PROPERTY RIGHTS, LABOR MARKETS, AND EFFICIENCY IN A TRANSITION ECONOMY: THE CASE OF RURAL CHINA

Dwayne Benjamin
and
Loren Brandt

Department of Economics
University of Toronto

First (Preliminary) Draft: June 5, 1998
This Draft: April, 2000

The most recent draft of this paper is available at:
http://www.economics.utoronto.ca/benjamin/bb03.pdf

This draft has benefitted from the comments of participants at several seminars: ANU, Maryland, Michigan, Princeton, Toronto, Rochester, Southern California, Sydney, Western Ontario and the World Bank; as well as the Econometric Society Meetings (Australasian) in Canberra, 1998, and the CEA meetings in 1999. We especially thank Al Berry, Justin Lin, Albert park, Matt Turner, and Chris Worswick for their comments. Benjamin and Brandt thank the SSHRC for financial support. We also thank IDRC for their financial support of the collaborative data collection efforts with the Research Center for Rural Economy in Beijing.

Abstract

This paper investigates the consequences of imperfect and uneven factor market development for farm efficiency in rural China. In particular, we estimate the extent to which an inverse relationship in farm productivity can be attributed to the administrative (instead of market) allocation of land, and the extent of unevenly developed non-agricultural opportunities. Using a recently collected household survey, we show that a considerable amount of inefficiency exists in the countryside, especially in the employment of labor. Our results show that this inefficiency is alleviated by the development of external labor markets, but perhaps more surprisingly, that in the context of the current imperfect market environment, administrative reallocations help improve both efficiency and equity on the margin. They do not go far enough, however, which raises important questions about constraints on rental activity and the link between administrative reallocation and decentralized land exchange.

JEL Classification numbers: D23, J43, Q15, O12

I. Introduction

The process of economic transition entails the substitution of decentralized market exchange for centralized administrative allocation. The ultimate goal is the establishment of a market-oriented economy. By its nature, this process involves institutional change and the development by the state of an array of inter-connected institutions– economic, financial and legal, as well as a system of property rights– that underpin a market economy. Because of the time required for these "supporting" institutions to develop, transition can be lengthy, with some markets emerging more rapidly than others.

In some cases, market development is deliberately impeded by state policy. In rural China, for example, farm land is not allocated to farmers through a land-market, but is owned by the village and allocated by village leaders. This administrative allocation of land among households persists despite rapid product market development and decentralization of decision-making in farming to households. In principle, the administrative control over land allocation might be justified on equity grounds, especially to provide an equalizing effect on household income during transition. In practice, however, it may actually have adverse consequences for *both* efficiency and equity. The inability to trade even short term land use-rights, let alone accumulate privately owned land over longer time periods, may especially hurt those farmers with the poorest off-farm opportunities.

In analyzing economic transition and the appropriate mix of public policy, it is important to identify the possibly counterintuitive interactions of market and non-market based institutions. In this paper, we explore the effects of administrative land allocation, with its attendant weak land property rights, and factor market development more generally, on efficiency in China's rapidly changing rural sector.

A. Background

The introduction of the Household Responsibility System (HRS) in rural China in the early 1980s marked a radical change in property rights and organization in agriculture. With the extension of use and residual income rights to agricultural land to rural households– typically on the basis of family size, labor endowment, or some combination of the two– agriculture shifted from collective-based to family-based. The incentive effects of the institutional change, combined with price and marketing reforms, explain much of the spurt in growth between 1978-1984. Subsequently, growth fell considerably.[1] The HRS did not lead to decentralized decisions regarding land allocation: land ownership remained vested in the village. Moreover, as part of the administrative decentralization that accompanied this process, village cadre were given discretion over the allocation of non-residual control rights to farm households including crop selection, input use, the right to rent, as well as the security of tenure and freedom from the administrative reallocation of residual rights.

The HRS law originally called for secure tenure for fifteen years. However, in slightly more than two-thirds of all villages, village cadre have reallocated land amongst households at least once. Although in most villages there does not appear to be outright prohibition on the hiring of farm labor or land rental, these markets are thin in most areas.[2] Village-administered reallocation of use-rights, therefore, is the most important way that land is reallocated across households. Based on village survey work covering eight provinces, slightly less than half of all village land has been reallocated through these village-administered reallocations since the introduction of the HRS.[3] This "central-planning" method of land allocation stands in stark contrast to the growth in off-farm labor markets and self-employment opportunities, although access to these opportunities varies significantly across households.

B. Objective

Our objective in this paper is to examine the effect that administrative land allocation mechanisms combined with unevenly emerging off-farm labor markets have on farm inefficiency in rural China. Moreover, how do the effects of current land allocation schemes compare to those we might

---

[1] On this point, see Lin (1992)

[2] Survey work described in Turner, Brandt, and Rozelle (1999) suggests that only 3 percent of cultivated land is currently rented.

[3] See Turner, Brandt, and Rozelle (1999) for a detailed investigation of the extent and motivation for these reallocations. For a related discussion, see Carter *et al* (1996).

expect to observe under decentralized exchange via land rental. Farmers differ in their farming abilities, and more likely, in their alternative opportunities. Administrative, "per-capita" based allocations may not take these factors into consideration. Similarly, those with better off-farm opportunities might prefer to rent out their land, rather than to cultivate it themselves or give up rights to the land entirely. Other farmers with more limited opportunities might gain from the opportunity to expand their land-holdings. Without well defined property rights and a land rental market, these trades cannot occur between households. Unless labor and other markets work perfectly, the resulting allocation of land and labor will be inefficient. In fact, it is important to underline the need to consider the operation of other factor markets, even when focusing on land.[4]

This paper asks whether imperfect factor market development in China has significant adverse consequences for farm efficiency. In order to identify the linkages between land allocation mechanisms, labor market opportunities, and farm efficiency, we exploit the extent of an "inverse relationship" between farm output and farm size as an indicator of inefficiency. The inverse relationship is a common empirical regularity in developing country agriculture.[5] The conventional explanation is that farmers with small farms utilize their land more intensively than big farms, and thus produce more output per acre. One leading interpretation of this correlation is imperfect factor markets. With constraints in market opportunities for hiring out (or renting more land), farmers with less land are more constrained in their labor supply than bigger farmers. Because of their lower opportunity cost of labor, they use more labor per acre, and generate higher yields.[6]

Our empirical strategy rests no so much on the inverse relationship itself, as its covariation with factor market development. Is the inverse relationship more severe where markets are more poorly

---

[4] See Benjamin and Brandt (1997) for a discussion of the linkages between the distribution of land rights and the extent of land rental and labor markets in a historical context.

[5] See Berry and Cline (1979) for the classic presentation of this relationship from a variety of developing country contexts.

[6] Not surprisingly, there are other explanations. See Carter (1984), Bhalla (1988), and Benjamin (1995) for discussions of the role of omitted land quality in generating the inverse relationship, for example.

developed, or where land is allocated on a criteria other than efficiency?[7] Drawing on a recent household level survey carried out in North China, we find that there is a great degree of static inefficiency in Chinese agriculture that can be linked to imperfect factor markets. The development of off-farm opportunities has helped considerably in eliminating constraints facing farm households, and indeed is probably the best mechanism for increasing farm efficiency. Unfortunately, this development is not under the direct control of policy makers. Administrative land reallocations also appear to be playing a modest role in this regard. Households remain constrained, however, suggesting the need for the development of institutions that help reallocate land among households. An obvious candidate is a system of well-defined rental rights and the development of a decentralized land rental market. Since it is the smaller farms that are most constrained, our findings also suggest that decentralized reallocation could serve both efficiency and equity goals.[8]

The paper is organized as follows. In Section II we provide an empirical framework for analyzing the effect of market development on efficiency. This is followed by a discussion of empirical implementation in Section III. Section IV discusses our basic results while Section V is devoted to a number of caveats of interpretation. Finally, Section VI concludes.

II. Empirical Framework

It is much easier to describe the effect of markets on efficiency than to measure it. Very few data sets contain measures of "efficiency", nor do surveys have accurate measures of how well markets work.

---

[7] This basic approach has also been undertaken in other papers. For example, Burgess (1997) compares the degree of the inverse output relationship in two provinces with different degrees of market development. See also Udry (1996 and 1999). In particular, Udry (1996) points to the possible inefficiency deriving from imperfect property rights assignment between men and women within a household.

[8] Providing households a combination of well-defined rental rights and security of tenure would go far to eliminating the static inefficiency, and providing the proper dynamic incentives for investment. For more discussion of the role of land property rights, see: Besley (1996), for a prominent example, Binswanger, Deininger, and Feder (1995) for a survey of the related literature, and Feder and Deininger (1998) for some of the policy implications of this research.

We therefore exploit some simple economic theory to suggest ways in which inefficient behaviour can be captured empirically, and as well use a variety of realized individual and institutional outcomes to capture the extent and nature of land allocation and labor market opportunities. We begin by discussing how the inverse relationship can be exploited to identify the degree of inefficiency.

A. Efficiency and The Inverse Relationship

The link between markets, institutions, and farm efficiency is captured by the inverse relationship as follows.[9] We begin with a very simple model of farm production. Output, Q_i, is produced with land and labor according to a production function, $Q_i = F(h_i, L_i)$. We assume that land, h_i, is predetermined for the farmer (in this case by the village leaders), so his decision only involves choosing the profit maximizing amount of labor, L_i. In a world of imperfect markets, this decision can be characterized as setting the marginal product of labor equal to the (endogenous) shadow wage, q_i:

$$F'(L_i; h_i) = q_i \tag{1}$$

In general, the shadow wage may depend on a variety of variables, such as household characteristics (like demographic structure), farm characteristics (like farm size), and institutional variables (like the extent of off-farm opportunities, or the development of agricultural labor markets). If markets are complete, the shadow wage would be replaced by the market wage.

For simplicity, we assume that the production function is Cobb-Douglas, with constant returns to scale, so that we have:

$$Q_i = A h_i^{a} L_i^{(1-a)} \tag{2}$$

This functional form, while restrictive, is consistent with our empirical modelling framework, and

---

[9] For further discussion of the inverse relationship and farm efficiency, see Benjamin (1995) and Udry (1999). Benjamin (1992) also provides a discussion of labor demand with shadow wages, and the role of imperfect labor markets.

captures most of the features we are interested in. More general functional forms yield virtually the same qualitative results.

