delay inherent in the implementation of reallocations.

of reallocations on rural economic outcomes between 1987 and 2002.[4]

3 Conceptual framework

3.1 Optimizing reallocation

Consider the decision made by a village leader of whether or not to undertake a reallocation in a given village in a given year. A reallocation has both costs and benefits. The advantages may include private benefits for the official in rent extraction or future opportunities for promotion, as well as quasi-public benefits such as an increase in equity that may also be valued by village households.

On the other hand, reallocations also have costs. Households that are uncertain about their long-term tenure on a given plot will not make investments whose returns accrue partly in the medium-term, thus resulting in a decline in agricultural investment and output; a simple model of a household production function demonstrating this result is presented in Appendix A. These costs are clearly highly salient to households. For simplicity, I will assume here that officials are themselves indifferent to this loss in output. They are, however, forced to take into account the preferences of households by bargaining over whether or not to hold a reallocation.

Assume that the official and each household face a variant of the single-seller, single-buyer problem; they need to bargain over the sale of a single good, a reallocation of land. The official places a value B on this reallocation, capturing benefits that include opportunities for rent-seeking and decreased intravillage inequality.

Each household i in the village places a value on a reallocation that can be written as follows, equal to the negative of the value of continued land tenure $\bar{v}$ plus the value of the expected change in land $w(E[\Delta L])$. For simplicity, I assume that every household in the village would have its land tenure disrupted by the reallocation.

$$v_i = -\bar{v}_i + w(E[\Delta L_i]) \tag{1}$$

$\bar{v}_i$ is defined more specifically as the loss in output due to foregone investments that are not made when tenure insecurity is introduced by a reallocation. Note that X_i^{NR} denotes a vector of agricultural investments made by household i in the absence of a reallocation,

[4]Data from the household survey employed here is not publicly available for the years after 2002; accordingly, it is not possible to employ data from the post-2002 period in a placebo test. There is, in addition, an ongoing debate about how well these subsequent reforms were implemented and thus how secure property rights in the post-2002 period are.

while X_i^R denotes investment in the case of a reallocation.

$$\bar{v}_i \equiv F(X_i^{NR}) - F(X_i^R) \tag{2}$$

Some households may place a negative value on reallocation if they face significant losses due to reduced long-term investment, and thus they will seek to avoid a reallocation. Others may place a positive value on reallocation if they expect to gain land in the process. Each household has the option to impose a bargaining or lobbying cost c_i on the official in the case of the outcome they do not prefer: i.e., a household that prefers a reallocation be avoided can inflict a lobbying cost at the time of the reallocation, and vice versa for a household that prefers a reallocation.[5]

Total bargaining costs are summed across all households in the case of a reallocation, defined $C(R = 1) = \sum_i c_i(R = 1)$, or a non-reallocation, defined $C(R = 0) = \sum_i c_i(R = 0)$. There is also a transactional cost of time and effort T needed to redefine land boundaries. This transactional cost is assumed to be higher for localities with more rugged topography; this assumption is consistent with the prior literature, as well as the intuition that implementing a land swap perceived to be fair is more challenging in areas with variable topography and thus more local heterogeneity in land quality.

The village official will reallocate if the benefits exceed the sum of bargaining and transaction costs:

$$B > C(R = 1) - C(R = 0) + T \tag{3}$$

Accordingly, the variable R_{vt}, defined as equal to one if a reallocation occurs in village v in year t and zero otherwise, can be viewed as a function of benefits of the reallocation for the official, its costs in lost output, and the topographic characteristics of the village.

$$R_{vt} = f(C_{vt}, B_{vt}, T_v) \tag{4}$$

This conceptual framework suggests that villages where households place a larger value on continued land tenure, i.e. $\bar{v}_i^h = F(X_i^{NR}) - F(X_i^R)$ is greater, should also exhibit a lower frequency of reallocations. In these villages, the cost in terms of foregone output of tenure insecurity is greater, and accordingly households will bargain more aggressively against reallocations.[6]

[5]This framework assumes that households can commit to imposing a certain cost on village officials. While this is clearly a strong assumption, it could be easily nested in a multi-period model in which households that fail to impose the postulated bargaining penalty on the official suffer a loss of credibility in future bargaining rounds.

[6]In addition, none of the preceding analysis precludes the possibility that the official himself also faces a direct loss from foregone output, via lowered tax revenue or other channels. In this case, the direct benefit B of reallocation will be lower in villages where the lost output as a result of tenure insecurity is larger; this serves only to strengthen the postulated negative correlation between the output costs of

Intuitively, in some villages there may be few profitable long-term investments available. A reallocation will decrease the probability that households in the village will make such investments, but this may not be a significant loss if the return on these investments is low. More specifically, the difference in investment and thus in output between the reallocation and the non-reallocation case is increasing in the returns to lagged agricultural inputs, a comparative static also demonstrated in Appendix A. If the returns to lagged inputs are higher, the loss in investment as a result of a reallocation is higher, the net benefit of reallocation for the official is lower, and accordingly reallocations should be observed less frequently.

3.2 Optimizing household-level land shifts

To sum up, the observed distribution of reallocations across villages and years can be understood as the outcome of a complex bargaining process that leads to some officials choosing to conduct reallocations in certain years while others do not. However, village leaders who have chosen to hold a reallocation then face another set of optimization decisions: how and to whom to reallocate land within the village. Some households will gain or lose land, while other households may not see changes to their landholdings.

The probability that a given household i in village v and year t will see its land reallocated is denoted D_{ivt}; it is assumed to be a function of household characteristics X_{ivt}, conditional on $R_{vt} = 1$. If there is no reallocation, then $D_{ivt} = 0$ for all households.

$$D_{ivt} = f(X_{ivt}) \mid R_{vt} = 1 \quad (5)$$

$$D_{ivt} = 0 \mid R_{vt} = 0 \quad (6)$$

Potential household covariates X_{ivt} relevant to the reallocation decision could include demographic characteristics that render the household a poor match with its current land allotment; the household's current position in the overall distribution of landownership, given the village leader's interest in equity; and the past history of land shifts for the household.

Accordingly, there are two sources of variation in insecure tenure for the households of interest, corresponding to two separate optimization margins for the village official. There is variation in the household probability of land shifts D_{ivt} conditional on a reallocation occurring ($R_{vt} = 1$), corresponding to the official's choice of which households to reallocate. There is also variation in the probability of reallocation across villages and

reallocation and their frequency.

years, corresponding the official's choice of whether or not to hold a reallocation. In this analysis, I will exploit both sources of variation in tenure insecurity.

4 Data

The dataset employed here is a panel collected by the China Research Center for the Rural Economy (RCRE), comprising a sample of 299 villages in 13 provinces in China every year between 1986 and 2002, excluding 1992 and 1994. Figure 1 shows the sample counties. A randomly selected sample of households in each surveyed village forms the panel; the mean number of households in a village-year cell is 69. Summary statistics for the dataset are shown in Table 1.

Measures of land reallocation are constructed using household reports of changes in their household landholdings from year to year, excluding land leased.[7] A shift in landholdings is identified at the household level if a household reports a change in land area owned of at least .1 mu, where a mu is the Chinese unit of land area (comprising .165 acres).[8] A reallocation is defined to have occurred when the proportion of households reporting a change in their landholdings in a given village in a given year exceeds the 75th percentile across all village-years or the proportion of land reported transferred exceeds the 75th percentile across all village-years. This definition is employed to exclude those cases where a small number of households report a change in landholdings as a result of measurement error or a private contractual arrangement that is not sanctioned by the village leadership.

Figure 2 shows histograms for both measures used to define the measure of reallocation. Both show a spike close to zero and a long right tail with a higher proportion of transfers. The reallocation measure employed captures this right tail. In addition, the first stage and the main results are robust to altering this definition; results employing varying definitions of reallocations will be shown in the robustness checks.[9]

Past literature on reallocations that has estimated their frequency has largely been drawn from two sources: surveys of village leaders, e.g. Kung (2000), or surveys of individual households conducted periodically by the Rural Development Institute that obtain retrospective statistics over a long recall period (Schwarzwalder et al. 2002). Survey data

[7] There is no uniform policy regarding the legality of land leasing arrangements in rural China. In this sample, leasing is rare; only around 8% of household-year observations report any land leased in or out. Leased land is thus of limited relevance to the rural economy overall.

[8] .1 mu represents around 2% of median land owned.

[9] This definition makes no distinction between different types of plots that households may hold (e.g., responsibility land versus private plots); though the dataset reports limited information on holdings of responsibility land, inputs and agricultural activity are not reported by type of plot. Accordingly, the resulting estimates should be viewed as mean effects across all household landholdings.

of leaders has the advantage of employing a clear definition of reallocation. However, leaders may also face incentives to bias reports of reallocations toward zero to avoid reporting reallocations that are not in line with national land policy guidelines. Retrospective data collected at the household level, on the other hand, may be imprecise and biased by recent events.

While survey data of leaders indicate that reallocations occur around every 5 years (Kung 2000), the reallocation measure constructed here R_{vt} shows reallocations occurring around every three years. It is plausible that a measure based on household reports of land shifts will be noisier and thus more likely to generate spurious reports of land reallocations; this may be a source of classical measurement error. However, this strategy for identifying household reallocations has the additional advantage of allowing the direct examination of the changes in landholdings at the household level that were induced as a result of the reallocation. A measure constructed from reports by village leaders, by contrast, provides no information about the mechanics of the implementation of the reallocation within the village.

5 Intravillage heterogeneity in security of tenure

5.1 First stage

In analyzing reallocations, I will begin by considering variation in tenure insecurity within a village, taking as given the observed distribution of reallocations across different villages and years. When a reallocation does occur in the sample villages, ex ante all households face the risk of the suspension of their use rights and the transfer of their plot. However, not all households experience a change in landholdings in every reallocation.

In order to evaluate the effect of variation in security of tenure on economic outcomes, it is useful to begin by analyzing the characteristics of households that do have their land reallocated. Assuming that the quantity of land already held is of first-order relevance, I first estimate the probability of a household's land being reallocated conditional on a reallocation occurring in the village for households in each decile of landownership. These probabilities are shown graphically in Figure 3.

The evidence indicates that land transfers are broadly progressive. The probability of receiving a positive transfer of land via a reallocation is generally decreasing by decile, and the probability of a negative transfer is increasing. Only the tenth (and richest) decile appears to be somewhat insulated from the effects of reallocations. Otherwise, households from the lower deciles are generally more likely to gain land, and households from the upper deciles more likely to lose it.

Now, assume one reallocation has already occurred in every village in the past. Both reallocation "winners" and reallocation "losers" have experienced a shock to their landholdings and, presumably, to other economic outcomes as well. Two groups of households can be defined based on whether their land was affected in the last reallocation: $DP_{ivt}^{-1} = 1$ defined for household i in village v in year t denotes a household that gained land in the previous reallocation three to five years before, and $DN_{ivt}^{-1} = 1$ denotes a household that lost land. These households have received opposite shocks, relative to the unaffected households, with the median (absolute) change in landholdings observed as a result of a reallocation around one third of median land owned.

There is, however, one characteristic common to all households that had their land reallocated in the previous round: a decline in the probability that their land tenure will be disrupted again in the next reallocation. Reallocating land incurs high transaction costs. Accordingly, it is logical to assume that village leaders will seek to minimize the number of land transfers they effect over time, conditional on reaching their goals of equity or an improved match between households and land, and this in turn implies that a series of incremental land transfers is unlikely to be welfare-maximizing for the official. Instead, he would seek to fully adjust a household's land to its optimal level when a reallocation is implemented, implying a lengthy period until either subsequent demographic shocks render the household's landholding suboptimal, or the household is again due for an equity-enhancing shift in land.

To test this hypothesis, I estimate the impact of past reallocation inclusion on a dummy variable capturing inclusion in the current reallocation, denoted D_{ivt} for household i in village v in year t. D_{ivt} is defined to be equal to one if a household reports any change in total land owned above the threshold (.1 mu) in the year of the reallocation. R_{vt} is defined as equal to one if a reallocation is observed in village v in year t.

