10-1-2010

Is There Surplus Labor in Rural India?

Andrew D. Foster

Mark R. Rosenzweig

ECONOMIC GROWTH CENTER
YALE UNIVERSITY
P.O. Box 208629
New Haven, CT 06520-8269
http://www.econ.yale.edu/~egcenter/

CENTER DISCUSSION PAPER NO. 991

Is There Surplus Labor in Rural India?

Andrew D. Foster
Brown University

Mark R. Rosenzweig
Yale University

October 2010

Is There Surplus Labor in Rural India?

Andrew D. Foster and Mark R. Rosenzweig

Abstract

We show empirically using panel data at the plot and farm level and based on a model incorporating supervision costs, risk, credit-market imperfections and scale-economies associated with mechanization that small-scale farming is inefficient in India. Larger farms are more profitable per acre, more mechanized, less constrained in input use after bad shocks, and employ less per-acre labor than small farms. Based on our structural estimates of the effects of farm size on labor use and the distribution of Indian landholdings, we estimate that over 20% of the Indian agricultural labor force is surplus if minimum farm scale is 20 acres.

JEL codes: O13, O16, O53

Keywords: Agriculture, India, scale, profits, labor, tractors

Although the generalization has many important caveats, across the world the most efficient and productive agriculture is situated in countries in which farms are family-owned, large-scale and mechanized. However, comparisons of farming productivity across countries of the world cannot easily identify the essential barriers to augmenting farming productivity, as countries differ in their property rights regimes, financial systems, labor markets, agroclimatic conditions and other institutional and environmental features. A vast literature has highlighted, usually one at a time, various market imperfections as constraining agricultural productivity in poor countries. These include, for example, credit market barriers, lack of insurance, problems of worker effort, and labor market transaction costs. However, many of these market problems are not confined to poor countries. Moral hazard and adverse selection afflict credit markets in all settings, and farmers do not have unlimited access to capital anywhere in the world. Nor do family farms in many developed countries use employment schemes that differ importantly from those used in those low -income settings where family farms also dominate. And most farmers in high-income countries do not participate in formal crop, income or weather insurance markets. It is thus unlikely that labor market problems or lack of insurance or even credit constraints can alone account for the large differences in the efficiency of farms across many developed and developing countries.

In contrast to agriculture in most developed countries where farming is very efficient, farming in India, while family-run, is neither large -scale nor mechanized. The 2001 Census of India indicates that farming in India is *very* small scale - 80% of farms are less than two acres in size and 95% are less than five acres in terms of owned holdings. Mechanization can be examined using data from a new panel survey of almost 5,000 crop-producing farmers in 17 of the major states of India in 2007-8, which we describe and employ extensively below. Figure 1 displays the fraction of farms in the survey data with a tractor, a mechanized plow or a thresher by land ownership size. As can be seen, less than five percent of farms below two acres own any of these mechanized implements, but mechanization increases significantly with ownership holdings, with 30% of farms above 10 acres owning a tractor and over 20% owning a mechanized thresher.

Are farms in India too small and under-mechanized? Our survey data also appear to indicate that small farms in India are substantially less efficient than larger farms. We use as our measure of efficiency profits per acre, which reflects the resource costs of farming, inclusive of the value of family labor (valued at their opportunity costs), supervisory labor, and own equipment use. By this measure, which does not take into account the likelihood that larger farmers may also have lower credit costs, landownership and farm productivity are strongly positively associated. Figure 2 provides a lowess-smoothed plot of per-acre profits and landownership from the survey data, net and gross of labor supervision costs. As can be seen, up to about12 acres, per-acre farm profits increase with land ownership size. The difference in the two profit measures is labor supervision costs. The plots thus indicate not only that per-acre profits rise but that per-acre supervision costs fall with owned acreage. These patterns appear to go against the conventional idea that small farms, which employ mostly family labor, have a cost advantage over larger farms who employ a higher fraction of hired labor. This presumption overlooks how mechanization, which evidently rises with farm size, reduces overall labor use.[1] And, Figure 2 also shows that total labor costs per acre decline with landholdings.

There is a large prior literature using Indian data from the 1970's and 1980's that has found both more intensive use of labor on smaller farms and a negative relationship between output per acre and cultivated area. But when profits are calculated, valuing family labor at prevailing wage rates, small Indian farms are found to be less efficient than larger farms (Carter, 1984; Lamb, 2003), consistent with the descriptive associations in Figure 2. One hypothesis explaining these findings has been that the shadow value of labor for small farms is less than the wage (deJanvry *et al.*, 1991). This idea is consistent with the original models of surplus labor (Lewis, 1954; Fei and Ranis, 1964) in which the true marginal product of labor is below observed wage rates. Indeed, Sen's formal model of Lewis' idea that removing labor from small farms would not reduce output, with family members increasing their work effort, is based on the assumption of the absence of labor market opportunities for family labor off the farm (Sen, 1966).

However, more recent evidence on rural employment and wage growth would appear to contradict the idea that rural labor in India is now in surplus, with subsistence needs binding wage rates and small-farm households unable to sell their labor. Plot-level data from India, again

[1]This also presupposes that family members do not require supervision to work efficiently. We show that both hired and family labor require the same amount of supervision time.

from the 1970's and 1980's, has led to the conclusion that differences in input costs across farms cannot account for any of the scale differences in productivity measures in India (Assunção and Braido, 2007; Lamb, 2003). Lanjouw and Murgai (2009), morever, documents the growth of Indian nonfarm rural employment since the 1970's, which now makes up half of the rural labor force in India. And our recent Indian survey data indicate that a least one family member works off the farm in 75% of farm households below half an acre, for 100 days on average per year. Finally, real Indian agricultural wage rates have risen substantially over the past 30 years (Bhalla and Das, 2005).

In this paper we show empirically, based on a simple model incorporating standard labor markets but with labor supervision requirements, risk, credit market imperfections and scale economies associated with mechanization that a reduction in the agricultural labor force can occur without a decline in farm output or an increase in rural wages if the assumption of the existing distribution of landholdings is relaxed - the surplus labor model prediction is conditionally correct for India. The source of labor surplus is not wage floors or labor-market imperfections, however, but the inefficiency of small-scale agriculture.

Of course, Figure 2 merely describes associations between scale, labor use and profitability. It is possible that within India larger farms are located where land is higher quality (Bhalla and Roy, 1988; Benjamin, 1995), where credit markets operate more effectively, or where agricultural conditions generally are more favorable to agriculture. Measurement error in farm or plot size also can bias estimates of per-area efficiency and scale (Barrett *et al.*, 2010; Lamb, 2003). Moreover, land holdings at the farm level are endogenous, and may reflect differences in property rights regimes or the capability of farmers. Many prior studies of scale effects and the role of market constraints on farm productivity have attempted to correct for particular dimensions of heterogeneity. A major shortcoming of the literature, however, is that there have not been credible methods of dealing with the endogeneity of machinery, input use and land ownership and most have examined only particular market constraints and not the interactions between them in terms of their effects on farm productivity.[2] In the absence of a feasible way of experimentally varying ownership holdings or farm scale, key to the empirical

[2]An exception is Feder (1985) who models how the interaction between the need for supervision and credit constraints affects the relationship between operational scale and landholdings. In that model, however, family labor is assumed to not require supervision and mechanization is ignored.

identification of scale and credit market effects on profitability and mechanization are the ability to control for unobserved differences across farmers in ability, preferences and in costs (e.g., interest costs and shadow labor costs) as well as differences in land quality.

We have collected panel data at the plot (across seasons in the same crop year) and at the farm level (over the span 1999-2008) on inputs, outputs and investments. Variation across plots for the same farmer can help identify pure scale effects, net of measured plot characteristics and measurement error, because such an analysis controls for all farm-specific costs. Variation over-time in the effects of lagged farm profits on contemporaneous profits for the same plot, by farm size, identify the role of ownership holdings on the ability to attain overall efficiency in input use in the face of uninsurable risk. To obtain causal effects of landownership on profitability and investments, we exploit the fact in the nine-year period 1999-2008, almost 20% of farm households split and/or received inherited land because a parent died. We follow an individual farmer before and after inheriting land and/or assets and use the inherited assets as instruments to explain the change in landownership and capital equipment in an instrumental-variables set-up.

Our estimates support the existence of scale economies: for a given farmer, per-acre profits and the use of capital equipment are higher on larger plots compared with smaller plots, while per-acre use of labor on larger plots is lower. A farmer who experiences an exogenous increase in owned landholdings exhibits an increase in profits per acre and is more likely to invest in capital equipment in villages where a bank is proximate. Moreover, profits per acre are higher on a given plot if a given farmer experiences a favorable farm-level profit shock in the prior period only for farmers with smaller overall landholdings. These latter results indicate that landownership helps overcome credit constraints on both investment and variable input use (Rosenzweig and Binswanger, 1993). Consistent with this and with the higher returns to land among larger landowners, we find that the marginal returns to capital and to fertilizer decline with owned landholdings. Finally, we show that in our data, consistent with the higher profitability of a larger scale of operation and with the relaxed credit needs associated with greater owned landholdings, farmers with small landholdings lease out to farmers with larger landholdings within a village. This reverse tenancy does not overcome the adverse ownership distribution of land, as only nine percent of farmers lease in land.