Taking logs of the first order condition yields the labor demand and output supply equations as functions of the shadow wage:

$$\ln L_i = b + \ln h_i - \frac{1}{\alpha}\ln\theta_i \tag{3}$$

$$\ln Q_i = c + \ln h_i - \frac{(1-\alpha)}{\alpha}\ln\theta_i \tag{4}$$

where b and c are constants that we ignore. The labor demand function is decreasing in the shadow wage, as is the output supply function. Notice that the wage elasticity of labor demand is larger in absolute value than the supply elasticity, unless " = 0, in which case only labor enters the production function. The most important implication of these equations, however, is that the elasticities of both labor and output with respect to land are one — doubling farm size should double labor input, and output. This is an implication of constant returns to scale, as well as the independence of the optimal input mix from scale (homotheticity).[10]

However, the implication that the land elasticity equals one is predicated on holding the shadow wage constant. What if the shadow wage depends on farm size? In fact, one common view of factor markets in developing countries is that the value of the farmers time is lower on small farms, reflecting constrained off-farm opportunities, and lower on-farm marginal productivity (because of diminishing returns). Small farmers thus cultivate their farms more intensively, reflecting the lower shadow value of labor. If the shadow wage is related to farm size as:

$$\ln\theta_i = d + \rho \ln h_i \tag{5}$$

Then substituting this into the labor demand and output supply equations yields:

[10] We explore the consequences for considering other inputs, especially capital, later in the paper.

$$\ln L_i = b' + \left(1 - \frac{1}{a} r\right) \ln h_i \tag{6}$$

$$\ln Q_i = c' + \left(1 - \frac{(1-a)}{a} r\right) \ln h_i \tag{7}$$

The elasticity of labor demand will then be less than one, the gap from one depending on D, the elasticity of the shadow wage with respect to land, as well as the importance of labor in the production function. Similarly, the output-land elasticity will also be less than one, though the gap from one is less than that in the labor equation, depending on the value of ". The inverse relationship is characterized by this gap in the output-land elasticity from one: small farmers cultivate their land more intensively, and output less than doubles as we double farm size. We can view D as capturing the degree of inefficiency caused by imperfect markets. The stronger the link between farm size and the shadow wage, i.e., the less shadow wages are related to markets, the greater will be the degree of the inverse relationship. We thus expect that the inverse relationship to be greater where markets are poorly developed, or where land allocation is most inefficient.

B. An additional source of inefficiency

The framework above focuses on the allocative inefficiency that follows from the absence of a common price allocating labor across farms. There may be other sources of inefficiency as well that may accentuate the inverse relationship, especially for labor. Consider the possibility that farmer effort also depends on the shadow wage. A measured day's work may not be the same on farms where farmers' opportunities differ. A simple way to capture this is to decompose the effective labor input, L_i , into the product of measured labor days, L_i^M , and labor efficiency, e_i:

$$L_i = e_i L_i^M \tag{8}$$

In logarithms we have:

$$\ln L_i = \ln e_i + \ln L_i^M \tag{9}$$

Substituting this into the labor demand equation, we see that measured labor demand is thus:

$$\ln L_i^M = b'' + \left(1 - \frac{1}{a} r - g r \right) \ln h_i \tag{10}$$

so that measured labor demand will be declining in labor efficiency for a given amount of effective labor. What this means is that measured labor demand may be affected by farm size for another reason. Not only will reduced farm size lower the shadow wage, and thus increase the optimal effective labor per acre, but if effort/efficiency also falls, this will further increase the measured days of labor per acre. More formally, if we write the efficiency relationship as:

$$\ln e_i = g + g \ln q_i \tag{11}$$

then, substituting equations (5) and (11) into (10), the labor-land elasticity will be reduced further:

$$\ln L_i^M = b'' + \left(1 - \frac{1}{a} r - g r \right) \ln h_i \tag{12}$$

while the output elasticity remains as before (it only depends on the response of effective labor to the shadow wage, not the decomposition of labor into effort and measured labor days).

We thus have at least two reasons why labor input, and output might be subject to an inverse relationship with respect to farm size. We have the further prediction that the inverse relationship can be expected to be stronger for labor, especially if the gap between effective and measured labor days is related to farm size through the shadow price of labor.

C. Empirical Implementation

Our empirical strategy rests on the identification of an inverse relationship, but more importantly, on estimating the covariation of the inverse relationship with various institutional and market features. The empirical framework follows the development outlined above. For example, we specify log-linear output and labor demand equations:

$$\ln Q_i = a_0 + a_1 \ln h_i + a_2 \ln q_i + a_3 ' Z_i + e_{Qi} \qquad (13)$$

$$\ln L_i = b_0 + b_1 \ln h_i + b_2 \ln q_i + b_3 ' Z_i + e_{Li} \qquad (14)$$

Z is a vector of other control variables that enter the production, while g is an unobserved error term. The economic model described in the previous section provides some predictions on the configuration of the land coefficients, depending on how well we control for the unobserved shadow wage.

Since the shadow wage is unobserved, we specify the following function based on observables:

$$\begin{aligned} \ln q_i = {} & l_1 + l_2 \ln h_i + l_3 ' D_i + l_4 ' Z_i \\ & + l_5 I_{Li} + l_6 I_{Li} \times \ln h_i + l_7 I_{Hi} + l_8 I_{Hi} \times \ln h_i + v_i \end{aligned} \qquad (15)$$

where D is a vector of household demographic variables, Z is a vector of control variables that enter the production function, and I_L and I_H are institutional variables describing the development of the land and labor markets. We also interact these variables with farm size, since the impact of farm size on the shadow wage might also depend on the extent of factor market development. In the extreme case of perfect markets, we expect farm size (and all other variables) to have no effect on the shadow wage, and for the shadow wage to equal the market wage.

Substituting the shadow wage into the equations above, and expressing effective labor in terms of measured days yields our main estimating equations:

$$\begin{aligned} \ln Q = {} & a_0 + a_1 \ln h + a_2 ' D + a_3 ' Z \\ & + a_4 I_L + a_5 I_L \times \ln h + a_6 I_H + a_7 I_H \times \ln h + v_Q \end{aligned} \qquad (16)$$

$$\begin{aligned} \ln L^M = {} & b_0 + b_1 \ln h + b_2 ' D + b_3 ' Z \\ & + b_4 I_L + b_5 I_L \times \ln h + b_6 I_H + b_7 I_H \times \ln h + v_Q \end{aligned} \qquad (17)$$

We focus on two main issues:

1) Whether, and to what extent, an inverse relationship exists in output and labor (how do α_1 and β_1 compare to one); and,

2) To what extent does this inverse relationship covary with institutional variables. This hypothesis can be tested by examining the coefficients on the interaction terms between farm size and institutions (α_5 and β_5 for the labor market, and α_7 and β_7 for the land market).

Note that we do not attempt to identify the theoretical parameters described in the underlying model. Instead, the model is used to interpret the magnitude of the inverse relationship across the labor and output equations.

III. Empirical Implementation

Our next task is to attach these notional empirical variables with ones based on real data. Especially with the institutional variables, this task is not straightforward.

A. Survey Data

The data that we draw on are the product of a household level survey that was carried out in 30 village and 6 counties in the northern provinces of Hebei and Liaoning in the summer of 1995.[11] The survey provides detailed household level income, expenditure, labor supply, and farm management data that are elaborated on below. The selection of the counties and villages was not entirely random. Each of the these counties was the site of an intensive household-level investigation carried out by Japanese investigators in either 1936 or 1937. In our re-survey, five villages in each of the six counties were selected, one of which had been fully enumerated in the 1930s. The other four villages in the county included one village located in the same township as the administrative capital of the county; one located in the same township as the village surveyed in the 1930s; and two villages drawn from a third township.

[11] The survey involved a collaborative effort involving Loren Brandt (University of Toronto), Paul Glewwe (World Bank), Scott Rozelle (U-C, Davis), and a team of researchers headed up by Bai Nansheng from the Research Center for Rural Economy, Ministry of Agriculture, Beijing.

The remaining township and villages were drawn to provide as representative of a cross-section as possible. A total of 130 households were surveyed in each county: Fifty from the village surveyed in the 1930s, and twenty from each of the remaining four villages. The households themselves were picked on the basis of random sampling using the most recent village registry.

B. Overview of the Villages

Agriculture, including farm sidelines are the primary source of income for households, and represents more than half of household total net income. All but a very small percentage of households engage in farming. Land per household ranges from slightly more than three mu (half an acre) to twenty-four mu (four acres). As reported in table 1, the mean cultivated farm size is ten mu, with farms larger on average in Liaoning than in Hebei. The smaller farm size in Hebei is largely offset by a longer growing season, and higher multiple cropping index.

The other two major sources of household income are family-run enterprises and wages earned from hiring off-farm. Slightly more than forty percent of all households report income from family-run enterprises, and more than half report wage earnings. On average, men work 264 days per year, and women 157 days. (These estimates of labor supply exclude housework.) The number of days supplied to farming and sidelines by the two sexes were similar (85 vs 81 days), while males worked considerably more off-farm and in family-run enterprises. Note that farm labor supply was slightly higher in Hebei than Liaoning, which largely reflects the better opportunities for multiple cropping in Hebei.

C. Quantifying the Institutional Environment

Estimation of equations (16) and (17) requires measures of I_H and I_L, the two institutional variables that describe the development of the land and labor markets. We begin with land.

Land

Although households enjoy use rights to land, they cannot buy or sell land. Ownership rights reside with the village. In principal, households have the right to rent their land to other households, however, only a small percentage of land is actually rented in these villages. Insofar as land is reallocated in these villages, the primary mechanism through which this occurs is in the form of a village-wide land reallocation process in which all or part of the land is taken back from households and reallocated to existing and possibly newly formed households.

Summarizing and interpreting these reallocations in a single village-level variable is not easy. From village-level surveys, for example, we know for each village the number of times that land has been reallocated since HRS was introduced as well as the size of each the reallocations. We also know the percentage of land that is currently farmed by the same households that received the land at the time of HRS. However, it is not just the size of the reallocations that likely matter for the efficiency of the current land allocation. Equally important is the *distributive* dimension of the reallocation of village land-holdings: How is the land being redistributed across households within the village? Is it going to a relatively small percentage of households, or is a much larger percentage of households benefiting from the reallocation?