The independent variables of interest are dummy variables for a household's past reallocation inclusion interacted with reallocation at the village-year level, R_{vt}. Having had land reallocated in the past may be correlated with the probability of a disruption to current land tenure, but only if a reallocation is actually occurring in the village; otherwise, the impact of past reallocation on current reallocation is precisely zero. A control for each strata of landownership L_{ivt} and village and year fixed effects are included, and the model is fully saturated in interactions with R_{vt}.[10]

$$\begin{aligned} D_{ivt} = \beta_1 DP_{ivt}^{-1} \times R_{vt} + \beta_2 DN_{ivt}^{-1} \times R_{vt} + \beta_3 L_{ivt} + \beta_4 L_{ivt} \times R_{vt} \\ + \nu_v + \gamma_t + \nu_v \times R_{vt} + \gamma_t \times R_{vt} + \epsilon_{ivt} \end{aligned} \tag{7}$$

[10] L_{ivt} is an integer variable controlling for each vingtile (5%) of landownership, i.e. ranging from 1 to 20. This controls for the position of the household in the overall distribution of landownership.

Here, the interactions between R_{vt} and γ_v and γ_t capture the main effect of R_{vt}. Rather than constraining the primary effect of a reallocation to be linear, the specification allows for a non-linear effect of reallocations in each village and year.

This equation is estimated for the household panel post-1995, to allow for coding of D^{-1}_{ivt} based on prior reallocations reported in the first section of the panel. 1995 is chosen as the cut-off year as a new and more comprehensive survey of household economic outcomes was administered for the first time in that year. The coefficients on β_1 and β_2 from estimating (7) are shown in Column (1) of Table 2, and are negative and significant. In other words, inclusion in a reallocation leads to a significant decrease in the probability that a given household will have its land adjusted again in the next reallocation for both past reallocation gainers and past reallocation losers, compared to households that were not reallocated. This is consistent with the intuition that multiple, sequential reallocations of land for the same household are unlikely to be optimal.

These results suggest that past reallocation history generates a plausible source of quasi-exogenous variation in current tenure insecurity. The implied exclusion restriction is that a reallocation has no differential impact on households that were included in a past reallocation and households that were not included, other than via the channel of differential probability of current reallocation and thus differential tenure insecurity. Column (2) shows the result of estimating the same equation with past reallocation participation pooled across gainer and loser households. The pooled dummy for past inclusion in a reallocation is denoted D^{-1}_{ivt}.

$$D_{ivt} = \beta_1 D^{-1}_{ivt} \times R_{vt} + \beta_2 D^{-1}_{ivt} + \beta_3 L_{ivt} + \beta_4 L_{ivt} \times R_{vt} + \nu_v + \gamma_t + \nu_v \times R_{vt} + \gamma_t \times R_{vt} + \epsilon_{ivt} \quad (8)$$

This is the first stage relationship of interest, and the same negative and significant relationship is evident.

Moreover, the heterogeneity of past reallocation patterns (including both winners and losers) can be used as an additional test of the exclusion restriction. Any bias in unobservables as a result of past reallocation-induced shocks to land is presumed to be of opposite sign for past winners, who now own more land than the mean household, and losers, who now own less land. Figure 4 shows estimated kernel densities of landownership for households with past positive and negative shocks to landholdings in reallocations, partialling out village and year fixed effects. Both a shift right in the distribution for past land gainers and a shift left for past land losers are evident.

Accordingly, if the reduced form impact of past reallocation status on economic outcomes in a reallocation year is observed for both past land-gainers and past land-losers, this suggests that the observed effect is plausibly interpreted as a causal estimate of the impact of tenure security on economic outcomes. A violation of the exclusion restriction would require that reallocation is correlated with a shock that affects both the relatively land-poor and the relatively land-rich, a non-monotonic pattern that would seem a priori implausible. Further evidence about the validity of the exclusion restriction will be presented in the next section.

5.2 Reduced form and 2SLS

The reduced form specification is the following, where Y_{ivt} denotes economic outcomes at the household level. Controls for lagged household reallocation D^{-1}_{ivt}, strata of landownership L_{ivt} and the full set of interactions with reallocation R_{vt} are included.

$$Y_{ivt} = \beta_1 D^{-1}_{ivt} \times R_{vt} + \beta_2 D^{-1}_{ivt} + \beta_3 L_{ivt} + \beta_4 L_{ivt} \times R_{vt} + \nu_v + \gamma_t + \nu_v \times R_{vt} + \gamma_t \times R_{vt} + \epsilon_{ivt} \quad (9)$$

The reduced form can also be estimated as the "split" reduced form, including both DP^{-1}_{ivt} and DN^{-1}_{ivt} and the corresponding interactions as explanatory variables. The 2SLS specification is the following, where D_{ivt} is a dummy for forced reallocation of land at the household level, instrumented by $D^{-1}_{ivt} \times R_{vt}$.

$$Y_{ivt} = \beta_1 D_{ivt} + \beta_2 D^{-1}_{ivt} + \beta_3 L_{ivt} + \beta_4 L_{ivt} \times R_{vt} + \nu_v + \gamma_t + \nu_v \times R_{vt} + \gamma_t \times R_{vt} + \epsilon_{ivt} \quad (10)$$

The assumed timing in each year is as follows: a signal about the reallocation is received prior to household's investment decisions. Investments are made and output is realized. Subsequently, land is reallocated. While the exact timing of the reallocation decision vis-a-vis household investment decisions doubtless varies, the assumption is that the considerable time required to implement a reallocation requires a decision to be made at a point that overlaps with the period of key investments, in line with the evidence that households are observed to respond strategically to early notifications about future reallocations in decisions about household formation and marriage. Such strategic behavior would be impossible if the decision to reallocate was simultaneous with the actual implementation.

Eight outcome variables are reported for each specification: land cultivated, fertilizer, agricultural labor, a dummy for agricultural structures, moveable capital, grain production, and dummies for labor in a non-agricultural household business and for labor outside the household. Land cultivated, fertilizer, agricultural labor and grain production are normalized by the area of land owned prior to the reallocation; all variables are then normalized by the mean and standard deviation of the outcome variable in the control (non-participating) households, following Katz, Kling & Liebman (2007).[11]

Reduced form results Panel A of Table 3 shows the results from estimating the reduced form, and Panel B the "split" reduced form with dummies for both past reallocation winners and losers, $DP_{ivt}^{-1} \times R_{vt}$ and $DN_{ivt}^{-1} \times R_{vt}$ respectively. Note again that all dependent variables are normalized to have mean zero and standard deviation one.

In Panel A, the coefficients on the interaction $D_{ivt}^{-1} \times R_{vt}$ are generally positive and significant with magnitude between .05 and .1, reflecting greater agricultural investments by households that were included in the last reallocation and accordingly enjoy greater tenure security. No effect is observed for moveable capital, labor input into household businesses or labor in outside enterprises. This is consistent with the intuition that the returns to moveable capital (an index of animals, tools and machines owned) and non-agricultural activities are unaffected by reallocations.

In the split reduced form, the coefficients are positive and significant for households that gained and lost land in the past. The fact that the estimated coefficients are generally slightly larger for past losers is consistent with the evidence of a larger first stage for these households (i.e., their relative tenure security is greater). However, the final row of Panel B reports a test of equality of the coefficients β_1 and β_2 on $DP_{ivt}^{-1} \times R_{vt}$ and $DN_{ivt}^{-1} \times R_{vt}$, and the hypothesis that the coefficients are equal is uniformly not rejected.

In addition, the assumption that there is no omitted channel that is biasing the estimated effect for both land losers and land winners can be tested by examining the estimated coefficients on DP_{ivt}^{-1} and DN_{ivt}^{-1}. In general, these coefficients are of opposite sign, though not statistically significant; households that gained land seem to employ less inputs per acre and are somewhat less likely to have non-agricultural businesses. Most importantly, the absence of any pattern of symmetric and significant coefficients on the dummies for past reallocation winners and losers suggests there is no common bias in observables across both sets of households. The only exceptions are positive coefficients on the probability of outside labor.

The exclusion restriction for the instrumental variables analysis requires that there is no shock correlated with a village-level reallocation R_{vt} that differentially affects households included in the previous reallocation. If households that had their land previously reallocated either positively or negatively showed characteristics that were significantly different from households with no previous reallocations in non-reallocation years, this

[11] Fertilizer is defined as the mean of total fertilizer and the most common subtype of fertilizer used, carbamide. Moveable capital is defined as the sum of animals and tools; agricultural structures is equal to one if a household reports any agricultural structures or associated machines. The top 5% and bottom 1% of observations of each continuous outcome variable are trimmed to remove the influence of outliers. The asymmetry reflects the much longer right tail in the distribution of agricultural input variables.

would suggest that the past history of reallocations generated different trends for both land winners and land losers. Furthermore, if there were an interaction between these trends and R_{vt}, the coefficients of interest would be biased. Given the absence of any evidence of significant difference in outcomes for past reallocation participants in non-reallocation years, however, it seems implausible that there is another, independent shock correlated with R_{vt} that affects only these households in reallocation years. Further evidence about differing trends for past land gainers and land losers in the years prior to a reallocation will be presented in the robustness checks.

2SLS results To reiterate the key assumptions underlying the two-stage least squares result, the exclusion restriction requires that a reallocation has no differential effect on households that were and were not included in the previous reallocation, other than via the channel of differential tenure security. Given the asymmetric nature of past reallocation inclusion, encompassing both reallocation winners and reallocation losers, the necessary assumption can be further refined: there is no shock correlated with a reallocation that affects both relatively land-poor and relatively land-rich households compared to the mean.

Under this assumption, Table 4 shows the results from estimating equation (10), the instrumental variables specification. The coefficients indicate that households that face a disruption to their continued land tenure with probability one exhibit a decline in area sown, fertilizer, agricultural labor and total agricultural production, all around one standard deviation in magnitude. There is no significant change in moveable capital or agricultural structures, and no change in the probability of non-agricultural employment.

These results are consistent with a model of household behavior in which households decrease the use of inputs that have medium-term returns and inputs that are complementary to those medium-term investments. The shift in sown area may reflect a decline in the prevalence of multicropping.[12] Optimized multicropping yields long-term benefits in terms of soil nutrition and health (Zhang, Shen, Li & Liu 2004), and thus households expecting short tenure may be less likely to multicrop. The decline in both multicropping and fertilizer use generates a decline in agricultural labor, presumably a complementary input.

No effect is observed for portable agricultural capital or agricultural structures, and there is also no effect for non-agricultural activities. Given that both the establishment of a non-agricultural household business and the search for outside employment (often rationed in rural China) may require considerable initial, and potentially irreversible,

[12]The average household in this dataset is multicropping around 50% of land owned, a rate consistent with previous estimates from agricultural censuses and remote sensing data (Frolking, Qiu, Boles, Xiao, Liu, Zhuang, Li & Qin 2004).

investments, it would be implausible to see a substantial divergence in non-agricultural investments between households with different short-term expectations of land tenure. No such divergence is observed. This evidence is consistent with the hypothesis that the observed impacts represent the effect of variation in short-term tenure, rather than other unobserved differences between households with different reallocation histories.

Panel B of the same table shows the results of estimating (10) controlling for a quadratic polynomial in land area held by each household, also interacted with R_{vt}. This specification tests whether differences in plot size between households that did and did not participate in past reallocations are a source of bias, and the estimated coefficients are consistent and in fact more precise. Panels C through E show the two-stage least squares results where the sample is restricted according to certain criteria. These specifications are discussed below in the robustness checks.

5.3 Robustness checks

This section presents a series of robustness checks on the above results.

Differing trends for households with different past reallocation histories It can also be hypothesized that households included in the last reallocation, who enjoy relatively greater tenure security, begin to show higher investments in years prior to the subsequent reallocation. This pattern could emerge for two reasons: first, while the previous specification assumed that households have perfect information about the timing of a reallocation, in fact this information may be noisy. Households may perceive some latent risk of a reallocation occurring in the year or two prior to its actual date. Second, even if they perfectly anticipate the next reallocation date, they may begin to taper down investments that have a time horizon longer than the anticipated time lapse to the next reallocation.