Our results indicate that scale and lack of mechanization is a barrier to greater farm productivity in India, and that as a consequence of credit market constraints and scale economies, most farms in India are too small to exploit the productivity and cost-savings from mechanization. The flip side of these findings is that there are too many farms and too many people engaged in agriculture. We provide illustrative calculations based on the existing distribution of Indian farms and our estimates of the causal effect of farm size on per-acre labor costs by farm size, that with a hypothetical consolidation that led to a minimum farm size of as little as 20 acres, 22% of the agricultural work force in India is surplus as defined by the classical surplus labor models.[3]

2. Theory

A. *Technical scale economies, cultivated land area and mechanization*

In order to illuminate the role of returns to scale associated with mechanization in a relatively tractable structure we develop a model in which there are constant returns to scale in land and all variable inputs. For ease of exposition, we define the services provided by labor and/or equipment as work to be done. The model is set up in such a way that scale in terms of land size affects the relative productivity of different sources of work but, *given area*, there are constant returns in terms of the amount of work done. In particular, for a farmer with given scale of production measured in acres a let output y per acre in a given crop cycle be

(1) $$y = ag(e/a, f/a)$$

where e denotes work, and f denotes a variable input such as fertilizer. We assume that (i) manual labor and machinery services are imperfect substitutes in producing work, (ii) that manual labor, regardless of the hired or family status of that labor,[4] must be supervised in order to be productive, and (iii) that machinery varies by capacity. These assumptions are embodied in the following function:

(2) $$e = (\omega_m(l_m(l_s/l_m))^{\delta} + \omega_k((\phi(a)-q)qk)^{\delta})^{1/\delta}.$$

where l_s denotes supervisory labor, l_m denotes manual labor, $l_m(l_s/l_m)$ is a constant-returns labor-services production function, q denotes the capacity of each machine, and k denotes the number

[3] In many Indian states the legal maximum landholding size is below this hypothetical floor. The average farm size in the United States in 2009 is 450 acres (US Economic Research Service). Allen (1988) documents that at the end of the 18th century in England, average farm size was 150 acres. He also shows that larger farms used substantially less labor per acre and exhibited higher profits per acre.

[4] We provide empirical evidence consistent with this assumption below.

of machines. Note that the loss associated with using higher-capacity equipment on small plots is embodied in the expression $\varphi(a)$-q with $\varphi'(a)$>0.

We assume that higher-capacity machines are more costly but that machinery cost does not rise linearly with capacity. In particular, the price per day of a machine with capacity q is $c_k q^v$, where v <1 and $v < 2\delta$ for an interior solution. We also assume that there is a perfect rental market for machines.[5] The cost of production is

(3) $\quad ,p_f f + c_k k q^v + w_m l_m + w_s l_s$

where k=the number of machines, p_f=price of fertilizer, w_m=wage rate of manual labor, and w_s=wage of supervisory labor. A profit-maximizing farmer maximizes the value of (1) minus costs (3) subject to (2) .

In solving this problem and to highlight the particular role that land-size plays in this structure it is helpful to consider first the cost function

(4) $\quad \tilde{c}(e,a) = \min\, p_k k q^v + w_m l_m + w_s l_s$ subject to (2)

Solving (4) first in terms of capacity yields an expression for optimal machine capacity

(5) $\quad q = \phi(a)(1-\nu)/(2-\nu).$

Expression (5) indicates that optimal capacity is determined only by area and the elasticity ν of the price schedule and, in particular, is not sensitive to the required total work. Larger operations will use higher-capacity equipment, but a decrease in the elasticity of the machinery price with respect to capacity (v), say due to technical change, increases machinery capacity particularly for large farmers.[6]

The first-order conditions to the cost minimization problem imply that the ratio of supervisory to manual labor is constant given prices and technologies and that the ratio of machinery to labor services is constant given area, prices, and technologies. Because of this proportionality, we can readily distinguish between how scale affects the demand for inputs conditional on the amount of work and on how scale affects total input demand by increasing work. In Appendix A we prove the following:

A. Profits per acre increase with area because the cost of work per unit area falls.

[5]We consider the own-versus buy decision once we introduce a credit market below.

[6]Note that substituting back into the (2) yields a work production function that is analogous to the CES production function with the exception that the share parameter $\omega_k(\kappa_1\phi(a)^2)^\delta$,where $\kappa_1 = (1-\nu)/(2-\nu)^2$,depends on area.

B. Larger operations are more profitable on a per-acre basis.

C. Larger operational holdings will use inputs more intensively. This is because per-acre work increases in unit area.

D. Fertilizer per acre increases if fertilizer and work are complementary.

E. The number of machines *k* per unit area will be increasing in area, for δ sufficiently close to 1. This is because (i) there will be an overall expansion of work and (ii) *k* is increasing in total work.

F. Whether total expenditures on machinery will rise for $\delta < 1$ as land size increases depends on whether the pricing of machinery is sufficiently elastic to capacity, on the magnitude of *v*. Regardless of whether the number of machines used per unit area increases or decreases, whether a farmer owns a machine of a given capacity or greater is rising in area.

Will larger operations use less labor per unit area? The effect of an expansion in area on the amount of manual labor used per acre is ambiguous. There is an increase in work intensity as the increasing returns associated with machinery lower unit work costs, but there is also a decrease in the amount of labor per unit work as we show in Appendix A. If the demand for work is price inelastic and/or labor and machines are sufficiently good substitutes, however, both manual and supervisory labor must decline,

$$(6) \qquad \frac{dl_m^*}{da} = \frac{dl_m^c}{da} e^*(1 + \varepsilon_{ec}(1 - \delta))$$

where l_m* is the per-acre amount of labor, l_m^c is the conditional amount of labor used for given work, and ε_{ec} is the own price elasticity of demand for work The expression for supervisory labor is the same except that the subscript *m* is replaced by *s*.

B. *Scale effects and credit market imperfections*

In the preceding analysis *a* was any contiguous plot of land used for an agricultural operation. We have thus ignored the distinctions between the ownership or rental of land, as well as equipment, and we have also assumed that over the agricultural cycle farmers can freely borrow against harvest revenues at a zero rate of interest. We now allow for the possibility of credit constraints. In doing so, we assume that farmers own their plots of land and also own capital equipment. We first take ownership of both assets as given, and then endogenize the ownership of equipment. To incorporate capital market considerations we assume that farmers borrow $b*$ per acre to finance agricultural inputs and repay this debt with interest during the harvest period. We assume that the interest rate *r* on this debt is dependent on the amount borrowed per acre as well as on total owned land area, with farmers who own a small amount of land *a* obtaining working capital at a higher interest rate than larger farmers. Formally, the per-acre amount that must be repaid in the harvest period is given by

(7) $$\rho(a,b^*) = (1+r(a,b^*))b^*$$

where the interest rate *r* is increasing in b^* and decreasing in owned land and * denotes per-acre values. The decrease in interest rates with land ownership might reflect the use of collateral, a requirement of most bank loans in rural India (Munshi and Rosenzweig, 2009). In this extended model, ownership of both land and machinery matters. By assumption owned landholdings reduce the cost of capital. But, while we retain the assumption that there is a perfect rental market for machinery, ownership (versus rental) of capital assets such as machinery also influences production decisions through its effect on the amount of debt that must be incurred to finance inputs. In short, if one owns a productive asset one does not have to finance the relevant rental cost.[7] Or equivalently one can rent the machine to other farmers and then use the cash to finance other inputs. Thus letting o^* denote the rental value of owned assets

(8) $$b^* = c(a)e^* + p_f f^* - o^* \ .$$

The farmer's maximization problem with credit market imperfections can thus be restated as

(9) $$\pi^*(a) = \max g(e^*, f^*) - \rho(a,b^*) - (1+r_0)o^*$$

where r_0 is the rate of return on savings and is assumed to be less than $r(a,b^*)$ for all positive levels of borrowing.

Profit maximization then implies that

(10) $$d\pi^*/da = -\rho_a - \rho_b c'(a)e^*$$

where $\rho_b = \partial\rho/\partial b^* = dr/db^* b^* + 1 + r(a,b^*) > 1$ and $c`(a)$ is the change in total cultivation

[7]In principle, a similar argument may be made for family labor. A farmer with a larger family size may have less need to finance hired labor inputs given area and thus will borrow less and face a lower interest cost per unit area. Profit estimates that did not remove variation in borrowing cost might underestimate his relative profitability. However, this ignores that family workers and their dependents must be fed throughout the agricultural cycle, which reduces the liquidity benefits of having a large endowment of family labor per unit of area farmed. We do not formally model consumption and family size here except to note that with food shares at 60-80% it is unlikely that the net liquidity effects of family labor will be substantial.

costs, conditional on work, which decline with area in the absence of credit market constraints, as shown in Appendix A. Profits per acre rise with the size of owned landholdings both because of scale economies and because of the existence of credit market imperfections. The latter, as modeled here, steepens the gradient of per-acre profits with respect to owned area relative to cultivated area, for given (or zero) credit costs. This is for two reasons. First, there is a negative effect of owned area on interest rates given input use per acre, $\rho_a < 0$. Second, any savings in cost per unit of work associated with scale lower the amount borrowed, thus further lowering interest costs and raising profitability.

In addition to affecting the input choices of farmers, the presence of credit market imperfections creates an empirical problem in measuring true profitability because of the difficulty of accounting for differences in interest rates and thus the true discounted costs of inputs across households in informal credit market settings. Expression (10) is relevant to the question of whether land consolidation will improve true profitability per acre. In Appendix B we consider the empirical question of whether it is possible to infer correctly the role of credit market constraints in the relationship between owned landholdings and (true) per-acre profitability when borrowing costs are ignored in computing farm profits. We thus consider the comparative statics associated with *estimated* profits, which exclude interest costs. Estimated profits is the most common measure of profits, and the one we use in the empirical work due to the difficulty of obtaining reliable and consistent measures of interest costs. We show that (i) both estimated profits and true profits have a steeper gradient with respect to scale if there are credit market imperfections, as in (10), and (ii) the finding that there is a non-zero return to owned capital assets using estimated or true profits would reject the hypothesis of perfect capital markets.