For each household we collected information on their current land-holdings, and all of the changes that occurred in their cultivated area in the last three years, i.e. 1991-1994. The former comes in the form of detailed plot-by-plot information, including the number of years that the household has farmed each plot. The latter documents any increase or reduction in the last three years in land the household farms, and provides the reason for the change. As noted above, most of the household-level changes are associated with village-wide reallocations.

We experimented with a number of alternative measures of land reallocation, but in this paper only report the results from using a measure (Re_all) that is the product of two components[12]:

1. The percentage of village land reallocated within the past three years.[13] This will be higher in

[12] Our empirical results are not sensitive to the choice of reallocation measure. As will be seen, the "Re_all" effect is quite small, though this particular variable yields the largest effects of the measured we experimented with.

[13] A couple of points of qualification. First, the extent of reallocation in the past three years may not fully capture permanent differences in land policy between villages. For example, all villages may pursue the same policies, but differed (ex post) only in the timing of their reallocations. Second, we only measure changes in the size of the household land endowment. We net-out the fact that a household may have the same amount of land before

those areas where village governments are more “meddlesome” in allocating land; or alternatively where village leaders are reacting to opportunities to allocate land on an efficiency basis, or possibly on a more egalitarian grounds. On its own, the effects of this particular measure are difficult to predict.

2. The proportion of households receiving positive increases in allocated land. We expect this variable to be highest in those areas with more egalitarian-motivated reallocations.

Re_all is designed to capture a number of alternative village scenarios regarding land reallocation decisions. At one extreme is a policy of “benign neglect” in which villages either do not reallocate, or reallocate relatively small amounts of land. Re_all in this case will be near zero. Broad-based reallocations in which more land changes hands, and more people receive land from the reallocations will result in higher Re_all. There are two intermediate cases: One in which a relatively high percentage of households receive relatively small amounts of land and the second in which a relatively small percentage of households are targeted for an major increase in land-holdings. The former is consistent with an egalitarian-motivated land reallocation designed to make modest adjustments for household size, while the latter might be more efficiency-oriented, with the land perhaps targeted to the most efficient farmers. Of course, there is no way by looking at the reallocations alone to be sure of the motivation, or the consequences.

In Figure 1, we graph for each village the two components of Re_all: 1) the proportion of land reallocated by each village; and 2) the proportion of households that reported an increase in their cultivated holdings. Figure 2, on the hand, graphs the product of these two measures, Re_all, which is the focus of our empirical analysis (as a measure of I_H). The contrast between the provinces of Hebei and Liaoning is striking: Over the course of the last three years, administrative reallocation in villages in Hebei has been minimal, with little overall change in the allocation of landholdings. In a majority of the

---

and after a reallocation, but that this might be comprised of different plots. This consideration is especially important if we were looking at issues of “dynamic” efficiency.

villages in Liaoning, on the other hand, reallocations have been fairly widespread, with a sizable percentage of both land and households affected by reallocation behaviour. In table 1, we show the proportion of households, and the average amount of land, reallocated in the sample.

Off-farm Employment Opportunities

Outside of crop production, households found alternative outlets for their labor (as seen in Table 1). These include farm-related sidelines, family-run businesses, and hiring-out off the farm, usually in local enterprises. There is considerable heterogeneity across villages in employment patterns and connections with off-farm opportunities. Figure 3 illustrates some of this variation. In several villages, in upwards of a third of all household members worked off-farm; in nearly a quarter of all villages, however, the proportion was less than ten percent. In our estimation, I_L is measured at the *household level* by the percentage of household members working off-farm. This variable will reflect both variation of off-farm opportunities across villages, as well as variation across households within villages. Of course, because this is a household level variable, it might also be endogenous to the labor input equation. For this reason, we instrument this variable with: 1) Village aggregates of off-farm labor market participation (which excludes non-agricultural self-employment); and 2) Household educational attainment. These instruments will be evaluated using conventional tools (like overidentification tests).

D. Dependent Variables

Two other key variables required for the estimation of equations (16) and (17) are farm output and labor input in agriculture. Farm output is the total value of crop production measured at market prices.[14] The actual market prices that farmer's received for their crops were used for valuing production except in the case when no output was actually sold. In this case, village-level averages were used.

The survey provides several alternative estimates of household member labor supply to agriculture. The first asks each household member the number of days they supplied to agriculture during

[14] Farm output that was sold at quota prices, which are below market prices, was re-valued at market prices.

the peak and non-peak periods, and the average number of hours they worked per day in these two periods. The second collects disaggregated information by task performed, e.g. ploughing, weeding, fertilizer application, harvesting, etc. The two measures are highly correlated, and in our analysis we use the disaggregated labor input data.

Total labor employed in agriculture is the sum of the household's own labor and labor hired-in, either for a wage, or as exchange labor. Hired labor (for a wage) averaged less than five days per farm, and exchange labor was six days per year. In contrast, household supplied labor was 130.3 days. In Figure 4, we graph the average number of days of labor hired-in per household. These figures suggest quite uneven development of agricultural labor markets. In Hebei, hardly any labor was hired in. In parts of Liaoning, however, the percentage of labor hired-in was in upwards of 15-20 percent, most of which was seasonal in nature and related to rice cultivation.

E. Other Control Variables

The following is a list of the control variables that we use:

1. ln h: Log Household farm size (for all plots).
2. Demographic variables: a) ln Family Size (log household size); b) The proportions of male and female children (0 to 10 years of age), male and female teens (11 to 19 years of age), prime age males and females (20 to 59 years old), and the proportion of elderly men (60 and older; old women are the excluded group). These variables help control for the productivity of family labor, and should also be insignificant in the regressions if labor markets are perfect. The ln Family Size coefficient shows the percentage change in output or labor associated with a percentage change in household size, holding constant the age-sex composition. Each of the age-sex composition coefficients show the effect of swapping the specified family member for an elderly woman.
3. We include village fixed effects in some of the specifications, as well as farm level measures of land quality ("low, medium, high", irrigated), and village level measures of soil characteristics

(pH level, levels of potassium, nitrogen, and phosphates). These coefficients are not shown in the tables.

4. The key coefficients, as noted before, are the levels and especially the interactions of the market variables, “Off farm opportunities” and “Re_all”, with farm size. If the market variables attenuate the inverse relationship, then the coefficient of their interaction with farm size should be positive in the output and labor input regressions.

IV. Results

A. Household Land Allocation

We begin with an exploration of the determinants of land allocation — just how closely does the land allocation approximate a constant per capita distribution? Is there evidence that allocation is more sensitive to the number of men than women in household? Does our measure of the extent of village reallocation (Re_all) lead to a more “egalitarian” distribution of land? The results are reported in table 2. The reported coefficients are from a regression of log household allocated land on household demographic structure, as well as a variety of other control variables. In this and subsequent tables, we report the results from a variety of specifications.

In the first column, we show results for straight OLS, without any controls for village heterogeneity. If land is allocated on a purely per capita basis, we expect that the coefficient on log family size will equal one — a doubling of family size should double the family land allocation. In column one we see that the estimate is significantly less than one. Furthermore, we see some correlation between demographic composition and allocated land. For example, the number of children is negatively associated with allocated land, especially for female children. This need not reflect active discrimination against children, but instead reflect a delay in the adjustment of the family allotment in response to changes in family size.[15]

In the second column, we add village fixed effects. This will absorb cross-village heterogeneity in farm size that might also be correlated with family size. Unless farm and family size vary exactly proportionately across villages, failure to control for these differences will bias the land coefficient. Indeed, once we add the village fixed effects, we see that the land coefficient moves to around 0.9, and we cannot reject that the elasticity is one. This means that controlling for demographic composition, family land allocations vary across households within villages in direct proportion to their family size. Some of the demographic composition variables remain significant. In particular, it appears that village leader are slightly quicker to adjust family land to the addition of a male than female child (though the difference is not statistically significant).

In the third column, instead of village fixed effects, we control for measures of land quality. This will control for the spatial differences in the interpretation of farm size (a large farm of lousy soil will not be the same as a large farm with high quality soil), as well as intra-village differences in land quality. The results are similar to the fixed effects specification, at least for the family size variable. This suggests that the village land quality measures closely approximate the inclusion of fixed effects. Again, we cannot reject the hypothesis that land is adjusted in direct proportion to family size, though the coefficient is less than one.

In the fourth column we add the village land reallocation variable (Re_all), as well as the household measure of off-farm employment. If Re_all captures village "fine-tuning" of the land distribution to changes in family composition, we expect that the interaction term between Re_all and family size will be positive: the per capita relationship will be strongest, i.e., the coefficient on family size will be closest to one, in those villages with the most widespread reallocations. In fact, the coefficient is positive, but not very well determined (the standard error is large). On the other hand, the

[15] In most villages, adjustments for changes in family size occur as part of village-wide reallocations.

off-farm employment measure is much stronger. Households with greater involvement in non-agricultural activities are allocated less land per person. This suggests that at least to some degree, village reallocations take into consideration the importance of farming to the household, implicitly allocating land from the more non-agricultural households to those engaged solely in farming.

B. Inverse Relationship: Output and Labor

The main results for the inverse relationship, the estimation of equations (16) and (17), are presented in Tables 3 and 4. We begin with the output equation in Table 3. The first column shows the OLS estimates, without controls for land quality. The coefficient on log land is 0.833, which is significantly below one. This suggests that there is a strong inverse relationship in output. There is also a slightly significant positive coefficient on family size, suggesting that larger families also produce more output per acre, possibly because of higher labor input. However, both of these variables might be correlated with omitted measures of land quality. In the case of land, if small farms are systematically of better quality than large ones, then the land coefficient will be biased, at least in terms of the interpretation developed in the modelling section. To address this possibility, we first add village fixed effects, which absorb the village-level land quality characteristics. With these controls, the inverse relationship vanishes. The land elasticity is 1.009, and precisely estimated. In this specification, the demographic variables are also (generally) insignificant. In the third column, instead of village fixed effects, we add direct controls for land quality. Here, the results are equally striking: The inverse relationship also disappears. Apparently, the entire inverse relationship for output can be explained by omitted land quality.[16] Two conclusions thus emerge: First, there appears to be no inverse relationship in output (and thus no inefficiency), and second, the village measures of land quality appear to parallel closely the addition of village fixed effects.