The objective of this robustness check is to evaluate whether the difference in agricultural investments between households that previously participated in a reallocation (who have relatively greater tenure security) and households that did not (who have less tenure security) is evident in years prior to the implementation of a reallocation. In order to test this hypothesis, the reduced form equation (9) is re-estimated for the reallocation year (denoted T=0) and each year leading up to a reallocation. For simplicity, the variable R_{vt} and associated leads R_{vt}^{+1} (one year prior to the reallocation), R_{vt}^{+2} (two years prior to the reallocation), etc. enter the equation linearly rather than interacted with village and year fixed effects.[13] Thus the equation of interest can be written as follows, for example

[13]The reallocation lead variables are coded as follows: moving backwards from the final observed reallocation, each previous year is coded as T=+1, T=+2, etc.; when a year with another reallocation is

for the first lead R_{vt}^{+1}:

$$E_{ivt} = \beta D_{ivt} + D_{ivt}^{-1} + L_{ivt} + R_{vt}^{+1} + \nu_v + \gamma_t + \epsilon_{ivt} \tag{11}$$

The estimated parameters, capturing the difference in outcomes between households that were included in past reallocations and those that were not in each specified year leading up to a reallocation, are then graphed in Figure 5 along with a 90% confidence interval.

The graphs show that for outcomes that are affected by reallocations (fertilizer, sown area, labor, structures and agricultural production), there is generally a pattern of increasing divergence between households previously included in reallocations and households not previously included in the two years prior to the next reallocation. Though the estimated coefficients are not statistically significant, they are positive and increasing in magnitude for variables R_{vt}^{+1} and R_{vt}^{+2}; there is little evidence of a significant trend in longer lags. However, for those outcomes that are hypothesized to be unaffected by tenure security, no significant trend is observed.

These results provide suggestive evidence that households at higher risk of losing their plots may begin tapering their investments in the years prior to a reallocation, though the largest effect is seen in the reallocation year. The absence of any systematic trend in longer lags or for other variables, however, suggests that there are no unobservable characteristics of households previously included in reallocations that are driving the results. In addition, the evidence of a divergence in agricultural inputs between households with different probabilities of future loss of their plot prior to the year of the reallocation suggests that the prior estimates of the impact of insecure tenure on investment may be conservative.

Miscoding partial versus full reallocations A second potential challenge to this estimation strategy is the possibility that some of the reallocations identified are what the literature has identified as partial reallocations, distinguished by the fact that only households that have had changes in their household composition experience incremental shifts in landholdings, without full swaps of their plots. In this case, households that need more land might receive an incremental, additional transfer, while households that have too much land would lose part of their holdings (Keliang et al. 2007).

In order to address this possibility, there are two separate cases that should be considered. One is that the past reallocation, on the basis of which DP_{ivt}^{-1} and DN_{ivt}^{-1} were defined, is in fact not a full reallocation. The second is that the current event that is generating insecurity in tenure, captured by dummy $R_{vt} = 1$, is not a full reallocation.

Under the first case, some households for which $D_{ivt}^{-1} = 1$ may have previously had

encountered, all the lead variables are re-set to zero. The regression sample size thus shrinks with each additional lead year; for each newly defined sample, the bottom 10% of outliers are trimmed.

their land reallocated partly or primarily because of their changes in composition. If this were the case, then the exclusion restriction required for the instrumental variables specification requires that a reallocation has no differential effect on households that previously experienced a change in composition compared to those that did not, other than via the channel of differing tenure security. A violation of this exclusion restriction would arise if there is a shock correlated with reallocation that differentially affects households with a past history of demographic shifts.

On the other hand, if what we identify as a current reallocation is in fact a partial reallocation or some other type of irregularity in land transfer, and households' expectations are rational, then only some households are subject to decreased tenure insecurity: more specifically, those households that expect to lose land based on a shift in their household's composition. The exclusion restriction implied by this specification requires that there is no shock correlated with land reallocations that differentially affects relatively land-rich (on a per capita basis) households.

In both cases, the exclusion restriction is weaker than that postulated for the primary analysis, primarily because the specification can no longer be interpreted as a symmetric and non-monotonic effect of greater tenure security observed for both relatively land-poor and relatively land-rich households. Accordingly, if miscoding is common and reallocation is correlated with other shocks that affect relatively land-rich households or households with previously unstable composition, this could generate bias. In order to test the robustness of the results to potential bias introduced by the miscoding of partial reallocations, I restrict the sample in several ways.

First, I restrict the sample to households that did not previously report a change in composition in the year of the previous reallocation. These are households that could not have been selected for inclusion in a past partial reallocation by virtue of their demographic instability. If the primary results represent bias introduced by correlated shocks for demographically unstable households, this specification should show no significant effect. The results are shown in Panel C, and the estimated coefficients are significant in both sign and magnitude.

Second, I evaluate the effect of households that can reasonably be assumed not to be relatively land-rich on a per capita basis. If the miscoding of partial reallocations as R_{vt} is common and the estimated effect reflects a correlated shock for these relatively land-rich households, these specifications should show no effect. Panels D and E show the results of re-estimating (10) restricting the sample first to households that have gained or remained constant in composition (Panel D), and second to households in the bottom half of the land distribution (Panel E). These are households that on the basis of demography and land ownership are plausibly land-gainers, not land-losers. The results again remain consistent in both sign and magnitude, suggesting that there is no bias introduced by miscoding and the presence of correlated shocks for only land-rich households.

Information as a channel for predicting reallocations The identification strategy also requires the assumption that there are no unobservable characteristics shared by households that experienced both negative and positive shocks to their landholdings in past reallocations that could co-vary with future reallocations. One potentially plausible assumption would be that village officials, who execute reallocations, systematically have more information about households that have similar characteristics to themselves and thus are more likely to participate in their social networks. Given the greater informational salience of these "socially proximate" households, officials may be more likely to alter their landholdings to a level the official regards as optimal. This unobservable proximity to village officials could also generate other time-varying effects if, for example, village leaders prefer to simultaneously implement a reallocation with another policy shift that also differentially affects households with close ties to the village leadership.

This hypothesis can be tested by examining whether there are any characteristics that, when shared with village officials, render proximate households symmetrically more likely to experience positive and negative reallocation shocks to their landholdings. A series of dummy variables are defined that capture households' economic specializations (whether they cultivate rice or wheat, and whether they report household businesses of any of the enumerated types), and a limited number of social characteristics enumerated in the survey (the presence within the household of an individual with education beyond high school, a veteran, resident grandparents or a Communist party member).

For each village-year cell, the mean of this dummy is calculated for households that are reported to be led by a village official, and this official mean is denoted O_{vt}. The equations of interest regress the dummies for positive and negative reallocation on the household indicator of interest I_{ivt}, the official indicator O_{vt} and the interaction $I_{ivt}O_{vt}$, all interacted with R_{vt}. The equation also includes a control for each strata of landownership L_{ivt} and village and year fixed effects interacted with R_{vt}. The specification is thus parallel to the original first stage, where the primary independent variable of interest is the interaction between official and individual characteristics. The objective is to test whether households with a particular economic or social characteristic are more likely to have their land reallocated in villages where officials share this characteristic, in years in which a reallocation is implemented.

$$\begin{aligned} D_{ivt}^{P/N} = & \beta_1 I_{ivt} O_{vt} \times R_{vt} + \beta_2 I_{ivt} \times R_{vt} + \beta_3 O_{vt} \times R_{vt} + \beta_4 L_{iv,t-1} + \beta_5 L_{iv,t-1} \times R_{vt} \\ & + \nu_v + \gamma_t + \nu_v \times R_{vt} + \gamma_t \times R_{vt} + \epsilon_{ivt} \end{aligned} \tag{12}$$

The results shown in Table 5 indicate no clear pattern of coefficients across the various interaction terms. The only household characteristics that seems to generate substantial shifts in both reallocation probabilities are rice cultivation, husbandry and Communist party membership, but the effects are not symmetric: when officials also cultivate rice, households that cultivate rice are more likely to gain land in a reallocation and less likely to lose it (i.e., they are favored), and likewise for husbandry. Thus it is plausible to conclude that there is very little evidence that informational proximity to village officials serves as a common source of bias for both households that gain land and households that lose land in a reallocation.

6 Cross-sectional variation in reallocation costs

If the exclusion restriction for the 2SLS specification estimated above is valid, then the resulting estimates represent the causal effect of a change in the probability of land reallocation on investment and economic outcomes within the same village and year, conditional on the observed distribution of official reallocation decisions. The estimated cost is uniform for all villages. However, the heterogeneity in the observed probability of reallocation suggests that the benefits and costs of reallocation are far from constant. Moreover, the bargaining process that generates the observed distribution of reallocations is itself a function of these benefits and costs. Accordingly, it is reasonable to hypothesize that there should be a negative correlation between the costs of reallocation and its probability.

The measure for relative costs of reallocation employed here is derived from the model of household optimization outlined in Appendix A. Households are assumed to equate the ratio of returns to labor N_t and capital F_t (fertilizer) in agriculture to the ratio of factor prices. In the case of a reallocation, there are no lagged returns to fertilizer and the solution given a Cobb-Douglas production function is standard:

$$F^R = \frac{w\alpha_F}{r\alpha_N}N \tag{13}$$

In the counterfactual case of no reallocation, the returns to fertilizer are realized both this period and next period and the ratio of returns to labor and fertilizer has a more complex form. The optimal level of fertilizer solves the following equation equating the ratio of returns to labor and fertilizer (both this period and next period) to the ratio of factor prices.

$$\frac{w}{r} = \frac{\frac{\partial \pi_t}{\partial N_t}}{\frac{\partial \pi_t}{\partial F_t} + \frac{\partial \pi_{t+1}}{\partial F_t}} \quad (14)$$

$$0 = \frac{\frac{\partial \pi_t}{\partial N_t}}{\frac{\partial \pi_t}{\partial F_t} + \frac{\partial \pi_{t+1}}{\partial F_t}} - \frac{w}{r} \quad (15)$$

The first-order condition indicates that when the returns to lagged capital are higher, fertilizer use increases, a result shown algebraically for the Cobb-Douglas case in the appendix. Accordingly, when the returns to lagged capital are higher, the difference in investment between the reallocation and the non-reallocation case and thus the cost of a reallocation is higher — presumably making reallocations less likely.

The objective of this section is to test the hypothesis that the frequency of reallocations is correlated with their relative cost by estimating a production function that allows the returns to agricultural inputs to vary cross-sectionally. First, I will describe the methodology used to estimate the production function. Second, I will present the primary results that test the correlation between returns to lagged capital and reallocations. Third, I will estimate optimal capital inputs and present evidence about the correlation between foregone investment as a result of a reallocation and reallocation frequency.

6.1 Estimating an agricultural production function

The production function postulated is Cobb-Douglas; inputs are labor, land area, fertilizer and lagged fertilizer. Sown area, seeds, labor and production for grain production are reported separately, and fertilizer employed is assumed to be devoted to grain cultivation proportionately relative to its share in total sown area. Lagged inputs are set equal to the amount of that input used in the previous year, provided that the household did not participate in a reallocation last year (i.e., conditional on the household cultivating the same land this year and last year). The objective is to identify lagged returns of inputs on land cultivated continuously by the same household.