Thus far we have taken the amount of owned capital assets as given. In practice, farmers both own and rent machinery, and the model incorporating credit constraints can explain variation in equipment ownership even in the presence of a perfect rental market. By the assumption of an effective rental market all farmers face the same equipment rental price. But due to credit market imperfections farmer with different landholdings face different borrowing costs. Given that the rental-equivalent price of owning machinery for one agricultural season depends on one's own cost of borrowing, individuals with relatively low borrowing cost will be more likely to own machinery and those with higher borrowing cost will rent it. This suggests that if, as in (7), financial intermediaries lower the cost of borrowing for larger versus smaller landowners, then given an active rental market, larger farmers will be more likely than small farmers to purchase rather than rent machinery following the entry of such intermediaries.

C. *Farm size and profit dynamics*

In the preceding section we assumed that the amount a farmer borrowed reflected only his demand for inputs and his ownership of equipment, ignoring own savings as a source of liquid capital. In this section we consider the role of landholdings in determining profitability in a dynamic setting in which profits are stochastic and liquid capital, or cash on hand, affects input allocations when credit market imperfections are in place. In this setting, if there are credit restrictions a farmer who has particularly high profits in one period may be able to finance more inputs and thus accrue greater profits in a subsequent period. If he has access to large amounts of capital at market rates no such effects should be observed.

There are other reasons why there may be a correlation in profits across time for a given farmer. For example, it is well-known that fertilizer use increases nutrient levels in the soil that persist over time. This persistence will influence fertilizer use and thus profitability in a subsequent period. Because past fertilizer use will augment past profitability, one might observe a negative correlation between past profits and current fertilizer use. Inattention to dynamic nutrient effects might lead to the false conclusion that credit constraints are unimportant even if credit imperfections were present.

Addressing these dynamics in a forward-looking model is complicated and thus we illustrate the basic structure using a simplified production function with one variable input, fertilizer, and assume that the production function and the cost of borrowing are quadratic in their respective arguments. In this model, farmers adjust their end-of-season savings based on unanticipated income shocks and subsequently use this savings to finance fertilizer purchases. We assume a stationary problem with state variables representing soil nutrition n^* and cash on hand h^*. Fertilizer levels are chosen prior to the realization of a shock θ_t. We define the value function recursively as follows:

(11) $$v(n_t^*, h_t^*) = \max E_t\,(\theta_t + g(f_t^* + n_t^*) - r(f_t^* - h_t^*) - h_{t+1}^* + \beta v(n_{t+1}^*, h^*_{t+1})),$$

where β is the discount factor and

(12) $$h_{t+1}^* = h_t^* + \lambda\,(\theta_t - E_t\theta_t),$$

where λ denotes the extent to which unanticipated shocks are saved. For $\lambda = 1$ unanticipated shocks are fully saved as in the permanent income hypothesis and for $\lambda = 0$ cash on hand is just a constant. Soil nutrients depend positively on both the previous period's stock of nutrients and fertilizer use and negatively on the output shock θ_t,

(13) $$n_{t+1}* = n_t* + f_t* - \alpha\theta_t$$

The idea is that more rapidly growing plants, for example, will deplete the soil of nutrients relatively quickly. For example, if θ_t is rainfall, more nutrients are used if rainfall and soil nutrients are complements. The production and credit functions are

(14) $g(x) = g_1 x - g_2 x^2$ and (15) $r(x) = x + r_2 x^2$,

where x are the respective arguments in (11) and for notational simplicity we set the fertilizer price to one. All of the parameters in (14) and (15) are positive; that is we assume that production is characterized by diminishing returns but the cost of credit r increases at a higher rate with the amount of credit.

Estimated profits in this model (again, profits that do not account for borrowing costs are):

(16) $$\hat{\pi}_1* = \theta_1 + g(f_1* + n_1*) - f_1*$$

Farmers optimally choose their level of savings and use of fertilizer. Given the soil dynamics and savings behavior, the effect of a previous period shock on next-period's profits is thus

(17) $$\frac{d\pi_1*}{d\theta_0} = -\alpha \frac{g_2 + g_1 r_2 - \beta v_{nn}}{g_2 + r_2 - \beta v_{nn}} + \lambda \frac{(1-\beta)(g_1 - 1) r_2}{g_2 + (1+\beta) r_2 - \beta v_{nn}},$$

where v_{nn} is the second derivative from the value function, with $v_{nn} < g_2 + r_2$ and $g_1 < 1$ for an interior maximum.

The two key parameters in (17) are α and λ, reflecting the influence of the dynamic nutrient and savings functions. If $\lambda = 0$ so that liquidity h does not depend on unanticipated income shocks the lagged profit shock only influences profits in the next period because of nutrient depletion. A positive shock in period 0 in that case leads to greater nutrient depletion and therefore reduces profitability in period 1. Conversely, if there is no nutrient carryover so that $\alpha = 0$ there is only a liquidity effect: a positive shock in period 0 induces higher savings and thus more cash on hand in the next period so that less credit is needed for fertilizer. The lower cost of borrowing increases fertilizer use and thus increase profitability in the current period. This effect vanishes if $r_2 = 0$, that is, if borrowing costs do not rise as the demand for credit increases.

The model thus implies that the finding of a positive lagged profit shock effect on (estimated) profits is indicative of the presence of liquidity effects. However, it also suggests that the liquidity effect may be obscured even in the presence of credit market failures due to soil nutrient dynamics. We show below that the nutrient depletion and credit-market effects can be separately identified using plot-specific information over time for farmers with multiple plots. The idea is that the nutrient effect only operates for a given plot but that the liquidity effect arises from an aggregate farm-level shock.

3. Data

Our empirical investigation of the relationships among scale, credit markets, labor use, and profitability uses four types of data from two surveys that form a panel. The main data sets are the 2007-8 Rural Economic Development Survey (REDS 2007-8) and the 1999 REDS both carried out by the National Council of Applied Economic Research (NCAER). The surveys were administered in 17 of the major states of India, with Assam and Jammu and Kasmir the only major states excluded. The two surveys are the fifth and sixth rounds of a panel survey begun in the 1968-69 crop year. The original sample frame was meant to be representative of the entire rural population of India at that time. To obtain nationally-representative statistic from the first round data, sampling weights must be used because a stratified sampling procedure was employed to draw the sample. This included the oversampling of high-income households within villages and selecting districts in areas particularly suitable for green revolution crops. By the sixth round, 40 years later, given household splits, the creation of new towns and villages, and out-migration, the original sampling weights no longer enable the creation of nationally-representative statistics from the later-round data. The data can be used to estimate relationships among variables that characterize behavior in the population. The oversampling of high-income households is an advantage for this study, given our focus on the relationships among scale, productivity and mechanization, because there is more variation in own landholdings at the upper tail where mechanization is prevalent.

The 2007-8 survey includes a listing, carried out in 2006, of all of the households in each of the original 242 villages in the panel survey.[8] The listing data, which included almost 120,000 households, will be used in the final section to examine land leasing patterns within villages. The survey of households in the 2007-8 REDS, took place over the period 2007-2009, and includes 4,961 crop cultivators who own land. The sample of farmers include all farmers who were

[8] Appendix Figure A provides the distribution of own landholdings in the set of sampled villages in comparison to that from the Census of 2001. The figure shows that landholding distribution in the sample villages is skewed to the right relative to the national figures. This is not due to the oversampling of high-income households, but reflects the geographical sampling.

members of households interviewed in the 1999 round of the survey plus an additional random sample of households. The panel households include both household heads who were heads in 1999 and new heads who split from the 1999 households. There are 2,848 panel households for whom there is information from both the 1999 and 2007-8 survey rounds. The 2007-8 survey is unique among the surveys in the NCAER long-term panel in that information on all inputs and outputs associated with farm production was collected at the plot level for each of the three seasons in the crop year prior to the survey interviews. There is input-output information for 10,947 plots, with about two-thirds of the plots observed at least twice (two seasons or more).

The plot/season data enable us to carry out the analysis across plots in a given season, thus controlling for all characteristics of the farmer, including the input and output prices he faces. Cross-plot, within-farmer estimates can be biased if plot sizes are chosen by farmers and plots vary by unmeasured characteristics that affect productivity. With respect to plot size, a farmer's ownership holdings consists mainly of parcels of inherited land.[9] A plot is a contiguous area of land that a farmer considers to be the basic unit of farming. Three-quarters of plots consist of a single parcel; the rest are sets of contiguous parcels. Plots are thus almost always sets of inherited parcels that are contiguous, as in less than 4% of cases is a contiguous parcel broken into more than one plot. The median distance between plots is 400 meters. Plot size is thus principally determined by the location and configuration of the farmer's inherited parcels of lands.

The survey includes seven characteristics that characterize the quality of plots. These include depth, salinity, percolation, drainage, color (red, black, grey, yellow, brown, off-white), type (gravel, sandy, loam, clay, and hard clay) and distance from the farmer homestead. In addition, there is within-crop-year panel data of up to three periods (seasons) on farm plots. The multiple season information by plot enables us to obtain estimates that control perfectly for plot characteristics as well as time-invariant farmer characteristics, as discussed below.