In Table 4 we report the corresponding results for labor input. The conclusions differ sharply

---

[16] See Benjamin (1995) for a discussion of related regressions for rice farms in Indonesia.

from the output equations. In the first column, we see a dramatic inverse relationship, with the land elasticity less than 0.5. Furthermore, household demographic variables are significant determinants of farm labor use. As discussed in Benjamin (1992), this can be interpreted as further evidence that labor is inefficiently applied to the farms, and that trade in labor across farms is somehow restricted. These results, while slightly attenuated, are essentially the same once we add controls for land quality or village fixed effects. The land elasticity lies between 0.5 and 0.6, suggesting significant inefficiency in labor use. The difference in inverse relationships between the output and labor equations, however, suggests that most of the inefficiency is related to "effort" rather than the application of effective labor. If effective labor ($e\times L^M$) alone was subject to an inverse relationship, we expect to see the inverse relationship in labor carrying through to an inverse relationship in output. Of course, we also saw that the inverse relationship in labor can be more severe, depending on the share of labor in production. However, a gap this large (0.4 or 0.5) can only be explained if we allow the efficiency of labor, e, to increase with farm size.

Before turning to our estimates that exploit spatial variation in institutional development, it is worth illustrating our methodology with a simple figure. What's more, this figure essentially captures the main result of our paper. Our empirical methodology is based on identifying "systematic" variation of the degree of the inverse relationship across households facing possibly different constraints. One source of this variation is spatial differences in the degree of labour market development, and spatial (village) differences in land reallocation. In its simplest terms, we wish to see whether the inverse relationship is weakest where shadow wages are likely to be highest, that is, where markets are most developed.

The basic methodology is illustrated in the two panels of Figure 5. Here we plot the estimated *village level* inverse relationship against our village-level institutional variables.[17] In panel 1 we plot the

[17] The estimated coefficients are retrieved from a regression identical to the specification in column 2, with village fixed effects, but with interaction effects between village and farm size. These interaction effects yield the village specific relationship between farm size and labor intensity. As an aside, F-tests for the significance of the interaction effects suggest that the inverse relations does indeed vary significantly across villages.

estimated land coefficients for 30 villages against the village-level “Re_all”. As can be seen on the “y-axis” the land coefficients range from essentially zero, to one, with a mean around 0.5 (which is unsurprising, given the results in Table 4). However, there is no obvious correlation between the land coefficient (inverse relationship) and “Re_all”. To some extent, this is driven by the fact that Re_all exhibits only minimal variation, as already seen in Figures 1 and 2. The more interesting picture emerges in the second panel, where we plot the same land coefficients against our measure of off-farm employment opportunities. It appears that there is a generally positive relationship, suggesting that the inverse relationship is weakest in villages where off-farm employment opportunities are greatest. Of course, these figures are only suggestive, not least because there is no formal evaluation of the “fit” of the correlations. The interaction effects in Tables 3 to 5 are designed to test these relationships more formally, with more control variables, and including the possibility that the inverse relationship varies systematically across households within villages.

Returning to Table 3, and the output equation, we see that the land coefficient remains close to one, and none of the “market” type variables are significant. This means that the relationship between output and farm size does not vary across villages in these dimensions. In the last column, we address the possible endogeneity of household participation in off-farm employment. We use household education and village measures of labor market activity to instrument household participation in non-agriculture. These variables should be correlated with the household non-agricultural participation rate, but hopefully independent of the unobservables that jointly affect this decision as well as farm output. In this case, the 2SLS results line up with the OLS findings. Furthermore, in testing the validity of the instruments, we find that the instruments are significant in the first stage equation, but also can be excluded from the second stage estimation, i.e., subject to the general weaknesses of these tests, the instruments appear to be reasonable.

In the fourth column of Table 4, we add the market variables and their interactions with farm size to the labor equation. We find that widespread village reallocations of land tend to reduce the inverse relationship, as the interaction between Re_all and farm size is positive and slightly significant. Since we are controlling for the extent of off-farm employment, the interpretation is that in villages with widespread land reallocation, (for reasons independent of off-farm employment patterns, such as adjusting to demographic changes) inefficiency is reduced. Thus, uneven factor market development allows these (possibly, distributionally-motivated) reallocations to improve efficiency by providing employment for otherwise underemployed small farmers. The results for off-farm employment are even stronger. First, it is clear that less farm labor is used when the opportunity cost is high. This interpretation is not driven by time-allocation identities, since it is even stronger once we use 2SLS to address this possible source of endogeneity. Where households have alternative opportunities they use less labor on the farm. Secondly, the inverse relationship is significantly reduced for those farmers with outside opportunities, i.e, farm efficiency is higher where off-farm employment opportunities are most developed. Again, the 2SLS results strengthen this finding. Furthermore, the interpretation of the 2SLS overidentification tests is that the effect of education on farm production primarily lies in its generation of non-farm opportunities, and a more economical application of farm labor. These findings thus reinforce the interpretation of the inverse relationship as reflecting inefficiency, and especially highlight the role played by the land reallocation mechanism, and its interaction with off-farm employment opportunities.

C. Labor Productivity

In Table 5 we summarize these findings by looking a labor productivity — that is output per day worked. The coefficients in this table are simply the difference in coefficients between the output and labor regressions (Tables 3 and 4). In the first columns, we see that output per worked-day is significantly related to farm size: Small farmers produce significantly less per day worked. Secondly, larger households tend to produce less output per day worked. This is also consistent with these households having lower shadow values of labor. Once we add controls for the land and labor market, we recover the other findings. First, labor productivity (per day) is higher when farmers have other activities in which to spend time productively. Second, the relationship between farm size and productivity is significantly reduced in those areas with more widespread land reallocation, and greater off-farm opportunities. These results suggest that in the absence of well developed markets, village land reallocations alleviate some of the inefficiencies, but that considerable inefficiencies remain, especially where there are few off-farm alternatives.

D. How serious is the inefficiency?

Our basic finding is that there is no inverse relationship for output, but a strong one for labor input. This pattern is consistent with wasted, or unproductive labor akin to old-fashioned "surplus labor." One way to evaluate the implicit economic cost of this inefficiency is to value the wasted labor at a "reasonable" opportunity cost. In this section, we provide some "back-of-the-envelope" calculations, using our estimates of the degree of wasted labor implied by the inverse relationship. Our thought experiment is as follows: How much wasted labor do we see reflected in the inverse relationship; and how much income is thus forgone, assuming that there were off-farm opportunities. So, we are asking how much of the inefficient (wasted) labor can be freed for off-farm employment.

First, note that our methodology does not identify "absolute," but only "relative" inefficiency. We find that small farms use disproportionately too much labor relative to larger farms. All we know is that if farm size is cut in half, labor input is cut by less than half. In order to pin down an absolute level of inefficiency, we need to identify an "efficient" land-labor ratio, and then estimate the implicit excess labor applied by all of the smaller farms. Obviously, there will be some sensitivity of our conclusions to the choice of benchmark. But the basic idea should be clear.

Imagine dividing the sample into three farm sizes: small, medium, and large, according to the criteria in Table 6. In our sample, we label the "large" farms as those in the top decile, i.e. with $\ln h > 3.135$ (or 33 mu). Assume that these farms are efficient (in terms of labor input). Their mean log labor input is 5.203 (per year). If there were no inverse relationship, we would expect that the labor input of medium-sized farms, with an average $\ln h$ of 2.512 will be:

$$5.203 - 1.0 \times \Delta \ln h = 5.203 - (3.505 - 2.512) = 4.210$$

or 67 days per year. But, given the inverse relationship apparent in Table 3, we estimate that the actual average labor input is:

$$5.203 - 0.60 \times \Delta \ln h = 5.203 - 0.6(3.505 - 2.512) = 4.607$$

or 100 days per year, which is significantly more than predicted. Thus, for medium sized farms, we estimate that they are wasting 33 days per year of labor input (i.e., employing 33 days of labor without any corresponding additional output.) If the opportunity cost of labor is the daily wage of a manual worker (on average, about 20 Yuan per day), then this works out to 660 yuan per year, or about 8% of median household income (8362 Yuan) for medium-sized farms. For small-sized farms, we calculate that about 29 days of labor are wasted, which represents a higher share of their farm labor input, and 9.3% of median household income (7090 yuan) for small-sized farms.

In summary, the estimated inverse relationship for labor implies significant income losses, in terms of forgone labor income. Of course, the source of this income loss is unidentified in the data, but in keeping with the evidence of our estimation, is driven by the combination of poorly developed off-farm employment opportunities, and the inability to trade land use-rights.

V. Caveats of Interpretation

It is worth reviewing our findings to this point:

1. We find a significant inverse relationship for labor, but not for output.
2. Household size is strongly related to total labor input on the family farm.

Together, these findings are consistent with a systematic positive relationship between shadow wages and farm size, and the over-application of labor on small farms. Indeed, the relative magnitudes of the inverse relationships are such that we infer a considerable amount of "wasted time" in the labor input of smaller farmers. This interpretation is bolstered by the covariation of the inverse relationship with factor-exchange institutions:

3. More “egalitarian” land reallocations reduce the inverse relationship, suggesting that smaller farmers make better use of their labor in villages that reallocate in this manner.

4. Households with better off-farm opportunities use less farm-labor, and more importantly, better off-farm opportunities reduce the link between farm size and labor efficiency.

Taken together, these correlations are consistent with the story outlined in the first part of the paper. Nevertheless, alternative explanations exist. In this section we explore a number of these alternatives. Our objective is not to address each of these separate research questions in the depth they merit, but rather to assess whether there is evidence that suggests our interpretation is misleading. Our guiding question is whether we can really believe that the farmers in our sample are wasting their time to the extent we estimate.

A. Measurement Error of Farm Labor

There may be measurement error in reported labor input that is systematically related to farm size. For example, small farmers might exaggerate (overstate) their labor input. Perhaps there are indivisibilities in labor input (especially by task), and smaller farmers “round up” their reported days on the farm. In this case, we might observe a spurious inverse relationship for labor, but not output.