First, I estimate the production function using OLS with village, year and crop fixed effects, employing both the full sample and the sample restricted to rice and wheat producers.[14] This specification follows the methodology employed in other production function analyses of Chinese agriculture (Lin 1992, Wan & Cheng 2001). The dependent variable is value added in grain production, equal to grain production valued at the market price minus the cost of seeds. X_{ijt} is the quantity of input j used by household i

[14]In order to identify these households, a dummy for rice or wheat production is set equal to one if a household reports positive sown area for one crop (rice or wheat) and no sown area for the other crop throughout the entire panel.

at time t and $F_{ij,t-1}$ denotes lagged fertilizer; lower-case letters denote log inputs.[15] μ_c, ν_v and γ_t denote crop, village and year fixed effects.

$$y_{it} = \sum_{j=1}^{J}(\alpha_j x_{ijt} + \alpha_f f_{ij,t-1}) + \mu_c + \nu_v + \gamma_t + \epsilon_{ivt} \tag{16}$$

Clearly, ordinary least squares estimates of the returns to agricultural inputs may be biased by the presence of unobserved shocks to productivity. For robustness, the production function is re-estimated using dynamic panel methods as described in Blundell & Bond (2000). A detailed description of the methodology and the results can be found in Appendix B, but in essence, the postulated production function imposes an AR(1) structure on the errors, yielding orthogonality of lagged levels of the independent variables and differences in those variables. Additional restrictions on the correlation between the household fixed effect η_i and differences ΔX_{it} and ΔY_{it} allow for the imposition of additional moment conditions that employ lagged differences as an instrument for the level equation. Given that the use of lagged levels as instruments requires dropping observations without observed lags, I estimate the production function only with the full sample of grain producers in order to maintain adequate power. While the primary specification employs the full set of lagged instruments, I also restrict the instruments employed to lags three and four to evaluate the robustness of the results to a change in the instrument set.

The results from the estimation of the production function using both methods can be found in Table 6. Each specification is reported with and without the returns to lagged fertilizer. The pattern of coefficients is relatively consistent across columns, with several caveats. The estimated returns to labor are more variable in the dynamic panel regressions and not statistically significant, and the estimated returns to lagged fertilizer are also noisy, though of substantial magnitude. The point estimates for returns to both fertilizer and lagged fertilizer are larger in the dynamic panel specifications, though the difference is not significant. Comparing across all the specifications, the returns to lagged fertilizer are between 10% and 40% of the returns to contemporaneous fertilizer, consistent with the intuition that a non-trivial component of the returns to fertilizer are realized in the medium-term.

The final rows of Table 6 report for the dynamic panel specification the results of the Sargan-Hansen test of overidentifying restrictions and the test of common factor restrictions imposed in the minimum distance model used to estimate the coefficients. (Again,

[15] The standard production function literature has generally employed either value added (Olley & Pakes 1996) or revenue (Levinsohn & Petrin 2003) as the dependent variable. The primary specification here employs value added, but in the cross-sectional analysis I will also present results using revenue (the grain harvest valued at the market price) as the dependent variable.

details can be found in Appendix B.) While the overidentifying restriction is rejected for specifications excluding lagged returns to fertilizer, for the primary specification the validity of lagged levels as instruments is not rejected at the 10% level (Column 3) or at the 5% level (Column 4). The test of common factor restrictions is uniformly not rejected.

6.2 Correlation between returns to capital and reallocations

In order to test the hypothesis that there is variation in the returns to lagged capital that is correlated with variation in reallocation behavior, I now estimate the production function allowing the returns to inputs to vary by province and crop. Interaction effects between crop fixed effects μ_c and province fixed effects λ_p and all agricultural inputs (x_{ijt} and $f_{ij,t-1}$) are included. The equation is again estimated using both OLS with crop, village and year fixed effects and dynamic panel GMM.

$$\begin{aligned} y_{it} = &\sum_{j=1}^{J}(\alpha_j x_{ijt} + \alpha_f f_{ij,t-1}) + (\sum_{c=1}^{C}\mu_c + \sum_{p=1}^{P}\lambda_p) \times (\sum_{j=1}^{J}\alpha_j x_{ijt} + \alpha_f f_{ij,t-1}) \\ &+\mu_c + \nu_v + \gamma_t + \epsilon_{ivt} \end{aligned} \tag{17}$$

I then calculate the return to lagged fertilizer for each household, corresponding to the linear combination of the returns in the province and for the crop cultivated. This allows for the calculation of the mean return to lagged fertilizer in a given village, which is normalized by the estimated standard error and denoted $\alpha_{f,v}$. R_{vt}, a dummy for reallocation in each village-year, is then regressed on the mean return to lagged fertilizer, standardized to have mean zero and standard deviation one.

$$R_{vt} = \beta\alpha_{f,v} + \epsilon_{vt} \tag{18}$$

The objective of this regression is to identify whether there is a cross-sectional correlation between the returns to lagged capital and the probability of reallocation: more specifically, whether there are fewer reallocations when the returns to lagged capital are higher and thus the cost of a reallocation in terms of foregone capital investment is higher. The standard errors are calculated by bootstrapping the two-step procedure (the estimation of the production function and the estimation of equation (18)) with clustering at the provincial level.

The results can be found in Table 7. In Panel A, the estimated coefficients capturing the correlation between returns to lagged capital estimated using OLS and the reallocation probability are significant and negative, consistent with the hypothesis that village officials respond to variation in the costs of reallocation. A one standard deviation increase in the returns to lagged fertilizer leads to a decline in the probability of reallocation of around 5 percentage points on a base probability of roughly 33%, a proportional effect of around 15%. These results are consistent across specifications in which the value of the grain harvest and value-added in grain production are employed as the dependent variable, outliers in returns to capital are trimmed, and the sample is restricted to only rice and wheat producers.

As a robustness check, equation (17) is re-estimated allowing for cross-sectionally varying lagged returns to agricultural labor and area sown, as well as fertilizer. The coefficient on lagged fertilizer and the mean coefficient on all lagged inputs are then employed as the independent variable in (18), and the results are shown in Panel B of Table 7. The coefficients are again negative and of comparable magnitude.

Panel C of Table 7 shows the results employing the return to agricultural inputs estimated using the dynamic panel methodology; details on the implementation of the bootstrap in this case can also be found in the appendix. The coefficients are generally of similar sign and magnitude, though the results employing grain value as the dependent variable in the production function are noisy. This suggests that the correlation between returns to lagged agricultural inputs and the frequency of reallocations does not reflect any systematic bias in the estimation of the production function.

Panel D shows one final robustness check in which the production function is re-estimated allowing the returns to inputs to vary at the level of the village v:

$$y_{it} = \sum_{j=1}^{J}(\alpha_j x_{ijt} + \alpha_f f_{ij,t-1}) + (\sum_{v=1}^{V}\nu_v) \times (\sum_{j=1}^{J}\alpha_j x_{ijt} + \alpha_f f_{ij,t-1}) + \mu_c + \nu_v + \gamma_t + \epsilon_{ivt} \quad (19)$$

This allows for the inclusion of province fixed effects in the estimation of equation (18), to test whether the correlation between returns to lagged fertilizer and reallocation behavior is evident within provinces. The results can be found in Panel E, and the coefficients are again negative and significant, though slightly smaller. Even within provinces, villages with higher returns to fertilizer report less frequent reallocations, consistent with the hypothesis that village leaders respond to the relative costs of reallocations is choosing whether or not to reallocate.

One potential challenge to this conclusion would be the possibility of reverse causation: villages with frequent reallocations exhibit lower returns to agricultural inputs precisely because they reallocate land frequently. There are two responses to this argument. First, the agricultural production function is estimated only using data from households with continuous tenure on the same plot last year and this year; thus any direct effect of reallocations (transactional costs of swapping plots, for example) should not be a source of bias. Second, the reduced form results have already shown that areas with greater frequency of reallocations show lower levels of agricultural inputs. This should generate an upward bias in the estimated returns to inputs in areas with higher frequency of reallocations, a bias that runs in the opposite direction from the detected effect.

6.3 Estimating optimal fertilizer use

To conclude this analysis, the previous estimates of the returns to agricultural inputs with variation at the province and crop level are then used to solve the household's optimization problem and calculate the optimal level of fertilizer use with and without a reallocation, as specified by equations (13) and (15).

There are several challenges inherent in this exercise. First, this analytical framework allows for only a partial equilibrium solution to the household's optimization problem, assuming a uniform decline in labor and sown area across all households corresponding to the estimated effect in the two-stage least squares results for a household at mean risk of losing their plot in a reallocation year. Only the estimated decline in fertilizer is allowed to vary across households, predicted by the varying returns to fertilizer. A full general equilibrium analysis would allow both capital and labor inputs to vary and add a second equilibrium condition derived from the output market for grain and/or the returns to capital and labor in other sectors for the household. Due to data constraints and the non-competitive nature of the output market, fully specifying and estimating this equilibrium would be challenging and is beyond the scope of this analysis.

Second, the accurate estimation of the agricultural wage is not easy in this context. The market for agricultural labor in rural China is extremely limited, and only a small number of households in the data report use of hired labor in agriculture, or working as a hired laborer on another farm. The median wage among these observations is then imputed as the relevant agricultural wage for all households in all years. However, this estimated wage is high, to the point that valuing the mean quantity of household labor used on a self-sufficient farm at this wage would render the household's profits in agricultural production negative. This presumably reflects the fact that hired labor has very different characteristics from household labor, and is often engaged in very different tasks. This weakness of the data cannot be easily surmounted; however, the cross-sectional analysis remains unbiased under the assumption that there is no differential bias in the true agricultural wage across villages with different frequencies of reallocation.

The final challenge encountered in estimating F^{NR}, optimal fertilizer in the absence of a reallocation, is that it requires the estimation of next period's return to this period's fertilizer, a function of next period's agricultural inputs. Estimating the counterfactual F^{NR} in a reallocation year in fact leads to an overestimate of the foregone investment induced by a reallocation, a result shown in Appendix A. This will generate an upward bias in the mean estimated cost of a reallocation in villages that have more frequent reallocations, generating a bias against the hypothesized negative correlation between the frequency of reallocations and their cost.

I define ΔF as the difference between optimal fertilizer use with and without a reallocation:

$$\Delta F \equiv F^{NR} - F^{R} \tag{20}$$

ΔF as a percentage of F^R is estimated at the household level using the estimated returns to agricultural inputs from the ordinary least squares specification, equation (17).[16] ΔF is then collapsed to the village-year level, normalized to have mean zero and standard deviation one and denoted Dif_{vt}. The following regression is then estimated, parallel to equation (18); standard errors are clustered at the village level.[17]

$$R_{vt} = \beta Dif_{vt} + \epsilon_{vt} \tag{21}$$

Panel E of Table 7 shows the results. Standard errors are again bootstrapped in a two-step procedure, with both the returns to agricultural inputs and ΔF re-estimated in each iteration. The estimated coefficients are consistent in sign and magnitude, though again noisier employing grain value as the dependent variable: an increase in the difference in optimal capital investment between the reallocation and the non-reallocation case leads to a decline in the probability of reallocation.

These results can also be used to compared the predicted decline in fertilizer and in agricultural production to the estimated magnitude in the instrumental variables specification. The mean predicted decline in fertilizer across the four specifications shown is between 10 and 20 percent, with a predicted decline in grain output of at least 50 percent. The instrumental variables specification predicts that at the mean level of risk of plot loss in the year of the reallocation, a 40% decline in fertilizer and a 53% decline in agricultural production is observed. Accordingly, the predicted effects based on the model employing returns to agricultural inputs are similar in magnitude to the reduced form effects, and well within the estimated confidence interval.

The correlation between reallocation propensity and the estimated returns to agricultural inputs is consistent with the hypothesis that village officials are selecting into reallocation on the basis of its relative costs in foregone agricultural investment. Thus despite the fact that reallocations and the associated tenure insecurity generate substan-

[16]Outlier observations implying an increase in fertilizer of more than 100% are dropped.

[17]The sample size contracts because estimation of ΔF requires the observation of all agricultural inputs employed in the year of a reallocation as well as the year before and after. Accordingly, some villages are truncated from the sample.

tial costs, the decision by village leaders to implement them does not seem to reflect pure irrationality. This raises the question of what the benefits of reallocations are, for both officials and rural households, and whether the observed pattern of reallocations could in fact be optimal under certain conditions.

7 Estimating the village leader's objective function

The conceptual model used to frame this analysis stated that village officials will choose to reallocate when the benefits exceed the costs, where the benefits were hypothesized to be greater intravillage equity and rents, and the costs were the loss in output induced by tenure insecurity, and the transactional burden of conducting the reallocation. Given a parameterization of these elements, the observed pattern of reallocations can be used to infer the relative weights assigned to greater equity vis-a-vis foregone output and transactional costs in the village official's objective function using a revealed preference approach.