The detail on inputs, outputs and costs enables the computation of farm profits at the plot and farm level for the 2007-8 survey round and at the household level for the 1999 round. Information is provided on the use of family, hired and supervisory labor by operation and by age and gender, along with own use of implements by type and the rental of implements, by type. Other inputs include pesticides, fertilizer (by type), and water. We subtract out the total costs of

[9] 96% of owned parcels were acquired, principally through inheritance, from an immediate family member or grandparent.

all of these inputs from the value of output using farm gate prices. Thus, our profit measure corresponds to 'empirical' profits in the model as it does not include interest costs associated with using credit to pay for inputs. Maintenance costs for own equipment is subtracted from gross income, but not maintenance costs (meals, clothing, shelter) for family labor.

The 2007-8 survey also includes retrospective information for each household head on investments in land and equipment, by type, since 1999. This includes information on land and equipment that is sold, purchased, destroyed, transferred or inherited. This information will be used to estimate the determinants of farm mechanization. The acquisition of land is primarily via inheritance that results from family division - less than 3 percent of farmers bought or sold land over the entire nine-year period. Division most often occurs when a head dies and the adult sons then farm their inherited land. Division sometimes occurs prior to the death of a head, which may result from disputes among family members (Foster and Rosenzweig, 2003). Time variation in the state variables owned landholdings and equipment thus principally stems from household splits.

Two key assumptions of the model are (i) that the rental of land does not overcome the limitations of scale associated with owned plots and (ii) that family labor does not have a cost advantage over hired labor in terms of the need for supervision. With respect to the first assumption, the 2007-8 data indicate that only 4.6 percent of cultivated plots, over the three seasons, are rented (4.9 percent of area). Moreover, the data indicate that in all states of India, except West Bengal, 72% of cultivated land is leased from immediate family members (parents and siblings). This is not unexpected, given the presumed efficiency of cultivating contiguous land area along with possible moral hazard issues that might arise in terms of farm maintenance.[10]

A key feature of the 2007-8 data, as noted, is that it includes information on supervision time at the plot level. We estimate a supervision cost equation across plots for the same farmer in a given season:

(18) $$l_{sijt} = \mathrm{a}_{0jt} + \mathrm{a}_{\mathrm{f}} l_{fijt} + \mathrm{a}_{\mathrm{h}} l_{hijt} + \mathrm{a}_{\mathrm{A}} A_{ijt} + \boldsymbol{X_{ij}} \mathbf{a}_x + e_{ijt},$$

where for plot i of farmer j in season t l_s=supervisory labor costs, l_f=family labor costs, l_h=hired labor costs, A=plot area, $\mathrm{a}_{0\mathrm{j}}$= farmer/season fixed effect, $\boldsymbol{X_{ij}}$ is a vector of plot characteristics, and e is an error term, where costs are simply days times the relevant wage per day. Our assumption is that $\mathrm{a}_\mathrm{f} = \mathrm{a}_\mathrm{h} > 0$, that an increase in hired or family labor equally increases supervision time. Note

[10] In contrast, in West Bengal, 26% of farmers rent from landlords, and only 7% from family. The third category in the data is 'other.'

that the farmer/season fixed effect picks up all market prices and all farm-level characteristics in a given season, including the number of family members.

It is important to control for farmer characteristics to obtain an unbiased estimated of a_f and a_h. Supervision is typically carried out by family members, presumably because family members benefit directly from profitability. This is one of the advantages of family farming. Supervision time thus may depend on family size if supervision is carried out less efficiently using hired labor. Farm households that have a greater number of family members thus may both use family labor more intensively and spend more time supervising. This would create a positive bias in the coefficient on family labor coefficient in (18).[11] The number of family members is impounded in the farmer/season fixed effect; the a-coefficients estimated with the farmer/season fixed effect included thus pick up how supervision costs vary across plots according to the allocation family and hired labor, for given family size.

Table 1 reports the estimates of (18). In the first column, only village and season dummy variables are included - there is no control for farm characteristics, including family size. The first-column estimates indicate that an increase in non-supervisory family labor use increases supervsion costs more than an increase in hired labor, $a_f > a_h$. This result is robust to the inclusion of the set of detailed plot characteristics. When we control for farmer and season and thus family size, however, the coefficient estimates for family and hired labor are essentially identical - we cannot reject the hypothesis that increasing family or hired labor use increases supervsion costs equally. This result is also robust to the inclusion of plot characteristics. The difference in the coefficient estimates across columns 2 and 4 do suggest that larger families may have an advantage in supervision, for given scale, but not because family manual laborers require less supervision than do hired laborers. Thus, as the model suggests and as the descriptive statistics confirm, mechanization also reduce the need for supervision by decreasing the total use of manual labor.

4. Identifying Scale Effects

As indicated in the model, larger landholdings potentially increase profitability by allowing the use of a higher-capacity (or any) mechanized inputs and also by lowering credit costs. In this section, we identify the effects of scale, net of credit cost effects, by estimating how

[11]We could include family size in the specification, but family size may be endogenous: on farms where supervision is particularly advantageous, or hiring supervisiory labor is difficult, more family members may be in place.

variation in the size of plots for a given farmer affects plot-specific per-acre profitabilty, the likelihood of tractor use and per-acre labor use. By using farmer fixed effects we are holding constant owned landholdings, access to credit (and family size) so that variation in area reflects only scale effects. We also examine the role of fragmentation. We estimate the equation

$$\pi_{ijt} = \mathrm{b}_{0j} + \mathrm{b}_{\mathrm{A}} A_{ijt} + \mathrm{b}_{\mathrm{-A}} A_{-ijt} + \mathrm{b}_{N} N_{ijt} + \boldsymbol{X}_{\boldsymbol{ij}} \mathbf{a}_{x} + u_{ijt}, \tag{19}$$

where π_{ijt}=profits per acre on plot i for farm j, b_{0j}=farmer fixed effect, A_{ijt}=plot area (acres), A_{-ijt}=total area of all other plots, N_{ijt}=total number of cultivated plots, and u_{ijt} is an iid error. The equation also includes season/state fixed effects to control for input and output prices. The interpretation of the coefficient on plot area b_{A} is straightforward - it is the effect of scale on profits. However, as noted by Lamb (2003) and Barrett *et al.* (2010), measurement error in plot size biases the own size coefficient negatively. Thus, we will get a lower-bound estimate of scale. For given total size of the other plots A_{-ijt}, an increase in the number of plots N_{ijt} is interpreted as a decrease in the average size of the other plots. If other plots are smaller, use of mechanized inputs is less likely so that more resources may be allocated to the larger plot because inputs will have a higher return. The coefficient b_{N} would then be positive. An increase in the total size of all plots might make the rental or ownership of higher-capacity equipment more profitable for the farm, thus also increasing profits on all plots ($\mathrm{b}_{\mathrm{-A}}$>0).

The first column of Table 2 reports the estimates of equation (19) without the inclusion of the seven plot characteristics. The second column reports estimates with the plot characteristics included. In both specifications, the estimates are consistent with the operation of scale economies - larger plots, given the farmer’s total ownership holdings, capabilities, preferences, and family size, are associated with higher profits per acre, despite the negative bias induced by measurement errors. And, if other plots are on average smaller or total cultivated area on the farm is greater, the plot is also more profitable. The inclusion of plot characteristics does little to alter the estimates, a finding consistent with Barrett *et al.* (2010), who also exploit variation in plot sizes to estimate scale effects.[12]

Are these profit estimates consistent with scale effects associated with mechanization?

[12] Many investigators use output value per acre to measure productivity. Assunção and Braido (2007) employ a similar cross-plot methodology based on ICRISAT data that indicates a negative relationship between per-acre output value and plot size, net of a farmer fixed effect and multiple measures of plot characteristics. However, measurement error biases negatively the coefficient. In our data, the relationship between own plot size and output value per acre is positive, small and statistically insignificant.

The third and fourth columns reports estimates of equation (19) with per-acre profits replaced by a dummy variable taking on the value of one if a tractor is used on the plot. By using a dummy variable, we avoid the problem of measurement error in plot size biasing the size coefficient negatively. The estimates, with and without plot characteristics included, indicate that, consistent with the theory, a tractor is more likely to be used on a larger plot and if the total amount of cultivated area is larger (for given owned area), but if the farm has smaller plots on average, a tractor is less likely to be used. In columns five and six we see that total labor costs per acre mirror the effects of scale on plot-specific tractor use - larger plots use less labor per acre and less labor is used per-acre, given plot size, the larger is the total cultivated area of the farm.[13] But, the smaller are farm plots overall, the higher are per-acre labor costs on any plot.

5. Owned Landholdings, Efficient Input Use and Profitability

In the preceding analysis we examined at how profitability varied across plots for fixed land ownership. To explore the overall effects of total farm size on input efficiency we exploit the plot level data to estimate the marginal returns to a variable input by farm size. Profit-maximization implies that the marginal return to an additional rupee spent on a variable input should be zero. We estimate the marginal returns to fertilizer expenditures based on variation across plots in fertilizer use for a given farmer, stratifying the sample by the size of owned landholdings. If farmers with small landholdings face higher borrowing costs and are unable to finance the efficient use of fertilizer, we should find that the marginal returns to fertilizer expenditure are positive for small farmers but decline as farm size increases. The equation we estimate includes a farmer/season fixed effect so that only plot-specific characteristics enter the specification. These again include plot size and soil characteristics. The fixed effect (FE) thus absorbs farmer characteristics and input prices faced by the farm:

$$\pi_{ijt} = c_{0jt} + c_A A_{ijt} + c_f f_{ijt} + \boldsymbol{X_{ij}}\mathbf{a}_x + \varsigma_{ijt}, \tag{20}$$

where f_{ijt}=plot-specific fertilizer expenditures per acre and ς_{ijt} is an iid error. Profit-maximization implies that c_f=0.