To some extent, this type of measurement error is reduced by having two independent measures of days spent on the farm: One based on a task-by-task summary of days on the farm (the measure we use); the other a question asking for a season-by-season accounting of total days spent in various activities (including aggregate days on the farm). Both measures yield virtually identical results.

However, there is a better way to check indirectly for farm-size related measurement error. If reported farm-days are exaggerated for small farmers, then total days worked in all activities will also be larger for small farmers. Stated differently, if there is no measurement error, so that a day spent on the farm represents one less day available for other activities, then total days worked and farm size will be unrelated (unless farm size affects labor supply).

We thus look at how total household labor supply varies with farm size, controlling for the same factors as in the other tables. We are interested in whether the coefficient on (log) land is negative, i.e., that smaller farmers work more total days, which is implied by the measurement error model. The results are shown in the first column of Table 7. First note that mean total days worked are 414 per household (median = 350), versus mean farm labor of around 144 (median=117). The coefficient on land is small and statistically insignificant. This suggests that per-capita labor supply is independent of farm size, so that there is no obvious evidence of exaggerated labor supply for small farmers. It thus appears that the opportunity cost of a day in the fields is a day in other activities for all farmers, both small and large. In other specifications (not shown) we find that where off-farm opportunities are greatest, farmers work more overall, but less on their farm, and this is consistent with our interpretation.

B. Seasonal Application of Labor

One conclusion we draw is that there are gains from reallocating land from larger to smaller farmers. The smaller farmers are wasting labor, and additional land will allow this labor to be put to good use. Underlying this interpretation is the evidence that labor is over-applied on small farms, with no corresponding increase in output. Another possibility is that labor is efficiently applied in the busy-season, but that time is wasted in the slack-season. If this is the case, extra land will not improve the efficiency of small farmers, since the wasted labor is entirely applied in the slack season, when the extra land will not make much difference. A simple way to test this alternative hypothesis is to disaggregate the labor input by busy and slack season. If the seasonal efficiency hypothesis is true, we will observe little or no inverse relationship in the busy season, but a significant one in the slack season. In Table 7 we conduct this exercise, looking at the separate application of family labor to the farm across seasons.

In the first columns for each type of labor (columns 2 and 4 in the table), we see that the inverse relationship is weaker in the busy season, but it is still large and significant. This suggests that the shadow wage of labor is higher in the busy season, which seems perfectly reasonable. Further support for this is provided by the coefficient on family size, which is larger in the slack season, suggesting that larger families waste more labor in the slack season. Regarding the specification with interaction terms, we find essentially the same results as before; however, there is some refinement in the way that the factor-allocation institutions affect labor application. During the busy season, off-farm employment opportunities have much the same effect as before (reducing the inverse relationship), but the effect is most pronounced in the slack season, when on-farm opportunities are lowest. In the slack season, land reallocations have no effect on reducing the inverse relationship, but in the busy season, it appears that extensive reallocations improve the efficiency of small farmers. This suggests that improved access to land can significantly improve the labor-efficiency of small farmers, especially in the busy season, when labor input is most valuable. In summary, the separate analysis by season reinforces the interpretation of the links between farm efficiency and factor-allocation institutions.

C. Endogenous Land Reallocations

Our interpretation of the interactions between farm size and the land reallocations suggests that the reallocations are (slightly) positively related to efficiency, principally through improved labor allocation. Can this be misleading? We are most concerned about getting the wrong-signed estimate. There are a couple of possibilities. First, if reallocations actually hurt productivity, but are only conducted where the efficiency cost is low (and productivity is high), then we will get biased, and potentially wrong-signed estimates. Alternatively, if reallocations are conducted instead where there is considerable inefficiency (say, where job opportunities are lowest) then the potentially positive effect of reallocation on efficiency will be understated.

One solution is to instrument these terms, as we instrumented the labor-market variables. We need to find instruments that are correlated with the frequency, extent, and nature of reallocations. This is a tall order, as the instruments must also be independent of the farm-productivity levels in the village. Our strategy is to use instruments that should be correlated with the preferences of village leaders, and the *costs* of reallocation (rather than the *benefits*). We use two different sets of instruments, one we call the “extended” set, and the other is a smaller “limited” set. The extended set includes 1) the number of plots in the village (which will increase the cost of reallocations); 2) an indicator of whether the township (not the village) makes the reallocation decision; 3) the number of households in the village (which will increase the administrative cost); 4) the number of production teams in the village; 5) the number of village cadres; 6) an indicator of whether the last village election was contested; and 7) the age, education, tenure, and farming-status of the village head and party secretary. For the limited instrument set, we use only indicators of whether the township reallocates and the number of households in the village. These variables are also all interacted with log farm size in order to aid the prediction of the interaction term between Re_all and farm size.

In Table 8 we show results from 2SLS estimation with the two sets of instruments. The "extended" set will have the benefit of predicting reallocations better, but riskier in terms of correlation with error term. The limited set is chosen to satisfy the overidentification test, but at the price of weaker predictive power in the first-stage regressions. In the first column for each specification we show the F-tests for how well the excluded instruments help predict each endogenous regressor in the first stage regressions. The instruments (extended and limited) are both significant predictors of the endogenous variables, though the extended set is much better. The extended set of instruments fails the overidentification test for output, but passes for labor (our main equation of interest anyway). The lean set (marginally) passes the overidentification test in both equations.[18]

That said, it can be seen that use of either set of instruments leads to qualitatively similar results, both with each other, and with the results in previous tables. The standard errors, however, are quite large, reflecting the imprecision with which we can estimate this effect. We find that the conclusions regarding off-farm employment opportunities are unaffected compared to the previous tables. Regarding the reallocation measures, using the extended set of instruments, we continue to find that the inverse relationship for labor is significantly reduced where reallocations are more extensive. With the limited

[18] Note that critical values could be adjusted upwards with use of Schwarz criterion (that accounts for sample size), in which case, overidentification tests would pass in all equations.

instrument set, the reallocation interaction effect is still positive, but is insignificant. Note, though, the very weak predictive power of the first stage equation for this variable.

Our conclusion from this exercise is that there is no evidence to suggest that our interpretation is misleading, or wrong-signed. Of course, our inability to find contrary evidence may merely reflect our failure to find excellent instruments for land reallocations. That said, it is still comforting (to us) that differing sets of reasonable instruments yield similar results, suggesting that our estimates our at least somewhat robust to the exogeneity assumptions about village land-reallocations.

D. Substitution with Capital

Perhaps small farmers are not wasting their labor, but are substituting their own labor for capital. If this is true, then the lower shadow wage for small farmers leads to input substitution, but entails a much smaller efficiency loss. For example, instead of hiring a tractor, a small farmer can accomplish the same task with so-many days of his own labor. If this is the case, then we expect to observe small farmers choosing a more labor-intensive method of farming, but producing as much output per acre. Furthermore, this choice of more labor intensive production can be related to the development of factor markets (as suggested here), or it can simply arise from technical features of farm production (some type of non-homogeneity in the production function).

A simple way to investigate this possibility is to examine directly the capital intensity of large versus small farms, essentially testing for a mirror image of the inverse relationship for capital input, seeing whether the labor-capital ratio varies negatively with farm size. Unfortunately, this is no easy task. The main difficulty is constructing a measure of capital input, analogous to “days of labor,” or “Yuan of output.” There are different kinds of capital (draft animals versus tractors), and intensity of capital use is difficult to infer from the stock of capital. Furthermore, some capital services are provided by the village, not the farmer. We are able instead to construct a vector of indicators of capital intensity. We have measures of the current value of farm machinery owned by the farmer, plus the value of hired capital services, and hired draft animal services. We also have indicators of whether the village provided capital services for irrigation, cultivation, and harvesting. Combining these measures into a single aggregate is impossible.

We pursue a different strategy. Loosely speaking, we want to see whether *controlling for capital use*, there is an inverse relationship in labor. Ideally, we want to compare two farmers, one big and the other small, and compare their labor input per acre, given they are using the same capital input per acre. If this were possible, then if we still found an inverse relationship for labor, we can conclude that labor was being over-applied by the smaller farmer. So, the simple exercise is to add controls for capital use, and see whether the inverse relationship disappears, or is at least attenuated.

Unfortunately, capital use is potentially endogenous to the labor demand equation. Since there are no obvious instruments for capital use, we must instead consider the consequences of estimating a potentially biased equation, treating the capital variables as exogenous. Our concern is less on the bias on the capital variables themselves, as opposed to the effect on the land coefficient. One can show analytically that the bias on the land coefficient is likely to be positive, that is, including the capital variables (when they are endogenous and correlated with land) is likely to make the land coefficient larger.

The intuition is quite straightforward. The capital substitution hypothesis suggests that smaller farmers will use less capital and more labor per acre (to produce the same output per acre). Controlling for the amount of capital used effectively controls for this substitution. If this substitution captures the entire effect of farm size on the choice of labor intensity, then the coefficient on land will be one, and the inverse relationship will vanish. Even if the controls are imperfect (which they most certainly are), then adding capital to the equation should increase the coefficient on land, attenuating the inverse relationship.

The results are reported in Table 9. Adding capital to the output equation has only marginal effects on the land coefficient, though it increases slightly to greater than one. Note, however that the nine capital variables are only marginally significant in the output equation. The absence of an effect on the inverse relationship should not come as a great surprise, since the addition of village fixed effects yields much the same as this specification (with controls for land quality). Village fixed effects will control for the village capital-provision variables, as well as the crop mix of the village (which may be correlated with capital intensity), and the village price of capital.

In the labor equations, we also see almost no effect of capital on the land coefficient. In the first specification, the land coefficient is 0.515 versus 0.591 in the corresponding column of Table 4. Thus, adding capital reduces the land coefficient, instead of increasing it as predicted by the substitution hypothesis. Even with the possible upward bias caused by the endogeneity of capital, we do not see a reduction of the inverse relationship. Certainly, there is no evidence that controlling for capital "explains away" the inverse relationship. In the next column, we also continue to observe the same relationships between land reallocation and off-farm opportunities and the inverse relationship. Finally, the labor productivity equations also show similar findings as before. While the capital variables are themselves significantly related to productivity, they do not affect the inverse relationship in any way. If anything, it appears that capital and labor are complements (not substitutes), and that farms that use more labor also use more capital. In summary, we find no evidence that the wasted labor interpretation can be discarded in favor of an interpretation that smaller farmers are optimally substituting their labor for capital.