The benefits of reallocation B are measured by the increase in equity as a result of a reallocation.[18] Reallocations on average do not result in a decrease in static measures of inequality in land distribution (e.g., the Gini coefficient). This presumably reflects the fact that the majority of land transfers implemented in reallocations are plot swaps, rather than reconfigurations of plots. Accordingly, households swap positions in roughly the same overall static distribution of landownership.

For this reason, a dynamic measure of inequality is employed. B is parameterized as the total number of shifts in landownership deciles induced by a reallocation. For every household i in the village, ΔL_{ivt} is the absolute value of the shift in deciles of landownership induced by the reallocation, and B_{vt} is defined as follows.[19]

$$B_{vt} = \sum_{i=1}^{H} \Delta L_{ivt} \tag{22}$$

In years that did not experience a reallocation, the counterfactual benefit $\tilde{B}$ is defined as a random draw from a normal distribution that has the mean and standard deviation corresponding to the observed mean and standard deviation of reallocation benefit over all reallocations observed in that village. T, the topographic burden of reallocation, is measured as mean elevation in the village.

[18]Unfortunately, there is no data available that would allow the econometrician to measure the benefit of a reallocation to the official in either promotion possibilities or rents.

[19]For example, a household that moved from decile 9 of landownership to decile 7 has $\Delta L_{ivt} = 2$, as does a household that moves from decile 3 to decile 5.

C, the cost of reallocation, is the estimated difference between output in the case of a reallocation and output in the absence of a reallocation. This difference is calculated using the decline in sown area, labor and fertilizer estimated in Section 6.[20] C_{vt}, the total bargaining cost of a reallocation in village v in year t, is the sum of the difference in output across all H households observed in the village, valued at the market price in hundreds of yuan.

$$C_{vt} = \sum_{i=1}^{H} \Delta Y_{ivt} \tag{23}$$

The net benefit of reallocations ψ_{vt} is then defined as a simple quadratic function of the benefits and costs. The weight on the quadratic function of C, lost revenue due to decreased agricultural output, is normalized to one.

$$\psi_{vt} = \alpha_1(B_{vt} + B_{vt}^2) - \alpha_2(T_{vt} + T_{vt}^2) - (C_{vt} + C_{vt}^2) \tag{24}$$

If $\psi \geq 0$, then a reallocation is optimal; if it is less than zero, a reallocation is not optimal. For postulated values of α_1 and α_2, a distribution of optimal reallocations can be generated and compared to the observed distribution of reallocations. The objective is to identify parameter values that best reproduce the observed pattern of reallocations.

More specifically, I wish to identify parameters that maximize the accurate prediction rate across all (reallocation and non-reallocation years), as well as minimizing the difference in accurate prediction rates between reallocation and non-reallocation years.[21] Define π_T as the percent of all reallocation and non-reallocation events that are accurately predicted by the postulated parameters, and π_R and π_{NR} as the percent of reallocations and non-reallocations that are accurately predicted, calculated separately. The goal of this exercise is to maximize $\hat{\pi}$, defined as

$$\hat{\pi} = \pi_T + \|\pi_R - \pi_{NR}\| \tag{25}$$

$\hat{\pi}$ is maximized by performing a grid search across potential values of α_1 and α_2. The range of parameters tested is 0 to 100 for both parameters; the increments of the grid are varied for each four specification, and reported in the results table. Standard errors are

[20] Note that this specification allows for cross-sectional variation only in the impact of reallocation on fertilizer inputs, not in the impact of reallocation on all agricultural inputs. While a more general framework would allow the household to equate the relative returns to both capital and labor inputs to the relative returns to those inputs in non-agricultural activities, the same caveat presented in Section 6 about the challenges of estimating a full general equilibrium solution to the household's optimization problem applies.

[21] If the objective was defined simply as maximizing π_T, given an observed reallocation rate of around one third, the resulting parameters predict the observed non-reallocations with a high degree of accuracy, while having little predictive power for reallocations.

bootstrapped across one hundred replications with re-sampling at the village-year level. For each specification, α_1 and α_2 are reported as well as π_R and π_{NR}.

Following this optimization process, I infer the predicted distribution of reallocations, conditional on the estimated weights, that would be optimal from the perspective of the official: namely, reallocating only when the net benefit is positive. I can then compare the estimated costs of the observed reallocations to the optimal reallocations. The difference ΔC as a percentage of the cost of the observed reallocations, is reported in the final row of Table 8.

The results show that the estimated weight on greater equity in the village leader's objective function is around 60, while the estimated weight on the transactional burden imposed by elevation is less than one. Converting these estimates to more easily understandable magnitudes, at the median level of C and B, the marginal benefit of an additional decile shift in landownership for a single household is valued at about six times the marginal benefit of one dollar less in foregone output as a result of a reallocation. At the median level of T, the marginal benefit of one meter lower in elevation is valued at about 10% the marginal benefit of one dollar less in foregone output.

The estimated parameters predicts around 50% of the observed events, both reallocations and non-reallocations, for a predicted reallocation probability of 49% (relative to the observed reallocation probability of 33%). Nonetheless, comparing the implied distribution of optimal reallocations given these weights and the observed distribution, the foregone output as a result of reallocations would be around 45% less if village leaders reallocated only where the net benefit as estimated by ψ were positive.

The results suggest first, that the observed pattern of reallocations is consistent with village leaders placing a high weight on the benefit of greater equity compared to potential output losses. Second, even given this greater weight on equity, and despite the fact that village leaders are partially optimizing the choice of reallocations, they are also making significant and costly errors. The objective of enhanced equity could be achieved at considerably lower cost given a different set of reallocation decisions.

8 Conclusion

Although secure property rights are perceived as immensely important to economic development, the literature on the impact of inframarginal variation in property rights on economic outcomes remains relatively sparse. This paper contributes to this literature by evaluating one of the most unusual and far-reaching experiments in land property rights over the last half-century, the system of village-based reallocations of land in China. Implemented in order to maintain relative equity among households and to allow for adjustment of landholdings in absence of any rural land market, this system generates periodic disruptions in property rights for rural households, who have no guarantee that they will continue to farm the plot they currently hold.

Using an identification strategy that exploits intra-village variation in security of tenure, as well as cross-village variation in the propensity to reallocate land, this analysis finds that a lower probability of land reallocation has a substantial impact on household's economic behavior. Households that are less likely to see their tenure on their current plot disrupted by virtue of their past inclusion in a reallocation employ more agricultural inputs and produce more output than other households, and this effect is of substantial magnitude.

At the same time, there is evidence that officials respond to variation in the costs of disrupting property rights in choosing whether or not to hold a reallocation. Village leaders are less likely to reallocate in villages where disruptions to property rights are costly, but they appear to make some significant mistakes in reallocating where the net benefit, even given a substantial weight on greater equity in the official's objective function, is negative. Thus while property rights institutions at a micro-level adapt to reflect the relative costs and benefits of different institutional structures in different economic contexts, this adaptation process is far from perfect.

9 Figures and tables

Figure 1: Map of sample counties

Figure 2: Land transfers

(a) Proportion land reported transferred

(b) Prop. households reporting land changes

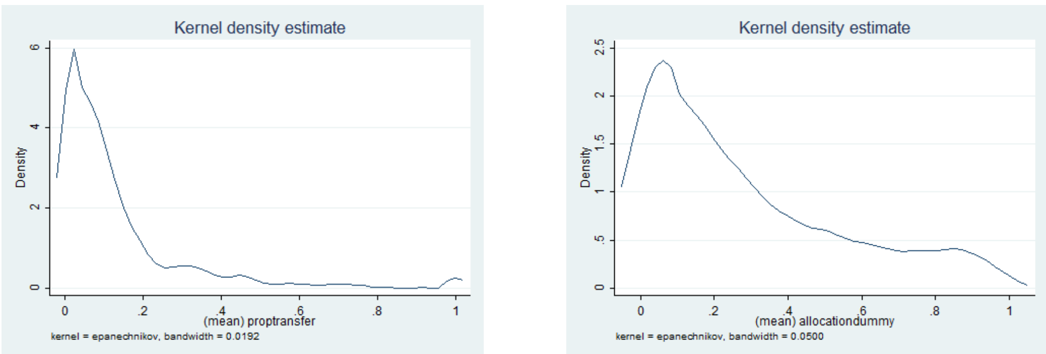


Figure 3: Probability of reallocation participation by decile of landownership

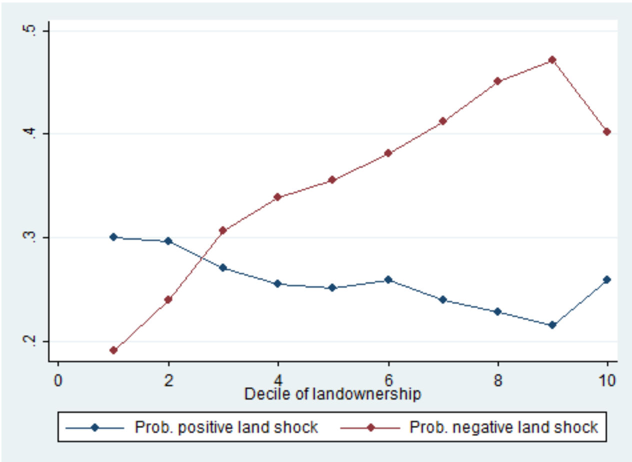
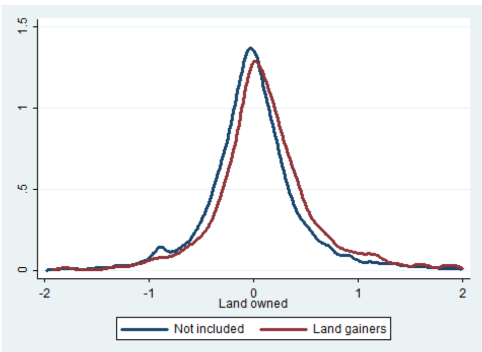


Figure 4: Kernel density estimates of landholding distributions

(a) Land gainers



(b) Land losers

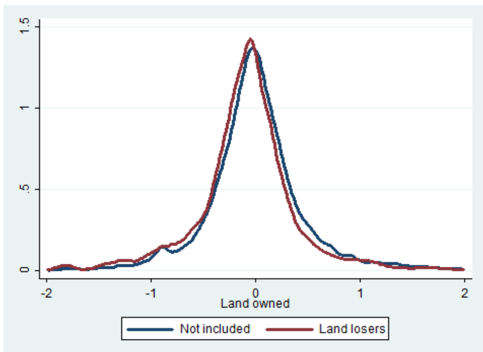
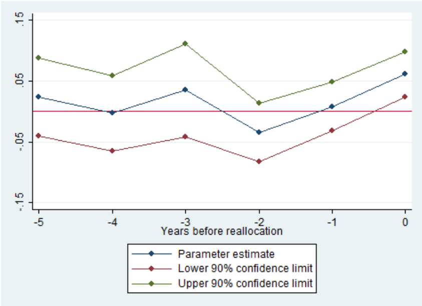
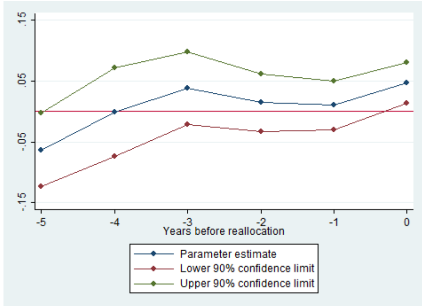