Table 3 presents FE estimates of (20) for farmers who own less than four acres of land, farmers with landholdings above four and less than 10 and for farmers who own more than 10 acres. Estimates are shown with and without the plot characteristics. We can reject the hypothesis

[13]Measurement error in plot size biases negatively the effect of own size on labor use. In the penultimate section we report estimates of landholdings on labor costs per acre that are robust to measurement error and to any bias due to the endogeneity of scale.

that farmers below 10 acres are using sufficient fertilizer; the coefficient on fertilizer is statistically significantly different than zero for both the <4 and 4-10 acre farmers. Indeed, the estimates imply that an additional rupee of expenditure on fertilizer yields more than a rupee of profit.[14] For farmers owning 10 or more acres of land, however, the marginal return is effectively zero on average; such farmers are evidently unconstrained in fertilizer use. The partition of farmers into three groups is somewhat arbitrary. Figure 3 presents smoothed local-area FE estimates of the effects of fertilizer on profits by land ownership from (20), along with one-standard deviation bands, for farms up to 50 acres in size. The pattern of estimates indicate that the marginal returns to fertilizer fall monotonically as landholdings increase and fertilizer is underutilized, given the prices that farmers face, for farms up to about 40 acres.

To estimate the effect of total land owned land and machinery on per-acre profitability we need to allow for the possibility that landownership and machinery are correlated with unmeasured attributes of farmers. We use the 1999-2007-8 panel data to estimate the causal impact of landownership on profits and on investment at the farm level. Prior studies have exploited panel data to eliminate time-invariant fixed farmer and land characteristics such as risk aversion or ability. However, this is not sufficient to identify the effect of variation in a capital asset. The equation we seek to estimate is

$$\pi_{jt} = \mathrm{d}_{0t} + \mathrm{d}_A A_{jt} + \mathrm{d}_k k_{jt} + \mu_j + \varepsilon_{ijt}, \tag{21}$$

where t is survey year, k=value of all farm machinery, μ_j=unobservable household fixed effect, and ε_{ijt}=an iid error. Controlling for farm machinery (mechanization), we expect that the coefficient $\mathrm{d}_A>0$ if there are scale effects and also that machinery has a positive marginal return ($\mathrm{d}_k>0$), perhaps a higher return for small farms that are unable to finance capital equipment purchases. The problem is that farmers who are unobservably (to the econometrician) profitable may be better able to finance land purchases and equipment, leading to a spurious positive relationship between landholdings, capital equipment and per-acre profits.

Taking differences in (21) across survey years to eliminate the farmer fixed effect, we get

$$\Delta\pi_{jt} = \Delta \mathrm{d}_0 + \mathrm{d}_A \Delta A_j + \mathrm{d}_k \Delta k_j + \Delta\varepsilon_{ijt}, \tag{22}$$

where Δ is the intertemporal difference operator. In (22), even if the errors ε_{ijt} are iid, investments in capital assets such as land or equipment will be affected by prior profit shocks in a world in

[14] Recall that our profit measure does not include credit cost. If costs of credit are high then we are in effect underpricing fertilizer in the computation of profits.

which credit markets are imperfect. By differencing we thus introduce a negative bias in the land and equipment coefficients - positive profit shocks in the first period make ΔA_{ijt} high when $\Delta\varepsilon_{ijt}$ is low. That is, even if the contemporaneous cov(ε_{ijt}, A_{jt}) = 0, because assets are measured prior to the profit shock, cov($\Delta\varepsilon_{ijt}$, $_{\Delta aj}$) ≠ 0. We show below that for most farms (small farms) in India there is underinvestment in machinery and that past profit shocks affect current variable input use.

To obtain consistent estimates of d_A and d_k we employ an instrumental-variables strategy.[15] We take advantage of the fact that over the nine-year interval between surveys 19.9% of farms divided and farmers inherited land. Moreover, for all heads of farm households in 1999, we know how much of the land and equipment was inherited before the 1999 survey round. The instruments we use to predict the change in landholdings of a farmer between 1999 and 2007-8 are thus the value of owned mechanized and non-mechanized assets inherited prior to 1999 and the value of assets and acreage of land inherited between 1999 and 2007-8. We also add variables that in our prior study of household division in India (Foster and Rosenzweig, 2002) contributed to predicting household splits and the size of inheritance. Splits, and the inheritance of land, most often occur at the death of the father. We thus include the age of the father in 1999. In some cases, sons choose to split from the family prior to the death of the father/head. Thus, the father of a farmer in 2007-8 may not have been co-resident in 1999, the son already having split. We thus also include a dummy indicating the resident status of the father in 1999. In our earlier study we found that inequality among claimants (principally siblings) was a significant predictor of household division for households in which the father was co-resident. We thus also include a measure of the educational inequality among the claimants to the head's land in 1999 and an indicator of whether the farmer in 2007-8 had brothers.

Appendix Table A contains the estimates of the first-stage equations predicting the change in landholdings and the value of farm equipment between 1999 and 2007-8. The Anderson Rubin Wald test of jointly weak instruments rejects the null at the .005 level of significance.
Indeed, post-1999 inheritance of land is a significant predictor of the change in landholdings over the period along with the indicator of whether the father was not co-resident in 1999, while inherited assets obtained prior to 1999 and inequality in claimants statistically and significantly affect the change in the stock of equipment. We estimate a variant of (22) in which we omit

[15]Note that using IV also eliminates the negative bias due to measurement error in landholding size afflicting both the rhs and lhs variables.

capital equipment in order to estimate the unconditional relationship between landownership size and profitability gross of mechanization. The first two columns of Table 4 report estimates of the two variants of the per-acre profit function (22) but only controlling for village and time effects, where the reported t-ratios are clustered at the 1999 farm level. This estimation procedure roughly, by village area, controls for land quality heterogeneity and prices, but not individual farm heterogeneity. The estimates indicate that larger farms are more profitable per acre, consistent with Figure 2, but capital equipment has little or no return, conditional on farm size. The farmer fixed-effects estimates are reported in the third and fourth columns of the table. These estimates suggest that there are no scale effects and that larger farms are not profitable. However, as discussed, these estimates are biased negatively to the extent that there are credit constraints on capital investments.

The last two columns of Table 4 report the FE-IV estimates that eliminate the bias in the farmer fixed-effects estimates. These show that an exogenous increase in landholdings gross of changes in capital equipment significantly increases per-acre profits - a one-acre increase in landholdings at the mean increases per-acre profits by 9.2%. A large proportion of this effect is evidently due to investments in equipment; controlling for farm equipment reduces the effect of farm size by 36%. And, the marginal return on capital assets is positive and statistically significant, at 3.5%. The Kelinberger-Paap and Hansen J diagnostic test statistics, reported in the table, indicate that we cannot reject the null that the second-stage estimates for either specification are not identified.

Does the data indicate that there is an optimal farm size? Or put differently, is there a farm size at which additional increases in owned land no longer increase profits per acre? Figure 4 reports the locally-weighted FE-IV land coefficient d_A by land ownership size ranging from 0.1 to 20 acres. The continuous line depicts the estimated coefficient from the specification that excludes capital equipment; the discontinuous line portrays the coefficient of land size conditional on owned capital equipment. As can be seen, for the entire range of farm sizes, increases in land increase profits per acre; the positive effects of size actually rise with land size for farms below 5 acres. That is, for 95% of the farms in India, increasing farm size would raise profits per acre at an increasing rate.

If credit costs decline with land size, as we have assumed, the marginal returns to capital should decline with land ownership size. Figure 5 reports the locally-weighted FE-IV estimates of the marginal return to capital equipment d_k, along with the associated one standard deviation bands, across the same range of owned landholdings. Marginal returns evidently do decline as landholding increase - for farm sizes at around three acres, the return to capital is between .04 and .10, while for farms of 10 acres, the return is between two and four percent. Smaller farms substantially under-invest in capital equipment.

6. Farm Size and Equipment Investment and Rental

Figure 5 suggests that credit costs fall with landownership, given the underinvestment in machinery characterizing small farms. In this section we estimate the effects of landholdings on equipment investment and rental. The model suggests that farms owning more land will purchase more capital equipment to take advantage of scale economies and because they face lower credit costs. For this analysis we use the retrospective information from the 2008-9 REDS that provides a yearly history of land and capital equipment acquisition from 1999 up to the survey interview date. In contrast to the panel data based on information from the 1999 and 2007-8 survey rounds in which the household unit is defined by the households in 1999, 19% of whom split, the unit for this analysis is the household in 2007-8. There are two consequences. First, the sample is larger than the 1999-2007-8 panel, because the latest survey round includes a new random sample of households. Second, if a farmer split from a household after 1999 his owned land and farm assets at the 1999 date is reported as zero if he was not formerly the household head. 25% of the sample farmers in 2007-8 experienced an increase in landholdings since 1999, of whom 79% inherited land due to household division. Less than 1.2% of farmers were observed to experience a decline in owned landholdings.

We create a panel data set from the retrospective history by computing any new investments made in farm machinery within the three-year period prior to the 2007-8 interview data and within the three year period 1999-2001. We also compute the stock of equipment and landholdings in 1999 and three years before the interview in the last round. Thus we create two observations on capital investment, landholdings and equipment stock value for each farmer.[16] We also examine the determinants of equipment rental. Here we must use information on the value of hired equipment services in 1999 and in 2007-8 from the 1999 and the 2007-8 surveys, so that the sample size is reduced to the matched 1999-2007-8 panel.