VI. Conclusions

The persistence of administrative forms of land allocation in the face of rapid marketization is a central feature of China's rural economy during transition. Our purpose in this paper has been to examine the effect that this non-market form of allocation and labor market development have on efficiency in the agricultural sector.

Drawing on the inverse relationship, we find no inverse relationship for output, but a significant one for labor input and labor productivity. These findings can only be reconciled by a view of labor inefficiency in the rural sector, with constrained households working more days per acre, but earning virtually nothing in return for the extra time farming. However, certain kinds of village-wide land reallocations and off-farm opportunities help attenuate the severity of the inefficiency. First, more "widespread" reallocations improve labor efficiency. These reallocations help mitigate differences in productivity by farm size, partially replicating the outcome that would occur if there was decentralized land rental. Second, it appears that the labor market, and other sources of off-farm employment are significant determinants of farm labor efficiency. Since low productivity of labor is a consequence of wasted labor, it appears that less labor is wasted where its opportunity costs is greatest.

Despite the positive role of these factors, our findings suggest that there remain significant efficiency gains to be realized by reallocating land from bigger to smaller households. This reallocation will tend to equalize the marginal product of labor across farms, and enable the same output to be produced with less labor. From a policy perspective, then, there is a need for the development of institutions that promote an efficient reallocation of land among households. It is highly unlikely that growth in off-farm opportunities by itself will be able to eliminate the inefficiency. (Parameter estimates suggest that off-farm opportunities will have to double from their current levels.) Moreover, we expect that as these labor market opportunities continue to develop, opportunities for specialization will emerge, and the need for reallocation of land will only increase.

The critical question is: Why don't we see more land reallocation carried out either administratively, or more likely, through a system of well-defined rental-rights and decentralized exchange? Possibly, land reallocations are limited by information problems concerning farmer productivity, administrative costs, or the potentially adverse effects that reallocations may have on investment (dynamic inefficiency). If so, what impedes the development of land rental markets? Is it a lack of a well-defined property right to rent land? Contract enforcement difficulties? Or possibly imperfections in other markets that increase the household's return to farming the land themselves. One thing that we can probably rule out is equity considerations. Our findings suggest that since it is smaller farmers who are most constrained, land reallocation from big to small farms, regardless of the method, will improve both efficiency and overall equity in these villages.[19] For future policy formulation, it is critical that we know the source of these underlying constraints on land rental and the potential interactions between the administrative reallocation and decentralized exchange.

[19] See Benjamin, Brandt, Glewwe, and Guo (2000) for a discussion of income inequality issues exploiting these same data.

VII. References

Benjamin, Dwayne. (1992). "Household composition, labor markets, and labor demand: testing for separation in agricultural household models." *Econometrica* , Vol. 60, (March) pp. 287-322.

Benjamin, Dwayne (1995). "Can Unobserved land quality explain the inverse productivity relationship." *Journal of Development Economics* , Vol. 46, no. 1 (February) pp. 51-84.

Benjamin, Dwayne and Loren Brandt (1997). "Land, Factor Markets, and Inequality in Rural China: Historical Evidence." *Explorations in Economic History* , Vol. 34, no. 4 (October) pp. 460-494.

Benjamin, Dwayne and Loren Brandt, Paul Glewwe, and Li Guo (2000): “Markets, Human Capital, and Inequality: Evidence from Rural China,” Manuscript, University of Toronto.

Berry, Albert and William Cline (1979). *Agrarian Structure and Productivity in Developing Countries*. (Baltimore, Maryland: Johns Hopkins Press.)

Besley, Tim (1995). "Property Rights and Incentives: Theory and Evidence from Ghana." *Journal of Political Economy* , Vol. 103, no. 5 (October) pp. 903-37.

Bhalla, Surjit (1988). "Does land Quality Matter? Theory and Measurement." *Journal of Development Economics* , Vol. 29, no. 1 (July) pp. 45-62.

Binswanger, Hans and Klaus Deinenger (1997). "Explaining Agricultural and Agrarian Policies in Developing Countries." *Journal of Economic Literature* , Vol. 35, no. 4 (December) pp. 1958-2005.

Binswanger, Hans, Klaus Deinenger and Gershon Feder (1995). "Power, Distortions, Revolt and Reform in Agricultural Land Relations." *Handbook of Development Economics*. J. Behrman and T. N. Srinivasan, Eds. (Amsterdam: Elsevier Science, North Holland.)

Burgess, Robin (1997). "Land, Welfare and Efficiency in Rural China." Manuscript, London School of Economics.

Carter, Michael (1984). "Identification of the Inverse Relationship between Farm Size and Productivity: An Empirical Analysis of Peasant Agricultural Production." *Oxford Economic Papers* , Vol. 36, no. 1 (March) pp. 131-146.

Carter, Michael, Shouying Liu, Michael Roth and Yang Yao (1996). "An Induced Institutional Innovation Perspective on the Evolution of Property Rights in Post-Reform China." Manuscript, University of Wisconsin.

Feder, Gershon and Klaus Deininger (1998). "The Evolution of the World Bank's Land Policy." Manuscript, The World Bank.

Lin, Justin (1992): “Rural Reforms and Agricultural Growth in China.” *American Economic Review,* Vol. 82, No. 1 (March), pp. 34-51.

Turner, Matt, Loren Brandt and Scott Rozelle (1999). "Property Rights Formation and the Organization of Exchange and Production in Rural China." Manuscript, University of Toronto.

Udry, Chris (1999). "Efficiency and Market Structure: Testing for Profit Maximization in African Agriculture." Chapter 16 in *Trade, Growth, and Development*, G. Ranis and L.K. Raut (eds.), Elsevier Science.

Udry, Chris (1996). "Gender, Agricultural Production, and the Theory of the Household." *Journal of Political Economy* , Vol. 104, no. 5 (October) pp. 1010-46.

TABLE 1
SUMMARY STATISTICS
(n=787)

	Mean	Percentage not 0
Farm Output	6505.0	93.2
<u>Labor:</u>		
Male days to farm	67.8	89.2
Female days to farm	62.5	87.3
Male days to sideline activities	16.9	59.1
Female days to sideline activities	18.1	67.6
Male days to wage labor	105.4	42.3
Female days to wage labor	39.4	16.3
Male days to family enterprise	73.3	40.7
Female days to family enterprise	36.9	20.2
TOTAL MALE DAYS	263.5	98.6
TOTAL FEMALE DAYS	156.9	94.5
Hired labor (on Farm)	4.4	14.6
Exchange Labor (on farm)	6.4	43.6
<u>Land</u>		
Cultivated land (mu)	10.5	93.3
Land Rented-in (mu)	0.45	10.8
Land Rented-out (mu)	0.21	6.0
Increases in allocated land (past 3 years, in mu)	1.81	22.5
Decreases in allocated land (past 3 years, in mu)	1.21	30.1

TABLE 2
DETERMINANTS OF ADMINISTRATIVE LAND ALLOCATION
(standard errors in parentheses)
Dependent variable: Log Household Land Allocation

	OLS	**OLS-FE**	**OLS-LQ**	**OLS-C**
Ln family size	.597 (.110)	.906 (.084)	.824 (.098)	.890 (.151)
Male children	-.723 (.386)	-.651 (.311)	-.436 (.360)	-.563 (.387)
Female children	-.784 (.396)	-.715 (.309)	-.524 (.363)	-.588 (.394)
Male teens	.138 (.379)	-.075 (.299)	.153 (.345)	.163 (.378)
Female teens	.145 (.395)	-.183 (.302)	.057 (.378)	.045 (.407)
Male prime age	-.150 (.385)	-.441 (.309)	-.331 (.347)	-.155 (.470)
Female prime age	.258 (.343)	.282 (.245)	.231 (.250)	.378 (.300)
Male elderly	-.525 (.523)	-.711 (.427)	-.598 (.426)	-.650 (.565)
Re_all				.046 (2.006)
Re_all × ln family size				.815 (1.415)
Off-farm employment				-.166 (.273)
Off-farm × ln family size				-.368 (.213)
Village Effects	No	Yes	No	No
Land Quality	No	No	Yes	Yes
R-Squared	.08	.50	.33	.37

Notes: 1) Re_all is the product of the fraction of village land reallocated within the past 3 years and the proportion of households receiving positive increases in allocated land. 2) Off-farm employment is a household level variable indicating the percentage of household members working off the farm. 3) Sample size is 728. 4) The demographic variables are the ln of households size, as well as proportion of household members in each demographic group. The omitted category is elderly women. 5) The column headings refer to the estimator: OLS is ordinary least squares; OLS-FE adds village fixed effects; OLS-LQ adds village and farm-level measures of land quality; OLS-C corrects standard errors for village clustering.

TABLE 3
LOG FARM OUTPUT REGRESSIONS
(standard errors in parentheses)

	OLS	**OLS-FE**	**OLS**	**OLS-C**	**2SLS**
Ln Land (ln h)	.833 (.035)	.971 (.022)	1.071 (.037)	1.028 (.050)	1.149 (.103)
Ln Family Size	.137 (.076)	-.003 (.054)	-.183 (.057)	-.204 (.059)	-.194 (.068)
Re_all				-.249 (.487)	.019 (.557)
Re_all × Ln Land				-.093 (.138)	-.142 (.164)
Off-farm employment				-.404 (.176)	.620 (.597)
Off-farm employment × Ln Land				.163 (.073)	-.035 (.178)
Cluster FE	No	Yes	No	No	No
Land quality	No	No	Yes	Yes	Yes
R-Squared	.63	.87	.82	.82	.80

Notes: 1) Re_all is the product of the fraction of village land reallocated within the past 3 years and the proportion of households receiving positive increases in allocated land. 2) Off-farm employment is a household level variable indicating the percentage of household members working off the farm. 3) The column headings refer to the estimator: OLS is ordinary least squares; OLS-FE adds village fixed effects; OLS-LQ adds village measures of land quality; OLS-C corrects standard errors for village clustering; 2SLS is IV estimation, with off-farm employment (and its interaction with farm size) treated as an endogenous variable. The excluded instruments are: village measures of household labor market participation, age and education of the household head, as well as the fraction literate. These variables are also interacted with log farm size. 4) The overid test is 12.13 (P^2(6) and the F-statistic (df=8,699) for the excluded instruments in the first stage regressions is 11.98 and 7.78 respectively for off-farm employment and its interaction with ln h. 5) Land Quality indicates controls for village and farm level land quality.