Figure 5: Anticipation of reallocation in pre-reallocation years

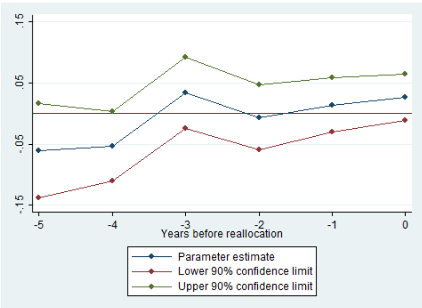
(a) Area sown



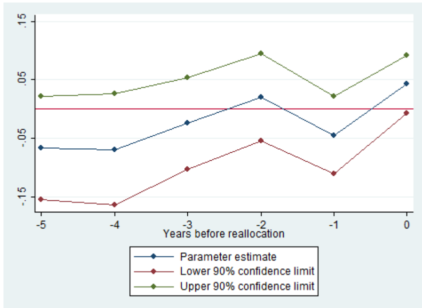
(b) Fertilizer



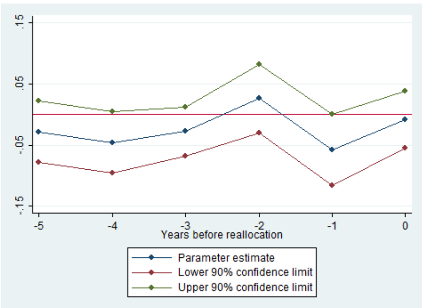
(c) Agricultural labor



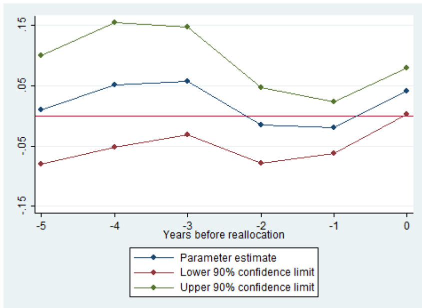
(d) Agricultural structures



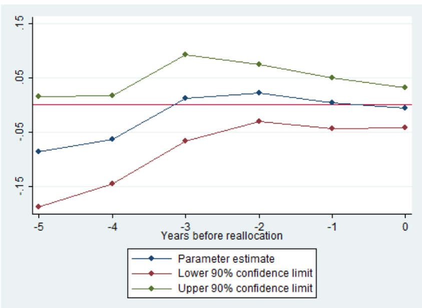
(e) Moveable capital



(f) Agricultural production



(g) Non-agricultural household business



(h) Outside labor

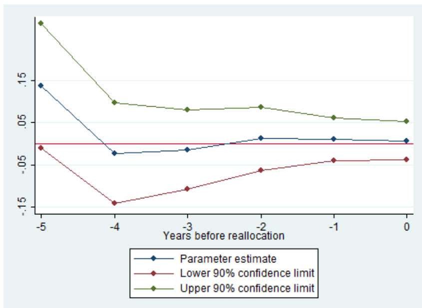


Table 1: Summary statistics

Village pop.	1690.794 (1134.75)
Village hh	417.90 (276.29)
Land per hh (hectare)	.40 (.37)
Plots per hh	5.96 (4.94)
Households sampled	68.42 (26.99)
Reallocation dummy	.53 (.50)
Forestry prop.	.22 (.29)

Table 2: Intravillage variation in reallocation probability: First stage

	Lagged participation	First stage
	(1)	(2)
$DP_{ivt}^{-1} \times R_{vt}$	-.038	
	(.014)***	
$DN_{ivt}^{-1} \times R_{vt}$	-.043	
	(.014)***	
$D_{ivt}^{-1} \times R_{vt}$		-.041
		(.013)***
Mean D_{ivt}	.169	.169
Mean $D_{ivt} \vert R_{vt} = 1$	.558	.558
F	5.436	10.831
Obs.	49376	49376

Notes: All specifications include village and year fixed effects interacted with R_{vt}, a control for vingtile of landownership also interacted with R_{vt}, and standard errors clustered at the village-year level. The independent variables in Column (1) are dummy variables equal to one if a household's land was reallocated positively (DP_{ivt}^{-1}) or negatively (DN_{ivt}^{-1}) in the previous reallocation interacted with a dummy equal to one if the village is holding a reallocation in that year, R_{vt}. The independent variable in Column (2) is the analogous interaction for a pooled dummy equal to one if the household had its land reallocated in a previous reallocation, D_{ivt}^{-1}. Asterisks denote significance at the 10, 5 and 1 percent level.

Table 3: Intravillage variation in reallocation probability: Reduced form

	Sown area (1)	Fertilizer (2)	Labor (3)	Structures (4)	Other cap. (5)	Agri. prod. (6)	Hh business (7)	Outside labor (8)
				Panel A: Reduced form				
$D_{ivt}^{-1} \times R_{vt}$	.067	.062	.041	.021	-.016	.062	-.005	-.011
	(.030)**	(.023)***	(.025)*	(.022)	(.023)	(.027)**	(.023)	(.026)
D_{ivt}^{-1}	.022	.013	.030	.007	.006	.016	-.00006	.031
	(.014)	(.011)	(.013)**	(.011)	(.009)	(.014)	(.013)	(.014)**
Obs.	46030	47841	46760	49376	49376	46465	49376	49376
				Panel B: Split reduced form				
$DP_{ivt}^{-1} \times R_{vt}$	.061	.042	.052	.031	-.012	.055	.016	.019
	(.036)*	(.024)*	(.029)*	(.026)	(.034)	(.032)*	(.030)	(.033)
$DN_{ivt}^{-1} \times R_{vt}$	.072	.074	.035	.015	-.018	.067	-.017	-.029
	(.034)**	(.026)***	(.028)	(.025)	(.025)	(.031)**	(.025)	(.029)
DP_{ivt}^{-1}	-.020	-.009	-.003	-.008	.015	-.020	-.002	.028
	(.015)	(.012)	(.015)	(.014)	(.013)	(.015)	(.015)	(.016)*
DN_{ivt}^{-1}	.047	.027	.051	.016	.00004	.038	.001	.032
	(.016)***	(.012)**	(.015)***	(.013)	(.011)	(.016)**	(.015)	(.016)**
Test: $\beta_1 = \beta_2$	.742	.165	.640	.651	.888	.702	.250	.139
Obs.	46030	47841	46760	49376	49376	46465	49376	49376

All specifications include village and year fixed effects interacted with R_{vt}, a control for vingtile of landownership also interacted with R_{vt}, and standard errors clustered at the village-year level. The independent variable in Panel A is the interaction between a pooled dummy equal to one if the household had its land reallocated in a previous reallocation, D_{ivt}^{-1}, and a dummy for a current reallocation R_{vt}. The independent variables in Panel B are the interactions between dummy variables equal to one if a household's land was reallocated positively (DP_{ivt}^{-1}) or negatively (DN_{ivt}^{-1}) in the previous reallocation and R_{vt}. The dependent variables are sown area, fertilizer, agricultural labor, agricultural structures, tools and animals owned, and dummies for participating in a non-agricultural business or in outside labor; sown area, fertilizer, agricultural production and agricultural labor are reported per acre owned, and all variables are normalized relative to the control observations following Katz, Kling & Liebman (2007). Asterisks denote significance at the 10, 5 and 1 percent level. The final row of Panel B reports the p-value for a test of equality on the coefficients of $DP_{ivt}^{-1} \times R_{vt}$ and $DN_{ivt}^{-1} \times R_{vt}$.

Table 4: Intravillage variation in reallocation probability: 2SLS estimates

	Sown area (1)	Fertilizer (2)	Labor (3)	Structures (4)	Other cap. (5)	Agri. prod. (6)	Hh business (7)	Outside labor (8)
Panel A: IV estimates								
Allocation dummy	-1.581	-1.347	-1.059	-.500	.376	-1.433	.114	.259
	(.829)*	(.633)**	(.742)	(.546)	(.544)	(.733)*	(.557)	(.621)
Obs.	46030	47841	46760	49376	49376	46465	49376	49376
Panel B: IV with polynomial in land area								
Allocation dummy	-1.604	-1.325	-1.055	-.457	.387	-1.410	.107	.243
	(.794)**	(.609)**	(.708)	(.525)	(.514)	(.693)**	(.535)	(.597)
Obs.	46030	47841	46760	49376	49376	46465	49376	49376
Panel C: IV excluding households with past demographic instability								
Allocation dummy	-2.993	-2.428	-2.580	-.612	.003	-1.916	-.896	.145
	(1.525)**	(1.139)**	(1.602)	(.866)	(.956)	(1.272)	(.900)	(.963)
Obs.	22724	23493	23177	24324	24324	22912	24324	24324
Panel D: IV excluding shrinking households								
Allocation dummy	-1.565	-1.388	-1.225	-.282	.565	-1.346	-.088	.280
	(.877)*	(.726)*	(.827)	(.567)	(.632)	(.826)	(.616)	(.666)
Obs.	40461	42108	41138	43514	43514	40855	43514	43514
Panel E: IV for households below median of landownership								
Allocation dummy	-1.530	-1.291	-1.077	-.828	.220	-1.383	-.464	-.142
	(.870)*	(.618)**	(.894)	(.585)	(.419)	(.818)*	(.568)	(.637)
Obs.	22336	23510	22530	24877	24877	22554	24877	24877

Notes: All specifications include village and year fixed effects interacted with R_{vt}, a control for vingtile of landownership also interacted with R_{vt}, a control for D_{ivt}^{-1} and standard errors clustered at the village-year level. The independent variable is a dummy for a household having its land reallocated, instrumented by $D_{ivt}^{-1} \times R_{vt}$. The dependent variables are sown area, fertilizer, agricultural labor, agricultural structures, tools and animals owned, and dummies for participating in a non-agricultural business or in outside labor; sown area, fertilizer, agricultural production and agricultural labor are reported per acre owned, and all variables are normalized relative to the control observations following Katz, Kling & Liebman (2007). Asterisks denote significance at the 10, 5 and 1 percent level. In Panel B, a quadratic polynomial in land area is added. In Panel C, the sample is restricted to households with no past history of demographic shifts in reallocation years; in Panel D, the sample is restricted to households that report either constant or increasing household size; in Panel E, it is restricted to households in the lowest five deciles of landownership.

Table 5: Information as a channel for predicting reallocations

	Allocation pos. (1)	Allocation neg. (2)		Allocation pos. (3)	Allocation neg. (4)
Rice int.	.136 (.021)***	-.110 (.026)***	Retail int.	-.046 (.035)	.018 (.037)
Wheat int.	-.043 (.036)	-.097 (.041)**	Fish int.	.056 (.045)	-.007 (.052)
Husb int.	.133 (.016)***	-.135 (.020)***	Educ. int.	-.020 (.020)	.047 (.021)**
Manu int.	-.020 (.037)	-.043 (.046)	Vet int.	-.064 (.112)	.051 (.062)
Trans int.	.001 (.079)	.068 (.095)	Grandparent int.	.039 (.019)**	-.030 (.021)
Cons int.	.022 (.147)	-.212 (.166)	Party int.	-.022 (.011)**	.042 (.011)***

Notes: Each cell corresponds to a separate regression including village and year fixed effects, a control for decile of landownership, and village and year fixed effects interacted with reallocation; standard errors are clustered at the village-year level. The dependent variable is a dummy for positive or negative changes in land in a reallocation as indicated; the independent variable reported is the interaction between a household dummy of interest and the mean of that dummy among government officials' households in that village-year. The dummy variables are indicators for whether the household engages in rice or wheat cultivation, or husbandry, manufacturing, transportation, construction, retail or fishing as a household business, as well as indicators for the presence within the household of a principal laborer with education beyond high school, a veteran of the armed forces, residential grandparents, or a member of the Communist party. Additional independent variables not reported are the household and official dummy entering linearly. Asterisks denote significance at the 10, 5 and 1 percent level.

Table 6: Agricultural production function

	OLS				Dynamic panel GMM			
	Full sample		Rice and wheat prod.		Full sample			
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Grain area	.620 (.022)***	.607 (.025)***	.576 (.029)***	.631 (.018)***	0.667 (0.192)***	0.664 (0.184)***	0.647 (0.178)***	0.611 (0.189)***
Labor	.126 (.013)***	.134 (.014)***	.130 (.019)***	.125 (.012)***	0.035 (0.193)	0.133 (0.185)	0.108 (0.180)	0.234 (0.191)
Fertilizer	.120 (.009)***	.118 (.010)***	.131 (.015)***	.102 (.009)***	0.324 (0.151)**	0.288 (0.102)***	0.308 (0.118)***	0.264 (0.106)**
Lagged fertilizer		.015 (.006)***		.014 (.006)**			0.122 (0.094)	0.158 (0.118)
Lags as instruments					Full set	Lags 3-4	Full set	Lags 3-4
Sargan-Hansen p-value					0.017	0.011	0.195	0.095
Test of overidentifying restrictions					0.824	0.616	0.875	0.912
Obs.	51073	42460	33789	28383	30019	30019	30019	30019

Notes: Each column represents a separate regression estimating the returns to agricultural inputs; the dependent variable is value added in grain production (the grain harvest valued at the market price in each village-year minus the cost of seeds). Columns 1 through 4 report estimates of the returns to agricultural inputs estimated in an OLS specification with village, year and crop fixed effects; in columns 3 and 4, the specification is restricted to rice and wheat producers. Columns 5 through 8 report estimation results employing a dynamic panel GMM methodology, employing the full sample. Asterisks denote significance at the 10, 5 and 1 percent level.