[16]In principle the data can be used to examine the determinants of net land sales. However, less than 2% of farmers sold or purchased land over the 9-year interval. In contrast, 18% of farmers invested in capital equipment.

Our model incorporates credit market imperfections as one of the factors that constrain mechanization, with owned landholdings serving to mitigate credit costs. We thus add to the household panel information on bank proximity. From the 1999 and 2007-8 village-level data providing comprehensive information on village institutions and facilities, we created a dummy variable indicating whether a commercial bank was within ten kilometers of the village in which the farm household was located. 84% of farmers were within 10 kilometers of a bank in 1999; 84% in 2007-8. However, banks were not stationary. 25% of the farmers experienced either the exit of a bank or a newly-proximate bank. To assess whether landownership plays a role in lowering credit costs, we interact landholdings and bank presence. The equipment purchase and hire equations we estimate are thus of the form:

(23) $$K_{kjt} = e_{0t} + e_A A_{jt} + e_k k_{jt} + e_B B_{jt} + e_{BA} A_{jt} \cdot B_{jt} + \mu_j + \eta_{ijt},$$

where K=equipment purchase or rental and B=bank proximity. We expect that $e_A>0$, $e_k<0$, and $e_{BA}>0$; that is, where banks are present, large landowners are more able to finance equipment purchases and/or rent equipment. To eliminate the influence of unobserved time-invariant farm and farmer characteristics (μ_j), we again difference across the two periods and use instrumental variables to eliminate the bias discussed in the previous section. Because a little over half of the observations in the retrospective-based panel are from the newly-drawn sample of households in 2007-8, we do not use information on family circumstances in 1999 as instruments, which is only available for the 199-2007-8 panel. We use as instruments for the change in owned landholdings, the change in the value of farm equipment and the change in bank presence, the value of farm assets inherited since 1999, the amount of land inherited since 1999 and bank proximity in 1999.

The estimates of (23) are presented in Table 5; the first-stage estimates are presented in Appendix Table B. Because here the second-stage estimates are exactly identified, we cannot use the standard diagnostics tests of identification. However, inherited land is a statistically significant predictor of the change in owned landholdings, inherited assets are statistically significant predictors of the change in the value of the stock of machinery, and bank presence in 1999 is a statistically significant predictor of subsequent bank location.

The first column of Table 5 reports fixed- effects estimates of the determinants of machinery investment that do not use the instruments and which exclude the interaction term. While the signs of the coefficients are as expected, the precision of the coefficient estimates is low for both land and the equipment stock. When instruments are used, however, as reported in the second column, both the capital equipment and land coefficients increase substantially and become statistically significant. In particular, an increase in owned landholdings increases equipment investment, given the existing stock of equipment, while for given landholdings, those farms that already own equipment invest less. The effect of bank presence is not precisely estimated, however, in the linear IV specification. When the interaction between bank proximity and landholdings is added (column three), we see that evidently the advantage of bank proximity is only captured by larger landholders - the interaction coefficient is positive and statistically significant while the linear bank and land coefficients become statistically insignificant. These results are consistent with land having value as collateral for obtaining bank loans to finance equipment purchases.

The estimates in columns four through six in Table 5 for equipment rental parallel those for equipment purchases, except that the interaction term is not statistically significant - the fixed-effects estimates are negatively biased, but once this bias is eliminated using instrumental variables large landowners are seen to rent more machinery than smaller landowners, for given owned stock. But bank proximity is not a statistically significant determinant of equipment rental in either the linear or interactive specification. Formal banks thus do not appear to play a major role in financing variable inputs. These results thus indicate that larger landowners are more likely to use and own farm machinery, and part of the reason is that they have better access to lower-cost bank credit for investment.

7. Credit market imperfections, size, and the effects of profit variability

The previous section provided indirect evidence on the role of credit market imperfections as a source of scale economies in rural agriculture. We used our model to show that a more direct test is possible of the interaction between credit market imperfections and scale by examining the sensitivity of profits to income shocks by land ownership. In this section we estimate how lagged profit shocks affect per-acre profits, taking into account that such shocks not only affect farmer liquidity but also soil nutrients. As noted in the theory section, assessment of the effects of past shocks on current profitability is complicated by the fact that past crop shocks may also affect the nutrient content of the soil, which, in turn, may also affect profitability. Removing farmer and/or plot specific fixed effects from estimates of a profit equation may remove fixed aspects of soil quality that affect profit but will not help if nutrient status is responding directly to past shocks.

To separate out credit effects from dynamic nutrient effects we make use again of the fact that we have plot-level data for each farmer over three consecutive seasons, which allows us to separate the effect of a crop shock on liquidity from the effect of the shock on soil nutrients. We augment the dynamic model by letting cash on hand depend on the unanticipated deviations in the across-plot average shock so that $h_{t+1}{*} = h_t{*} + \delta(\overline{\theta}_t - E_t\overline{\theta}_t)$. The key distinction is between lagged profits and fertilizer use on a given plot and lagged profits on all other plots. The coefficient on the lagged profits specific to a plot will capture the combined nutrient and (a small fraction of) liquidity effects; the coefficient on the lagged profits from other plots will only reflect the liquidity effect. To achieve identification, we use a subsample of farmers who cultivate at least two plots over three seasons.

The results, reported in Table 7 for three categories of farmers based on own landholdings, are strongly consistent with the notion that liquidity shocks affect input use and thus profitability among small farmers. In particular, conditioning on the lagged profits and fertilizer use on a given plot, a 1000 Rupee increase in profits per acre on a farmer's other plots leads to a substantial 140 Rupee increase in profits per acre among farmers with less that four acres of land. In farmers with 4-10 acres of land the corresponding figure is substantially less (62.8 Rupees) and in the largest farmers (10+ acres) the estimate is even smaller (36.6 Rupees), with neither estimate differing significantly from zero.[17] The corresponding coefficients on the profits on the particular plot also decline, consistent with the idea that the own profit effect combines both a technological effect (in this case nutrient depletion) that is constant across landholding and a declining liquidity effect.

9. Lease markets

The preceding results suggest that there are substantial unrealized returns to profitability in rural India that are a consequence of current small farm sizes. If this is indeed the case, then even in a setting in which there are important barriers to wide-scale land amalgamation, one should expect to see transfers of land that on net move land from smaller to larger farms.[18] Leasing, however, cannot overcome the scale effects associated with credit markets. Moreover, given that scale economies arise in part from contiguous land, the opportunities for productive

[17]Appendix Figure 6 plots the locally-weighted IV estimates of the lagged farm profit effects for the span of land sizes ranging from .01 to 20 acres, which shows that lagged profit effects are positive for all farm sizes below 15 acres.

[18]As we have noted land sales are too scarce to characterize patterns. Moreover, our data suggest that land sales too are intrafamily - 95% of land sales are from parent to child.

trade given ownership holdings appear to be small, as reflected in the fact that less than 10% of farmers lease and most rental contracts are with family members. However, the 2006 village listing data gives a large enough sample size to examine the distribution of this relatively rare event across farms stratified by ownership size to assess if Indian farmers seek to exploit scale economies.

The relationship between ownership holdings and the probability of leasing in and leasing out in the 2006 listing data, net of village fixed effects that characterize the relevant supply of existing land plots, are plotted in Figure 6. The figure strongly supports the hypothesis that leasing goes in the direction of capturing scale economies - reverse tenancy. In particular, relative to a household that is 5 acres below the village mean a farmer with 5 acres above the village mean has .018 (over 50%) higher probability of leasing in and a .014 lower probability of leasing out. Indian farmers appear to behave as if they also believe that increasing operational scale, when feasible, is profitable.

10. Estimates of surplus labor and surplus farms

The evidence we have obtained from our plot and farm-level panel data suggests that increasing farm size in the range of sizes characterizing most Indian farms would raise agricultural profits due to scale economies associated with mechanization and reduced credit costs. An implication of these findings is that a process of land amalgamation that results in larger farms in India would increase the overall efficiency of Indian agricultural production without any innovations in credit markets or technological progress in production. Our model and the plot-level estimates also implied that increased profitability is due in part to the reduction in labor costs. This suggests that increasing the average scale of Indian agriculture could release agricultural labor without any reduction in agricultural wages. As noted earlier, the ability of the agricultural or other traditional sectors to release labor to the industrial sector at prevailing wages seems to conform closely to the idea of surplus labor as posited by Lewis and others as an important ingredient to the development process. What do our estimates imply for how much labor can exit from agriculture without impacting wage rates based on unexploited scale economies?

To gage the amount of the Indian agricultural labor force that is surplus we first estimate the direct of effects of ownership holdings on total labor costs per acre using the 1999-2007-8 panel data and the same two-stage procedure (and instruments) used to identify land scale effects on profits. By using instrumental variables we avoid the bias due measurement error in farm size that may afflict the plot-level estimates of labor savings. Moreover, the total landholdings effect on labor use, unlike the FE-farmer plot-based estimates also incorporates the effects of cost savings associated with lower credit costs.

Consistent with the plot-level results reported in Table 2, the estimated land size coefficient for per-acre labor costs is negative and statistically significant. The point estimate, -273.4 with a standard deviation of 135.3, indicates that a one acre increase in landholdings reduces total labor costs per acre by 13% at the sample mean. Labor savings from increasing scale are evident across the bulk of the distribution of Indian farms. Figure 7 plots the locally-weighted IV estimates for the span of land sizes ranging from .01 to 20 acres. As can be seen, per-acre labor savings are associated with increasing ownership size at all landholding sizes up to 20 acres, with the cost-saving effect increasing for farm sizes below five acres, where almost all farms in India are located, and reaching a maximum at five acres, where our estimates indicated (Figure 4) the maximum profit gain from increasing scale also occurred.