TABLE 4
LOG LABOR INPUT REGRESSIONS
(standard errors in parentheses)

	OLS	**OLS-FE**	**OLS**	**OLS-C**	**2SLS**
Ln Land (ln h)	.495 (.034)	.558 (.033)	.591 (.039)	.424 (.068)	.170 (.137)
Ln Family Size	.470 (.095)	.322 (.079)	.343 (.080)	.351 (.081)	.329 (.110)
Re_all				-.366 (.707)	-.916 (.690)
Re_all × Ln Land				.321 (.157)	.433 (.164)
Off-farm employment				-.810 (.214)	-2.895 (.977)
Off-farm employment × Ln Land				.282 (.093)	.717 (.266)
Cluster FE	No	Yes	No	No	No
Land quality	No	No	Yes	Yes	Yes
R-Squared	.33	.60	.53	.55	.42

Notes: 1) Re_all is the product of the fraction of village land reallocated within the past 3 years and the proportion of households receiving positive increases in allocated land. 2) Off-farm employment is a household level variable indicating the percentage of household members working off the farm. 3) The column headings refer to the estimator: OLS is ordinary least squares; OLS-FE adds village fixed effects; OLS-LQ adds village measures of land quality; OLS-C corrects standard errors for village clustering; 2SLS is IV estimation, with off-farm employment (and its interaction with farm size) treated as an endogenous variable. The excluded instruments are: village measures of household labor market participation, age and education of the household head, as well as the fraction literate. These variables are also interacted with log farm size. 4) The overid test is 4.89 ($P^2(6)$ and the F-statistic (df=8,699) for the excluded instruments in the first stage regressions is 11.98 and 7.78 respectively for off-farm employment and its interaction with ln h. 5) Land Quality indicates controls for village and farm level land quality.

TABLE 5
LOG PRODUCTIVITY REGRESSIONS
(standard errors in parentheses)

	OLS	**OLS-FE**	**OLS**	**OLS-C**	**2SLS**
Ln Land (ln h)	.339 (.048)	.413 (.039)	.481 (.050)	.604 (.077)	.979 (.187)
Ln Family Size	-.334 (.122)	-.325 (.094)	-.526 (.097)	-.555 (.101)	-.522 (.148)
Re_all				.117 (.926)	.935 (1.002)
Re_all × Ln Land				-.414 (.217)	-.575 (.245)
Off-farm employment				.406 (.281)	3.515 (1.277)
Off-farm employment × Ln Land				-.119 (.131)	-.751 (.370)
Cluster FE	No	Yes	No	No	No
Land quality	No	No	Yes	Yes	Yes
R-Squared	.33	.58	.53	.55	.42

Notes: 1) Re_all is the product of the fraction of village land reallocated within the past 3 years and the proportion of households receiving positive increases in allocated land. 2) Off-farm employment is a household level variable indicating the percentage of household members working off the farm. 3) The column headings refer to the estimator: OLS is ordinary least squares; OLS-FE adds village fixed effects; OLS-LQ adds village measures of land quality; OLS-C corrects standard errors for village clustering; 2SLS is IV estimation, with off-farm employment (and its interaction with farm size) treated as an endogenous variable. The excluded instruments are: village measures of household labor market participation, age and education of the household head, as well as the fraction literate. These variables are also interacted with log farm size. 4) The overid test is 6.14 ($P^2(6)$ and the F-statistic (df=8,699) for the excluded instruments in the first stage regressions is 11.98 and 7.78 respectively for off-farm employment and its interaction with ln h. 5) Land Quality indicates controls for village and farm level land quality.

TABLE 6
IMPLIED DAYS OF WASTED LABOR

Farm Size	Mean ln h	Mean ln L "Efficient" labor	Mean ln L "Actual" labor	Implied Wasted ln L (days)
Large	3.505 (33 mu)	5.203 (182 days)	5.203 (182 days)	0
Medium	2.512 (12 mu)	4.210 (67 days)	4.607 (100 days)	0.397 (33 days)
Small	1.393 (4 mu)	3.091 (22 days)	3.936 (51 days)	0.845 (29 days)

Notes: (1) Large farms are defined as those farms above the 90th percentile in size (i.e., ln h $ 3.135); Medium farms are those above the median size (but below the 90th percentile; i.e., 2.079 # ln h < 3.135); and small farms are below the median size (i.e., ln h < 2.079). (2) "Efficient" labor is estimated as the predicted labor input, assuming no inverse relationship. (3) "Actual" labor is the implied days of labor, given the degree of the inverse relationship estimated in Table 4 (column 3)

TABLE 7
ALTERNATIVE MEASURES OF LABOR
(standard errors in parentheses)
All Specifications Estimated by OLS

	Ln Household Labor Supply	**Ln Labor Applied to Farm in Busy Season**		**Ln Labor Applied to Farm in Slack Season**	
Ln Land (ln h)	-.038 (-.050)	.631 (.034)	.471 (.061)	.504 (.047)	.374 (.083)
Ln Family Size	1.032 (0.100)	.349 (.092)	.360 (.093)	.545 (.126)	.521 (.126)
Re_all			-1.775 (.595)		1.508 (.806)
Re_all × Ln Land			.563 (.223)		-.315 (.303)
Off-farm employment			-.668 (.237)		-1.105 (.321)
Off-farm employment × Ln Land			.211 (.010)		.367 (.135)
Cluster FE	No	No	No	No	No
Land quality	Yes	Yes	Yes	Yes	Yes
R-Squared	.30	.47	.49	.39	.41

Notes: 1) Re_all is the product of the fraction of village land reallocated within the past 3 years and the proportion of households receiving positive increases in allocated land. 2) Off-farm employment is a household level variable indicating the percentage of household members working off the farm. 3) All specifications correct standard errors for village clustering. 4) Land Quality indicates controls for village and farm level land quality. 5) Cluster FE refers to village fixed effects. 6) All specifications also include controls for the age-gender mix of the household, as presented in Table 1. 7) Household Labor Supply is the total number of days worked in ALL activities by household members; Labor Applied in the Busy Season is the number of days of family labor supplied to the farm in the self-reported busy season, while labor applied in the slack season is the analogous measure for the self-reported slack season.

TABLE 8
LAND REALLOCATION TREATED AS ENDOGENOUS
2SLS Estimates
(standard errors in parentheses)

	Extended Instrument Set			**Limited Instrument Set**		
	1st Stage F-stat (36 Excl.)	**ln Output**	**ln Labor**	**1st Stage F-Stat (12 Excl.)**	**ln Output**	**ln Labor**
Ln Land (ln h)		1.087	.154		1.176	.225
		(.064)	(.108)		(.088)	(.123)
Ln Family Size		-.177	.328		-.188	.310
		(.061)	(.102)		(.072)	(.100)
Re_all	7.77	.757	-.885	3.18	1.151	-.292
	(0.00)	(.899)	(.814)	(0.00)	(1.313)	(1.099)
Re_all × Ln Land	15.56	-.098	.471	2.15	-.303	.124
	(0.00)	(.177)	(.212)	(0.05)	(.401)	(.445)
Off-farm employment	46.20	.029	-2.783	9.95	.773	-2.699
	(0.00)	(.377)	(.711)	(0.00)	(.522)	(.885)
Off-farm employment ×	37.07	-.048	.806	7.17	-.069	.726
Ln Land	(0.00)	(.118)	(.218)	(0.00)	(.150)	(.246)
Over-id test		100.88	36.6		13.16	19.8
R-Squared		.82	.46		.79	.45

Notes: 1) Re_all is the product of the fraction of village land reallocated within the past 3 years and the proportion of households receiving positive increases in allocated land. 2) Off-farm employment is a household level variable indicating the percentage of household members working off the farm. 3) 2SLS is IV estimation, with off-farm employment (and its interaction with farm size) and Re_all (and its interaction with farm size) treated as endogenous variables. The excluded instruments are: (i) The same as in previous tables: village measures of household labor market participation, age and education of the household head, as well as the fraction literate; and (ii) Additional village-level political economy variables designed to help predict the nature of land reallocation. (a) For the extended set: number of plots, an indicator of whether the township reallocates, the number of households, the number of production teams, the number of cadres, whether the last election was contested, and the age, education, tenure, and farming-status of the village head and party secretary. (b) For the limited instrument set, only the township reallocates and number of households are used. These variables are also all interacted with log farm size. 4) The overidentification (overid) test is distributed as P^2(32) and P^2(8) in the extended and limited specifications. The 5% critical values would be 45 and 16, respectively, and at the 1% level, 55 and 20. The first stage F-statistic tests for the joint significance of the excluded instruments in the first stage regressions. 5) All specification include the demographic variables, as well as the Land Quality measures.

TABLE 9
CONTROLS FOR FARM CAPITAL USE
(standard errors in parentheses)

	Ln Output		Ln Labor		Ln Productivity	
	OLS	**2SLS**	**OLS**	**2SLS**	**OLS**	**2SLS**
Ln Land (ln h)	1.085	1.121	.515	.119	.570	1.002
	(.039)	(.091)	(.043)	(.141)	(.044)	(.185)
Ln Family Size	-.194	-.192	.355	.426	-.549	-.517
	(.045)	(.049)	(.076)	(.092)	(.088)	(.124)
Re_all		-.365		-1.358		.200
		(.559)		(.716)		(1.263)
Re_all × Ln Land		-.045		.424		-.448
		(.139)		(.229)		(.280)
Off-farm employment		.357		-2.351		3.123
		(.495)		(.848)		(1.202)
Off-farm employment × Ln Land		-.018		.500		-.714
		(.156)		(.245)		(.343)
F-Capital	2.02	2.48	3.61	10.42	8.44	2.78
	(0.07)	(.03)	(.00)	(.00)	(.00)	(.02)
Over-ID		16.52		8.11		7.97
R-Squared	.83	.82	.55	.35	.49	.35

Notes: 1) The capital variables are a) indicators for the use, and the level of i) The current value of household farm capital; ii) the hiring of capital services; and iii) the hiring of draft animal services, and b) Indicators of whether the village provides capital services (separately for harvesting, irrigation, and cultivation. 2) Re_all is the product of the fraction of village land reallocated within the past 3 years and the proportion of households receiving positive increases in allocated land. 3) Off-farm employment is a household level variable indicating the percentage of household members working off the farm. 4) 2SLS is IV estimation, with off-farm employment (and its interaction with farm size) and Re_all (and its interaction with farm size) treated as endogenous variables. The excluded instruments are the same as in previous tables: village measures of household labor market participation, age and education of the household head, as well as the fraction literate. 5) All specifications include the demographic variables, as well as the Land Quality measures.