Table 7: Returns to lagged fertilizer and reallocation probability

	All households				Rice and wheat producers			
	Value added		Grain prod.		Value added		Grain prod.	
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Panel A: Lagged fertilizer								
Returns to lagged fertilizer	-.097	-.092	-.090	-.084	-.069	-.038	-.080	-.073
	[.030]***	[.033]***	[.037]**	[.043]**	[.025]**	[.032]	[.032]**	[.042]*
Outliers trimmed	No	Yes	No	Yes	No	Yes	No	Yes
Obs.	896	847	811	733	896	811	811	761
Panel B: Lagged fertilizer, area and labor								
Returns to lagged fertilizer	-.106		-.043		-.088		-.058	
	[.036]**		[.035]		[.038]**		[.031]*	
Mean return to lagged inputs		-.058		-.045		-.068		-.046
		[.031]*		[.041]		[.034]**		[.044]
Outliers trimmed	No	No	No	No	No	No	No	No
Obs.	896	896	896	896	811	811	811	811
Panel C: Dynamic panel coefficients								
Returns to lagged fertilizer	-.072		.023		-.082		-.034	
	[.045]		[.041]		[.037]**		[.045]	
Outliers trimmed	No		No		No		No	
Obs.	848		644		816		585	
Panel D: Lagged fertilizer with variation at the village level								
Returns to lagged fertilizer	-.023	-.021	-.030	-.030	-.035	-.033	-.030	-.020
	(.009)***	(.020)	(.010)***	(.020)	(.015)**	(.019)*	(.020)	(.030)
	[]***	[]***	[]**	[]**	[]**	[]	[]**	[]*
Fixed effect	Prov.	Prov.	Prov.	Prov.	Prov.	Prov.	Prov.	Prov.
Outliers trimmed	No	Yes	No	Yes	No	Yes	No	Yes
Obs.	853	845	844	836	837	831	818	812
Panel E: Optimal capital								
Optimal capital	-.048		-.016		-.062		.002	
	(.022)**		(.025)		(.024)***		(.014)	
Obs.	704		566		490		564	
Outliers trimmed	Yes		Yes		Yes		Yes	

Notes: The dependent variable is reallocation at the village-year level; all standard errors are clustered at the province level, and the independent variables are standardized to have mean zero and standard deviation one. The independent variable in Panels A through D is the mean estimated return to lagged fertilizer or to other lagged inputs in the village-year, normalized by the standard error. In Panels A and B, returns to agricultural inputs are allowed to vary by province and crop, and value-added or grain value denotes the dependent variable. In Panel C, the coefficients in the agricultural production function are estimated using dynamic panel GMM. In Panel D, returns to agricultural inputs are allowed to vary by village. In Panel E, the independent variable is the difference in optimal fertilizer input with and without a reallocation, calculated using a model of household optimization and the estimated returns to agricultural inputs estimated in Panel A, and again collapsed to the village-year mean.

Table 8: Parameters of the village leader's objective function

	(1)	(2)	(3)	(4)
α_1	57	57	56.75	56.75
	(14.5199)***	(20.621)***	(13.821)***	(11.469)***
α_2	.5	.25	.5	.25
	(.165)***	(.185)	(.171)***	(.092)***
π_R	0.516	0.516	0.516	.516
	(0.024)***	(0.026)***	(0.026)***	(.024)***
π_{NR}	0.513	0.513	0.513	.518
	(0.024)***	(0.026)***	(0.025)***	(.024)***
Δ^C	-0.452	-0.452	-0.452	.451
	(0.094)***	(0.097)***	(0.093)***	(.098)***
Grid step (α_1)	.5	.5	.25	.25
Grid step (α_2)	.5	.25	.5	.25

Notes: the coefficients correspond to the estimated weights on equity and transactional costs in the village leader's objective function. π_R and π_{NR} are the proportion of reallocation and non-reallocation events respectively predicted by the estimated parameters. ΔC is the percent difference in cost between the observed distribution of reallocations and the optimal distribution of reallocations conditional on the estimated weights. Standard errors are calculated using a bootstrap with 100 replications.

A Household production function under conditions of uncertainty

Assume the household seeks to maximize value-added profits in agricultural production (i.e., profits minus the cost of seeds); the production function is not constrained to be constant returns to scale, and evidence suggests it is in fact decreasing returns to scale. I postulate a standard Cobb-Douglas production function in which there are lagged returns to capital (fertilizer). Note that fertilizer is assumed to be a flow variable: F_t is equal to fertilizer applied in period t only. However, fertilizer applied in period t-1 is allowed to continue to have a direct effect on soil productivity.

γ is equal to the probability of reallocation; in the case of a reallocation, lagged returns to fertilizer are lost. Accordingly, the production function and value-added profits take the following form. The contemporaneous return to fertilizer will be denoted α_C, and the lagged return to fertilizer denoted α_F.

$$
\begin{aligned}
Y_t &= \tilde{A}_t L_t^{\alpha_L} N_t^{\alpha_N} F_t^{\alpha_C}((1-\gamma)F_{t-1}^{\alpha_F}) && (26) \\
\pi_t &= P_t Y_t - P_t^s S_t && (27)
\end{aligned}
$$

Assume further that the household optimally chooses F_t and N_t, fertilizer and labor inputs, and that land cultivated L_t is a mechanical function of inputs chosen: i.e., when a household optimally uses more inputs, it will cultivate more of its land allotment. For simplicity of notation, in the subsequent analysis I will denote $A_t = \tilde{A}_t L_t^{\alpha_L}$.

Define σ as the return next period to this period's capital investment in the absence of a reallocation.

$$
\begin{aligned}
\sigma_t &= \frac{\partial \pi_{t+1}}{\partial F_t} && (28) \\
&= A_{t+1}\alpha_F F_t^{\alpha_F - 1} F_{t+1}^{\alpha_C} N_{t+1}^{\alpha_N} && (29)
\end{aligned}
$$

The first-order condition governing optimal capital and labor can then be written as follows.

$$
\begin{aligned}
\frac{w}{r} &= \frac{P_t \alpha_N A_t F_t^{\alpha_C} N_t^{\alpha_N - 1}}{P_t \alpha_C A_t F_t^{\alpha_C - 1} N_t^{\alpha_N} + P_{t+1}\sigma_t(1-\gamma)} && (30) \\
0 &= \frac{P_t \alpha_N A_t F_t^{\alpha_C} N_t^{\alpha_N - 1}}{P_t \alpha_C A_t F_t^{\alpha_C - 1} N_t^{\alpha_N} + P_{t+1}\sigma_t(1-\gamma)} - \frac{w}{r} && (31)
\end{aligned}
$$

Denoting the right-hand side of (31) by ψ, the implicit derivative of capital with respect to the probability of reallocation can be calculated as $\frac{\partial F}{\partial \gamma} = -\frac{\psi'(\gamma)}{\psi'(F)}$. For ease of notation, define λ_1 and λ_2:

$$\lambda_1 = P_t \alpha_N A_t F_t^{\alpha_C} N_t^{\alpha_N - 1} \tag{32}$$

$$\lambda_2 = P_t \alpha_C A_t F_t^{\alpha_C - 1} N_t^{\alpha_N} + P_{t+1}\sigma_t(1-\gamma) \tag{33}$$

The implicit derivative can then be calculated employing the following formula:

$$\begin{aligned} \psi'(\gamma) &= \frac{\lambda_1}{P_t \alpha_C A_t F_t^{\alpha_C - 1} N_t^{\alpha_N} + P_{t+1}\sigma_t(1-\gamma)^2} \\ \psi'(F) &= \frac{1}{\lambda_2^2}(\lambda_2 P_t \alpha_N \alpha_C A_t F_t^{\alpha_C - 1} N_t^{\alpha_N - 1} \\ &\quad - \lambda_1 P_t \alpha_C(\alpha_C - 1) A_t F_t^{\alpha_C - 2} N_t^{\alpha_N} \\ &\quad - \lambda_1 P_{t+1}(1-\gamma) A_{t+1} \alpha_P (\alpha_F - 1) F_t^{\alpha_F - 2} F_{t+1}^{\alpha_C} N_{t+1}^{\alpha_N}) \end{aligned} \tag{34}$$

Given that both the numerator and denominator are positive, the implicit derivative formula yields that $F'(\gamma) < 0$. This is intuitive: optimal fertilizer investment declines when the probability of reallocation increases.

Assuming that households know whether or not a reallocation will occur at the end of the year at the time they make their investments (i.e., γ is a dummy variable equal to zero or one), optimal capital investment with or without a reallocation can be specified as follows. $F^R = \frac{w\alpha_C}{r\alpha_N} N$, while F^{NR} solves the following equation:

$$0 = \frac{P_t \alpha_N A_t F_t^{\alpha_C} N_t^{\alpha_N - 1}}{P_t \alpha_C A_t F_t^{\alpha_C - 1} N_t^{\alpha_N} + P_{t+1}\sigma} - \frac{w}{r} \tag{35}$$

Define the difference in investment between the reallocation and the non-reallocation case as follows.

$$\Delta F \equiv F^{NR} - F^R \tag{36}$$

ΔF is increasing in α_F, a comparative static that can be established again using the formula for implicit differentiation. Note that F^R is independent of the returns to lagged capital, while $\frac{\partial F^{NR}}{\partial \alpha_F} = -\frac{\psi'(\alpha_F)}{\psi'(F)}$. The denominator is positive, and the numerator can be written as follows:

$$\psi'(\alpha_F) = -\frac{\lambda_1}{\lambda_2^2} P_{t+1}[A_{t+1} F_t^{\alpha_F - 1} K_{t+1}^{\alpha_C} N_{t+1}^{\alpha_N} + A_{t+1}\alpha_P F_{t+1}^{\alpha_C} N_{t+1}^{\alpha_N} F_t^{\alpha_F - 1}(log F_t)] \tag{37}$$

Accordingly, $\frac{\partial F^{NR}}{\partial \alpha_F} > 0$ and thus $\frac{\delta \Delta F}{\partial \alpha_F} > 0$. The investment gap between years with and without reallocation is increasing in the returns to lagged capital.

In the empirical exercise, σ_t in reallocation years is estimated from the data using observed inputs in period t+1. In the data it is evident that γ_{t+1}, the probability of reallocation in period t+1, is lower after a reallocation has been conducted at time t, i.e. $E[\gamma_{t+1}|R_{vt} = 1] < E[\gamma_{t+1}|R_{vt} = 0]$. This suggests that the direction of bias should be that agricultural inputs in year t+1 following a year with a reallocation would be higher than the counterfactual agricultural inputs in year t+1 when year t had no reallocation. The intuition for this result is that in the absence of a reallocation, households would remain at risk of a future reallocation. Accordingly, the partial derivative estimated from the data $\frac{\partial Y_{t+1}}{\partial F_t}|R_{vt} = 1$ will be higher than the counterfactual, $\frac{\partial Y_{t+1}}{\partial F_t}|R_{vt} = 0$.