To estimate labor surplus from unexploited scale economies based on these estimates of labor savings we carry out a simulation based on the actual land distribution in India in which we set a land-size floor and amalgamate the land of any farms whose size is below this floor to the level of the floor. Based on that floor, we compute the amount of labor that can be released without changing agricultural wages. We ignore the effects of the floor on profitability except to note that our estimates indicate that increasing average farm scale would raise profitability per acre even if wages stayed fixed.[19] In particular, we combine the local-area IV estimates of the causal effects of land size on labor utilization by land size from Figure 7, the level of per-acre labor expenditures by farm size obtained from the 2007-8 survey data (Appendix Figure C), and estimates of the national distribution of land ownership among rural households, constructed from the National Sample Survey (NSS) in 2007 (Appendix Figure D). The nationally-representative NSS data are used, based on linear interpolation, to construct at intervals of .001 acres from .01 to 20 acres, an estimate of the actual density of farms in India.

[19]This floor is not envisioned as an actual policy but does provide a measure of the extent to which labor could be released without increasing the agricultural wage if opportunities outside of the agricultural sector were to expand and exit was accompanied by land amalgamation. In practice, land consolidation absent a growing industrial or service sector that could absorb these workers would increase farm profitability but would lead to substantial declines in the incomes of agricultural wage workers due to the reduction in labor demand.

Define surplus labor, $SL(\alpha)$, as the amount of labor that would be released via farm expansion to a minimum level of acreage α and surplus farms, $SF(\alpha)$, as the number of separate farms that would be removed at that minimum scale. Thus if $n(a)$ denotes the actual density of farms with area a and $\beta(a)$ is the local-area IV estimate of the effect on labor costs per acre of a unit increase in land among farmers with area a, then

$$SL(\alpha) = -\int_{a_1=0}^{\alpha} n(a_1)a_1 \int_{a_2=a_1}^{\alpha} \beta(a_2)da_2da_1 \text{ and } SF(\alpha) = \Delta n(a) = \int_{a=0}^{\alpha} n(a)(1-\frac{a}{\alpha})da\,.$$

Figure 8 reports the estimates of surplus labor for hypothesized floors of 2, 5, 10, 15 and 20 acres. The figures are expressed, respectively, as the fraction of the total agricultural labor force employed below the farm size floor and of the total agricultural labor force.[20] The latter is indicative of the overall extent to which the given floor would release labor without a fall in wages or farm profitability and the former gives an estimate of the degree of surplus labor among farms of a particular size. The two sets of bars are surprisingly similar, but this reflects the fact that only a small fraction of all labor is employed on large farms, both because large farms are relatively rare and because such farms employ relatively few workers. In particular, if the smallest farms in India were two acres, only 0.4 percent of the total labor force is 'surplus.' This fraction rises, however, to a substantial 22.7% of all workers for a floor of 20 acres. Put another way, if all farms under 20 acres were consolidated to a level that our data suggest yields close to maximal profitability per acre, over one in five workers could be released from agricultural crop production at current wages. This latter figure suggests that the absolute size of surplus labor in Indian agriculture is about 21 million full-time agricultural workers.[21] The decline in the number of separate farm operations is even more dramatic - a two acre floor would lead to a 37% reduction in the number of farms while a twenty-acre minimum would lead to a 88.2% reduction in the number of farms. It clear based on these estimates that there are both too many farms and too much agricultural labor employed in rural India, given available technology and existing labor and credit market conditions.

11. Conclusion

[20]These fractions are based on the cumulative estimated reductions in total labor expenditures divided by total labor costs up to the floor and total overall labor costs.

[21]To estimate the number of workers we have assumed that a full-time worker works 250 days per year and the average wage is 50 rupees per day. Our estimates of the fraction of the labor force that is surplus do not depend on any assumption about days worked.

The focus of surplus labor models is on the relationship between industrialization and the productivity of the labor force in agriculture. A key prediction is that in the early stages of industrialization agriculture neither output nor wages would not decline as laborers exited the agricultural sector. In this paper we suggest that, in contrast to the original surplus labor models, rural labor in India is not in surplus conditional on the distribution of land. However, we show that labor could be released from agriculture without a decline in wages or profitability if there were an increase in the scale of agricultural production. In particular, we develop a model in which labor markets are competitive and there is an effective rental market for machinery, consistent with empirical evidence over the past twenty years that labor and product markets in India, and many other low-income rural areas are well approximated by standard competitive models, in contrast to the assumptions of the original surplus labor framework. Within this context however technical scale economies associated with mechanization and credit market imperfections give rise to hidden or disguised surplus labor arising from unrealized gains from larger-scale operations.

Scale economies as a source of surplus labor has not been considered in the modern development literature in part due to data limitations. Older data sets from India that have been used in many studies do not take into account recent changes in the rural sector in India, including improvements in the scalability of mechanized implements and the rapid rise in rural nonfarm income sources that employ workers from small farms. Few data sets also contain the necessary information to measure profitability and/or to control for cross-household variation in the shadow prices or transaction costs or provide a mechanism for addressing the potential endogeneity of landownership size.

We make use of a newly available data set that provides detailed information on farm production and assets in rural India. These data permit the linking of farmers over an eight year period, and includes information on multiple seasons and across multiple plots in the 2007-08 crop-year. Our empirical results, which exploits the prevalence of household division and land inheritance, indicate that larger farms in contemporary India are more efficient than their smaller counterparts. On farms which experienced a contraction in owned landholdings there is less mechanization, more labor use per acre inclusive of supervisory labor, and, most importantly, lower profitability per acre. Our findings suggest that the greater efficiency of larger farms is partly a scale effect associated with the use of mechanized inputs but also is related to credit market constraints. Larger landowners appear to have an advantage in the credit market. They face lower credit costs due to superior collateral and are better protected from income shocks. Consistent with this we find that larger farms use variable inputs more efficiently, are better mechanized, and are more insulated against fluctuations in profits associated with weather variation in terms of input efficiency.

Our quantitative assessment of the magnitudes of these effects using the current size distribution of land in India as a basis for analysis suggests that the amount of potential or hidden surplus labor in India is large - given the scale economies we identify and the large number of farms with less than 2 acres, amalgamation to minimum-scaled, 20-acre farms can free up 22 percent of the agricultural labor force with no decline in wage rates. Of course, we do not envision this as a policy but instead as a rough guide to the amount of surplus labor that is available at the given wage if existing cultural or political barriers to amalgamation were removed and there were sufficient growth in industrial and service employment to absorb these workers.[22] We also do not claim that increasing the scale of agriculture would, by releasing 'surplus' labor, lower the costs of expanding the nonfarm sector in all contexts or that the classic surplus labor models are no longer relevant in all countries. Evidence from contemporary India where markets are reasonably competitive and a large farm is one with as few as ten acres, may provide little insight into the relative profitability of larger farms or the costs to the agricultural sector from industrialization in Latin America where land inequality is far greater than that in India, or in parts of Africa where labor markets and product markets are far less complete.

These findings suggest that there are barriers to land amalgamation. While a full investigation of the sources of these barriers is beyond the scope of this paper, we note that in contrast to labor markets, there is ample evidence of a thin market for land and tenancy in India. Certainly there are a host of possible reasons for this thinness. In this paper we have highlighted the importance of contiguous land as a friction that complicates amalgamation, as suggested by the fact that a large fraction of current tenancy and land transfers occur among family members. It

[22]Regardless of the efficiency consequences associated with consolidation, any policy-maker promoting amalgamation should consider possible consequences for rural poverty of an increase in landlessness among poor households. In the presence of thin land and capital markets and weak political institutions and the absence of a growing industrial sector, it is not a given that small landowners would be appropriately compensated and have adequate means of productively using their labor and capital in the event that such an amalgamation were to take place.

is interesting to note that while our model indicates that there may be a substantial individual benefit to selling of land by small farmers given existing wages, it also suggests the presence of a negative pecuniary externality associated with the aggregate selling of land by individuals who participate in the local labor market: aggregate land sales by small farmers will lead to reduced labor demand in agriculture. In the absence of growing alternatives for employment outside of agriculture, i.e., in manufacturing, small landholders may have a collective incentive to discourage land sales using cultural sanctions and/or imposing constraints through the political process. This set of ideas about possible constraints on land transfers and tenancy provide a productive avenue for exploration in future research.

References

Allen, R.C. (1988). The Growth of Labor Productivity in Early Modern English Agriculture. *Explorations in Economic History* 25, 117-146.

Assunção, J.J., and L.H.B. Braido. (2007). Testing Household-Specific Explanations for the Inverse Productivity Relationship. *American Journal of Agricultural Economics*, 89 (4), 980-990.

Barrett, C.B., M.F. Bellemare and J.Y. Hou (2010). Reconsidering Conventional Explanations of the Inverse Productivity–Size Relationship. *World Development*, Vol. 38, No. 1.

Benjamin, D. (1995). Can Unobserved Land Quality Explain the Inverse Productivity Relationship? *Journal of Development Economics*, 46 (1), 51-84.

Berry, R.A., and W.R. Cline (1979). *Agrarian Structure and Productivity in Developing Countries*. Baltimore: Johns Hopkins University Press.

Bhalla, S.S., and P. Roy. (1988). Mis-Specification in Farm Productivity Analysis: The Role of Land Quality. *Oxford Economic Papers*, 40, 55-73.