FIGURE 1 — VILLAGE LAND REALLOCATIONS

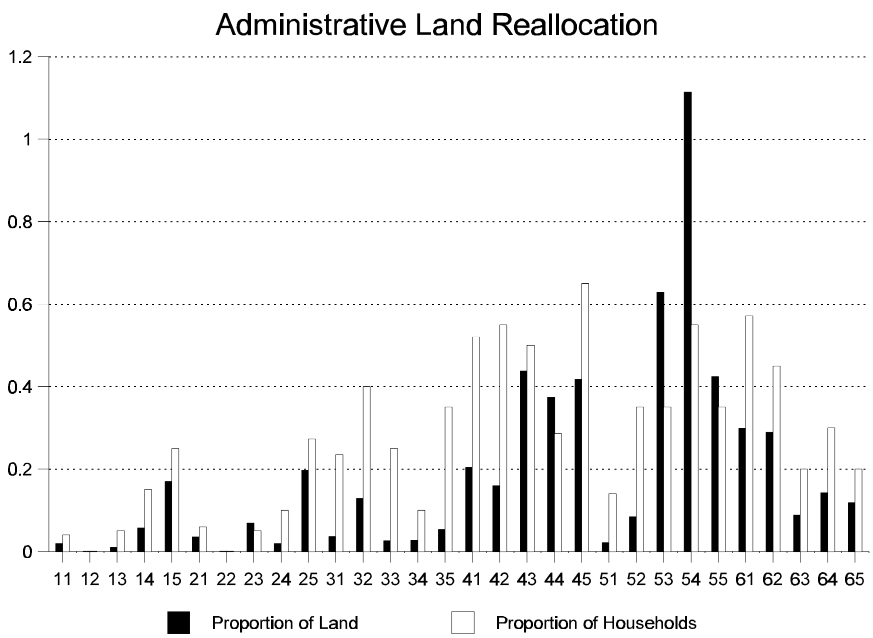


FIGURE 2 — VILLAGE LAND REALLOCATIONS: "RE_ALL"

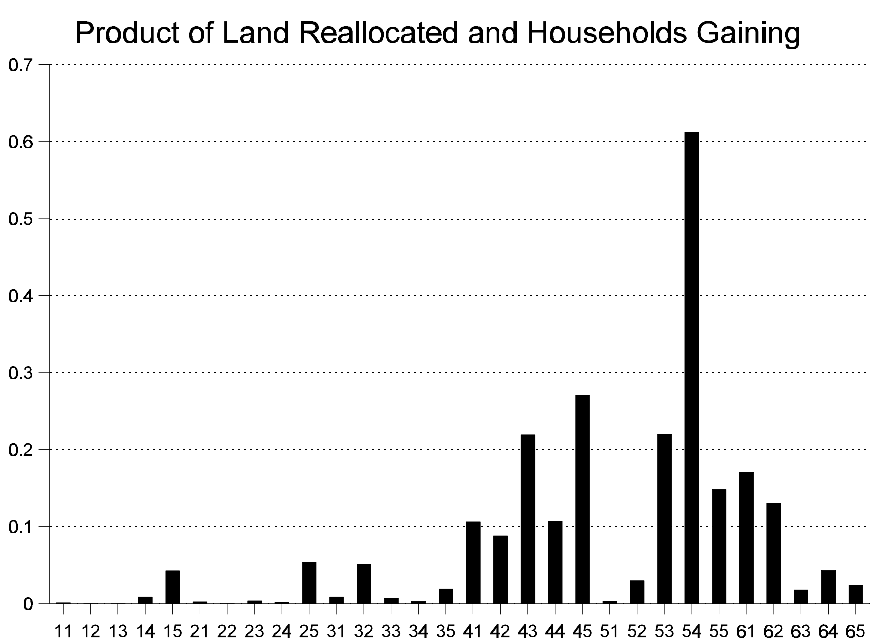


FIGURE 3 — INTENSITY OF HOUSEHOLD ENGAGEMENT IN OFF-FARM EMPLOYMENT

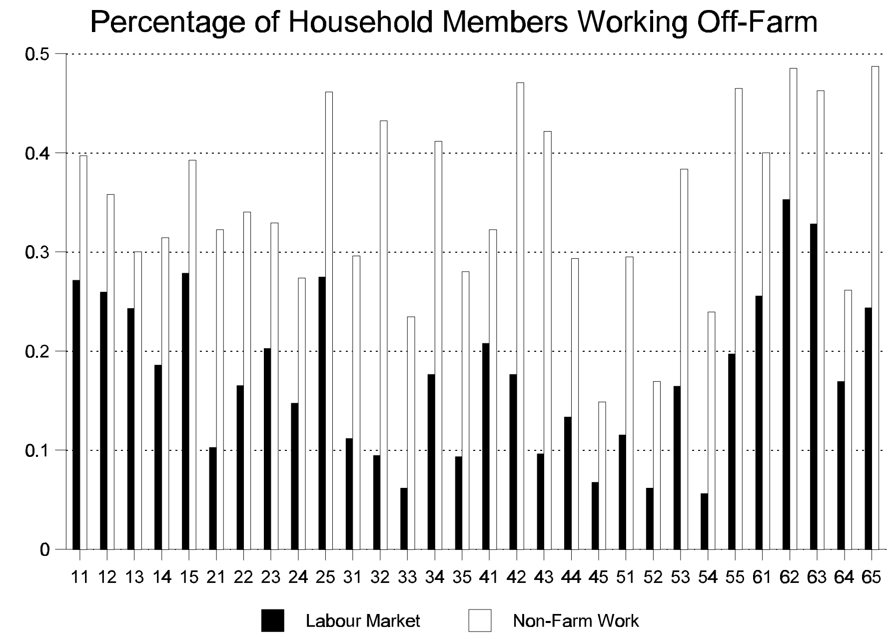


FIGURE 4 — PROPORTION OF FARM LABOR HIRED-IN BY VILLAGE

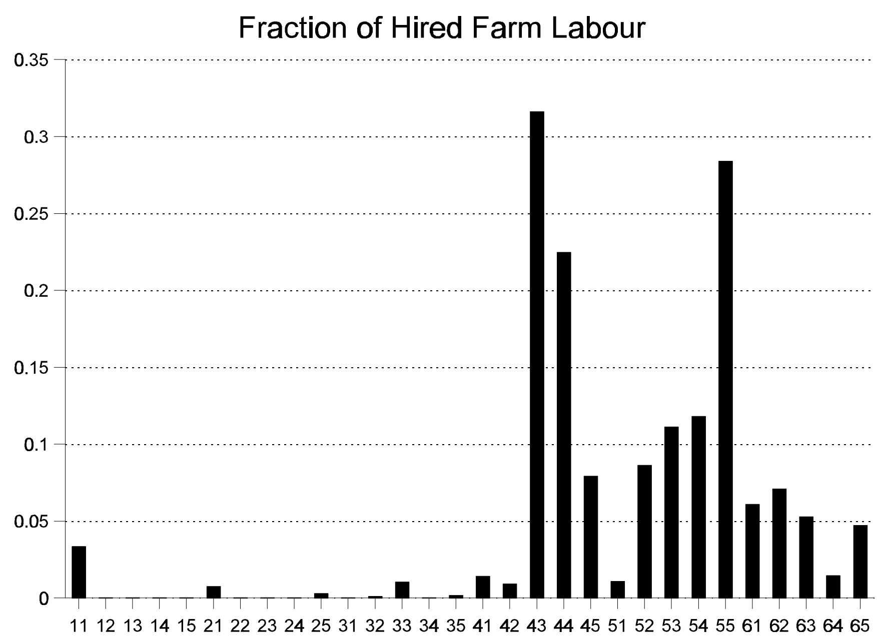
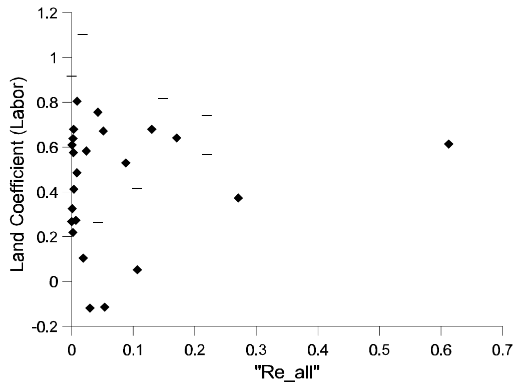
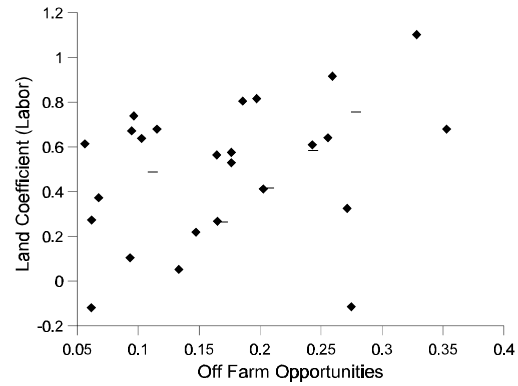


FIGURE 5 – THE INVERSE RELATIONSHIP ACROSS VILLAGES AND ITS CORRELATION WITH LOCAL "INSTITUTIONS"





Notes: The inverse relationship land coefficients are estimated from a regression identical to the specification in column two of Table 4 (with village fixed effects), but with interaction effects between village and (log) farm size. These interaction effects yield the village specific relationship between farm size and labor intensity. These coefficients are then plotted against the institutional variables ("Re_all" and off-farm opportunities) as described previously.