What bias does this generate in optimal F^{NR}? Returning to the first-order condition, we can see by inspection that if $\tilde{\sigma}_t = A_{t+1}F_{t+1}^{\alpha_C}N_{t+1}^{\alpha_N}$ is biased upwards, F^{NR} is also biased upwards. This can be confirmed by again employing the implicit derivative.

$$\psi'(\tilde{\sigma}_t) = -\frac{\lambda_1}{\lambda_2^2}P_{t+1}\alpha_F F_t^{\alpha_F - 1} \tag{38}$$

Accordingly, $\frac{\partial F}{\partial \tilde{\sigma}_t} = -\frac{\psi'(\sigma_c)}{\psi'(F)} > 0$. This suggests that estimating F^{NR} from the first-order conditions using observed data to estimate $\tilde{\sigma}_t$ will lead to an overestimate of F^{NR} in reallocation years, and thus an overestimate of $\Delta F = F^{NR} - F^R$. In the empirical analysis testing for a correlation between the estimated cost of reallocations and their frequency, this generates a bias against the expected negative correlation between the frequency of reallocations and their estimated cost.

B Dynamic panel estimation of the agricultural production function

Following Blundell & Bond (2000), an $AR(1)$ error structure is now imposed on the production function.

$$
\begin{aligned}
y_{it} &= \alpha_l l_{it} + \alpha_s s_{it} + \alpha_n n_{it} + \alpha_f f_{it} + \alpha_p f_{i,t-1} + \gamma_t + (\eta_i + \nu_{it} + m_{it}) && (39)\\
v_{it} &= \rho v_{i,t-1} + e_{it} && (40)
\end{aligned}
$$

This model has a dynamic representation:

$$
\begin{aligned}
y_{it} = {} & \alpha_l l_{it} - \rho\alpha_l l_{i,t-1} + \alpha_s s_{it} - \rho\alpha_s s_{i,t-1} + \alpha_n n_{it} - \rho\alpha_n n_{i,t-1} \\
& + \alpha_f f_{it} + (\alpha_p - \rho\beta_f) f_{i,t-1} - \rho\beta_p f_{i,t-2} + \rho y_{i,t-1} \\
& + (\eta_i(1-\rho) + e_{it} + m_{it} - \rho m_{i,t-1})
\end{aligned} \tag{41}
$$

The dynamic model can be rewritten as follows:

$$
\begin{aligned}
y_{it} = {} & \pi_1 l_{it} + \pi_2 l_{i,t-1} + \pi_3 s_{it} + \pi_4 s_{i,t-1} + \pi_5 n_{it} + \pi_6 n_{i,t-1} \\
& + \pi_7 f_{it} + \pi_8 f_{i,t-1} + \pi_9 f_{i,t-2} + \pi_{10} y_{i,t-1} + \gamma_t^* + (\eta_i^* + w_{it})
\end{aligned} \tag{42}
$$

subject to the following non-linear common factor restrictions,

$$
\begin{aligned}
\pi_1 &= -\pi_2/\pi_{10} && (43)\\
\pi_3 &= -\pi_4/\pi_{10} && (44)\\
\pi_5 &= -\pi_6/\pi_{10} && (45)\\
\pi_7 &= -\pi_8/\pi_{10} - \pi_9/\pi_{10}^2 && (46)
\end{aligned}
$$

as well as equalities in π_1, π_3, π_5, π_7, π_8 and π_{10}.

Given consistent estimates of the unrestricted parameter vector π and $var(\pi)$, these restrictions can be tested and imposed using a minimum distance model to obtain the restricted parameter vector.

B.1 Estimating the unrestricted parameter vector

The unrestricted parameter vector is estimated using dynamic panel methods; the following exposition largely follows Blundell & Bond (2000). A standard assumption on the initial conditions ($E[x_{i1}e_{it}] = E[x_{i1}m_{it}] = 0$ for $t = 2 \ldots T$) yields the following moment conditions.

$$E[x_{i,t-s}, \Delta w_{it}] = 0 \tag{47}$$

for $s \geq 3$ where $w_{it} \sim MA(1)$. This allows for the use of lagged levels of the variables as instruments after the equation is first-differenced.

However, the resulting GMM estimator in first differences can have poor finite sample properties when the instruments (lagged levels) are weak. Imposing additional conditions on the correlation between the fixed effect and first-differenced variables allows for the generation of additional moment conditions that can be used to estimate the parameters. The additional assumptions needed are as follows:

$$E[\Delta x_{it} \eta_i^*] = 0 \tag{48}$$

$$E[\Delta y_{it} \eta_i^*] = 0 \tag{49}$$

The moment conditions thus implied can be written as follows, for $s = 2$ when $w_{it} \sim MA(1)$.

$$E[\Delta x_{i,t-s}(\eta_i^* + w_{it})] = 0 \tag{50}$$

In other words, lagged first differences of the variables can be used as instruments in the equations in levels. Both sets of moment conditions can be employed in a linear GMM estimator using both first-differenced and levels equations; this is what Blundell-Bond deem the system GMM estimator.

B.2 Estimating the minimum distance model

The minimum distance model entails minimizing the distance between the unrestricted parameter vector and the previously enumerated set of common factor restrictions $g(\hat{\pi})$.

$$f(\beta, g(\hat{\pi})) = H\beta - g(\hat{\pi}) = 0 \tag{51}$$

where $g(\hat{\pi})$ can be written as

$$\begin{bmatrix} \pi_1 \\ -\pi_2/\pi_{10} \\ \pi_3 \\ -\pi_4/\pi_{10} \\ \pi_5 - \pi_6/\pi_{10} \\ \pi_7 \\ -\pi_8/\pi_{10} - \pi_9/\pi_{10}^2 \\ \pi_8 \\ \pi_{10} \end{bmatrix}$$

The minimum distance estimator is given by the minimization of

$$D(\pi) = f(\beta, g(\hat{\pi}))'\hat{V}[g(\hat{\pi})]^{-1} f(\beta, g(\hat{\pi})) \tag{52}$$

where $\hat{V}[g(\hat{\pi})]$ denotes the estimated variance-covariance matrix of $g(\hat{\pi})$, estimated using the delta method. Minimization of D yields the following:

$$\hat{\beta} = (H'\ \hat{V}[g(\hat{\pi})]^{-1}\ H)^{-1}\ H'\ \hat{V}[g(\hat{\pi})]^{-1}\ g(\hat{\pi}) \tag{53}$$

with variance-covariance matrix

$$\hat{V}[\hat{\beta}] = (H'\ \hat{V}[g(\hat{\pi})]^{-1}\ H)^{-1} \tag{54}$$

B.3 Estimating by province and crop

Estimating and imposing the minimum distance restrictions in an equation including interactions between the primary agricultural inputs and province and crop dummies imposes too large a computational burden. Accordingly, in the results restricted to rice and wheat producers, the model is estimated separately for each province-crop pair provided there are adequate observations. In the full-sample specification, the results are estimated for each crop-province pair for which there are adequate observations, and then for the remaining pool of households in that province.

The bootstrap is estimated by bootstrapping with replacement at the household level for each province-crop, estimating the agricultural production function for each province-crop, and then estimating the mean return to lagged fertilizer in each village-year. 100 replications are employed.

References

Banerjee, Abhijit, Paul Gertler & Maitreesh Ghatak. 2002. "Empowerment and efficiency: Tenancy reform in West Bengal." *Journal of Political Economy* 110(2):239–280.

Besley, Timothy. 1995. "Property rights and investment incentives: Theory and evidence from Ghana." *Journal of Political Economy* 103(5):903–937.

Besley, Timothy & Robin Burgess. 2000. "Land reform, poverty reduction, and growth: Evidence from India." *Quarterly Journal of Economics* 115(2):389–430.

Blundell, Richard & Stephen Bond. 2000. "GMM Estimation with persistent panel data: an application to production functions." *Econometric Reviews* 19(3):321–340.

Brandt, Loren, Jikun Huang, Guo Li & Scott Rozelle. 2002. "Land rights in rural China: Facts, fictions and issues." *The China Journal* pp. 67–97.

Chinn, Dennis. 1979. "Team cohesion and collective-labor supply in Chinese agriculture." *Journal of Comparative Economics* 3:375–94.

de la Rupelle, Maelys, Deng Quheng, Qi Shi & Thomas Vendryes. 2009. "Land rights insecurity and temporary migration in rural China.".

De Soto, Hernando. 2000. *The mystery of capital: Why capitalism triumphs in the west and fails everywhere else.* Basic Books.

Deininger, Klaus & Hans Binswanger. 1999. "The evolution of the World Bank's land policy: Principles, experience and future challenges." *The World Bank Research Observer* 14(2):247–76.

Feder, Gershon, Lawrence Lau, Justin Lin & Xiaopeng Luo. 1992. "The determinants of farm investment and residential construction in post-reform China." *Economic Development and Cultural Change* 41(1):1–26.

Field, Erica. 2005. "Property rights and investment incentives in urban slums." *Journal of the European Economic Association Papers and Proceedings* 3(2-3):279–290.

Frolking, Steve, Jianjun Qiu, Stephen Boles, Xiangming Xiao, Jiyuan Liu, Yahui Zhuang, Changsheng Li & Xiaoguang Qin. 2004. "Combining remote sensing and ground census data to develop new maps of the distribution of rice agriculture in China." *Plant and Soil* 16(4):89–99.

Galiani, Sebastian & Ernesto Schargrodsky. 2010. "Property rights for the poor: Effects of land titling." *Journal of Public Economics* 94(9-10):700–729.

Goldstein, Markus & Christopher Udry. 2008. "The profits of power: Land rights and agricultural investment in Ghana." *Journal of Political Economy* 116(6):981–1022.

Hornbeck, Richard. 2010. "Barbed wire: property rights and agricultural development." *Quarterly Journal of Economics* 125(2):767–810.

Jacoby, Hanan, Guo Li & Scott Rozelle. 2002. "Hazards of expropriation: Tenure insecurity and investment in rural China." *The American Economic Review* 92(5):1420–1447.

Katz, Lawrence, Jeffrey Kling & Jeffry Liebman. 2007. "Experimental analysis of neighborhood effects." *Econometrica* 75:81–119.

Keliang, Zhu, Roy Prosterman, Ye Jianping, Li Ping, Jeffrey Reidinger & Ouyang Yiwen. 2007. "The rural land question in China: Analysis and recommendations based on a seventeen-province survey." *International Law and Politics* 38:761–839.

Kung, James. 2000. "Common property rights and land reallocations in rural China: Evidence from a village survey." *World Development* 28(4):701–719.

Levinsohn, James & Amil Petrin. 2003. "Estimating production functions using inputs to control for unobservables." *The Review of Economic Studies* 70(2):317–341.

Lin, Justin Yifu. 1992. "Rural reforms and agricultural growth in China." *American Economic Review* 82(1):34–51.

North, Douglass & Robert P. Thomas. 1973. *The rise of the Western world: A new economic history*. Cambridge University Press.

Olley, G. Steven & Ariel Pakes. 1996. "The dynamics of productivity in the telecommunications equipment industry." *Econometrica* 64(6):1263–1297.

Rozelle, Scott & Guo Li. 1998. "Village leaders and land-rights formation in China." *The American Economic Review* pp. 433–438.

Schwarzwalder, Brian, Roy Prosterman, Ye Jianping, Jeffrey Riedinger & Li Ping. 2002. "An update on China's rural land tenure reforms: Analysis and recommendations based on a seventeen-province survey." *Columbia Journal of Asian Law* 16:141–225.

Unger, Jonathan. 1985. "The decollectivization of the Chinese countryside: A survey of twenty-eight villages." *Pacific Affairs* 58:585–606.

Unger, Jonathan. 2005. "Family customs and farmland reallocations in contemporary Chinese villages." *Social transformations in Chinese society* 82:113–130.

Walker, Kenneth. 1984. "Chinese agriculture during the period of the readjustment, 1978-83." *The China Quarterly* 100:783–812.

Wan, Guang & Enjiang Cheng. 2001. "Effects of land fragmentation and returns to scale in the Chinese farming sector." *Applied Economics* 33(2):183–194.

Wen, G.J. 1989. "The current land tenure system and its impact on long term performance of farming sector: The case of modern China.".

Zhang, F., J. Shen, L. Li & X. Liu. 2004. "An overview of rhizosphere processes related with plant nutrition in major cropping systems in China." *Global biogeochemical cycles* 260:1–10.