Bhalla, S.S. and T. Das (2005). Pre- and Post-Reform India: A Revised Look at Employment, Wages and Inequality. *India Policy Forum 2005* 2, 183-233.

Carter, M.R. (1984). Identification of the Inverse Relationship between Farm Size and Productivity: An Empirical Analysis of Peasant Agricultural Production. *Oxford Economic Papers*, 36, 131-145.

de Janvry, A., M. Fafchamps and É. Sadoulet. (1991). Peasant Household Behavior with Missing Markets: Some Paradoxes Explained. *Economic Journal, 101* (409), 1400-1417.

Feder, G. (1985). The Relation between Farm Size and Farm Productivity: The Role of Family Labor, Supervision and Credit Constraints. *Journal of Development Economics, 18* (2-3), 297-313.

Fei, J.C.H. and G. Ranis (1964). *Development of the Labor Surplus Economy: Theory and Policy*. Richard D. Irwin, Homewood, Ill.

Foster, A.D. and M.R. Rosenzweig (2002). Household Division, Inequality and Rural Economic. Growth", *Review of Economic Studies*. 69(4): 839-869

Lamb, R.L. (2003). Inverse Productivity: Land Quality, Labor Markets, and Measurement Error. *Journal of Development Economics, 71* (1), 71-95.

Lanjouw, P and R. Murgai (2009). Poverty Decline, Agricultural wages, and Nonfarm Employment in Rural India: 1983–2004. *Agricultural Economics* 40(2): 243-263.

Lewis, A. (1954). Development with Unlimited Supplies of Labor. *The Manchester School* 2:139-191.

Rosenzweig, M.R. (1988) Labor Markets in Low-income Countries. *Handbook of Development Economics, Volume I*, H. Chenery and T.N. Srinivasan, eds.,Elsevier Science Publishers B. V., 713-762.

Rosenzweig, M.R. and H.P. Binswanger. (1993). Wealth, Weather Risk and the Profitability of Agricultural Investment. *Economic Journal* 103(1): 56-78.

Sen, A.K. (1966). Peasants and Dualism with or without Surplus Labor. *Journal of Political Economy, 74* (5), 425-450.

Figure 1. Mechanization and Owned Landholdings (Acres), 2007-2008

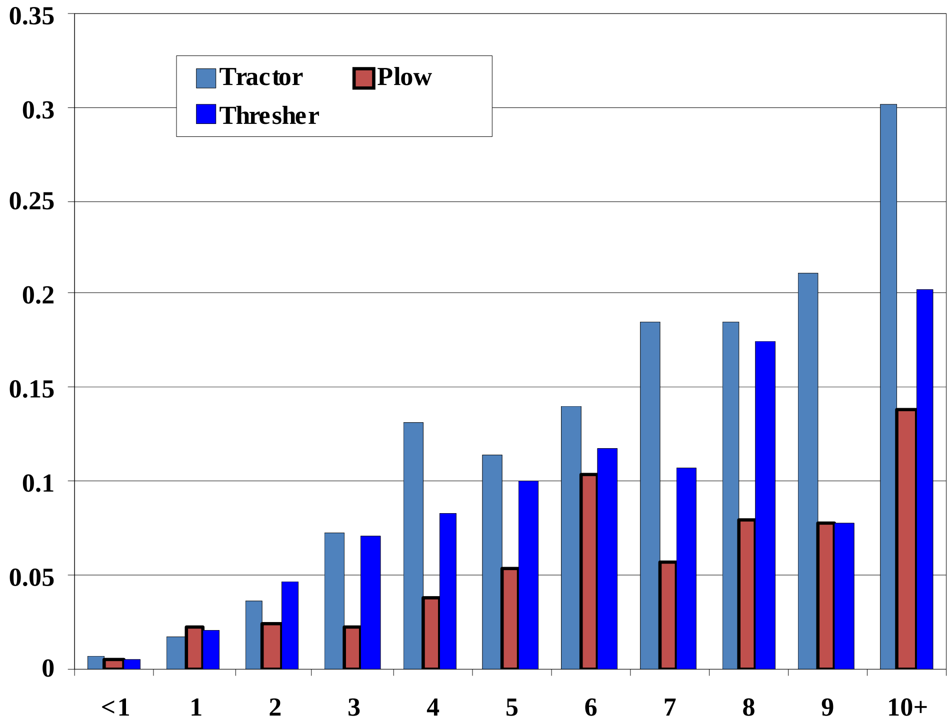


Figure 2. Per-Acre Profits, Per-Acre Profits less Supervision Costs, and Per-Acre Total Labor Costs, by Owned Landholdings, 2007-8

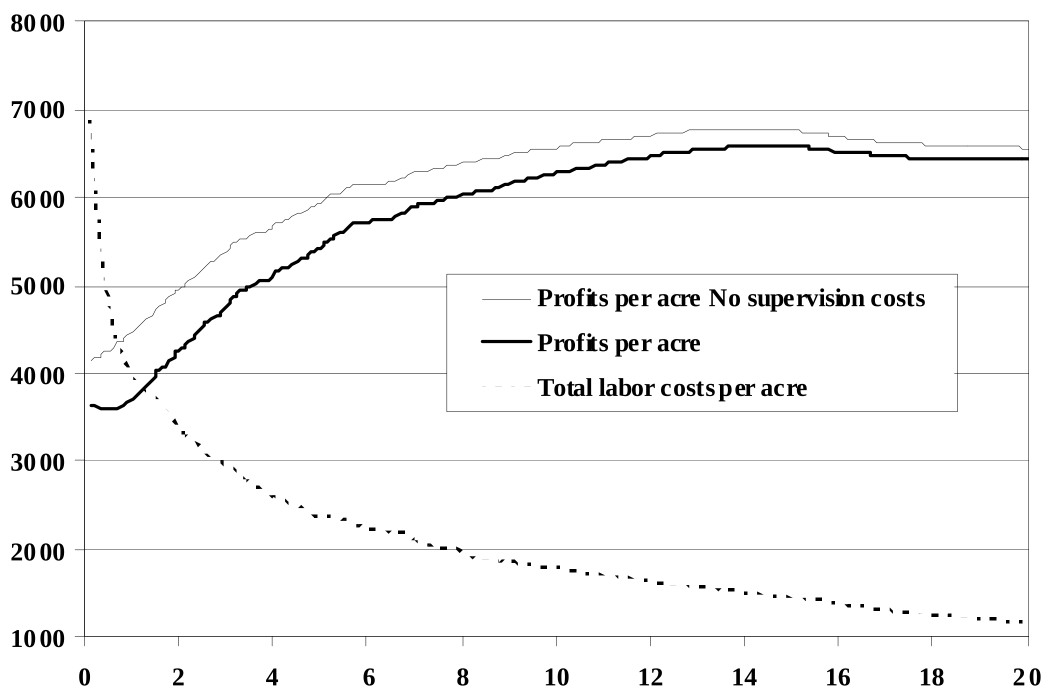


Figure 3. Locally-weighted Within-Farmer and Within-Season Estimates: The Effects of Plot-Specific Fertilizer on Plot-Specific Profits per Acre (one sd Confidence Bounds), by Landholding Size

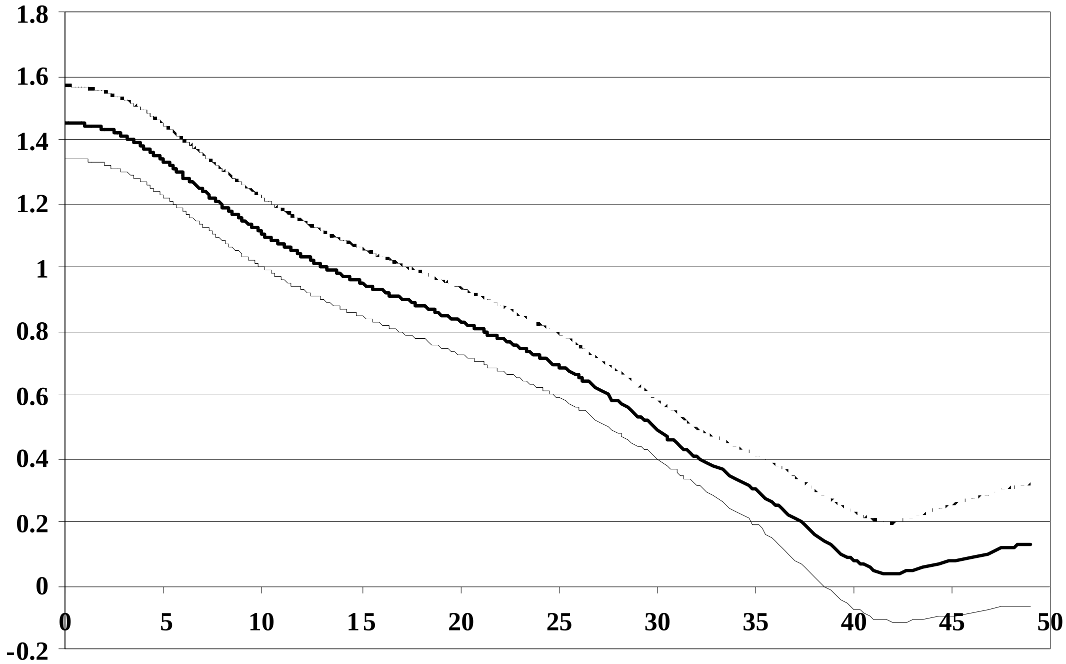


Figure 4. Locally-weighted FE-IV Estimates of the Effects of Owned Landholdings on Profits per Acre, Net and Gross of Farm Equipment Owned, by Landholding Size

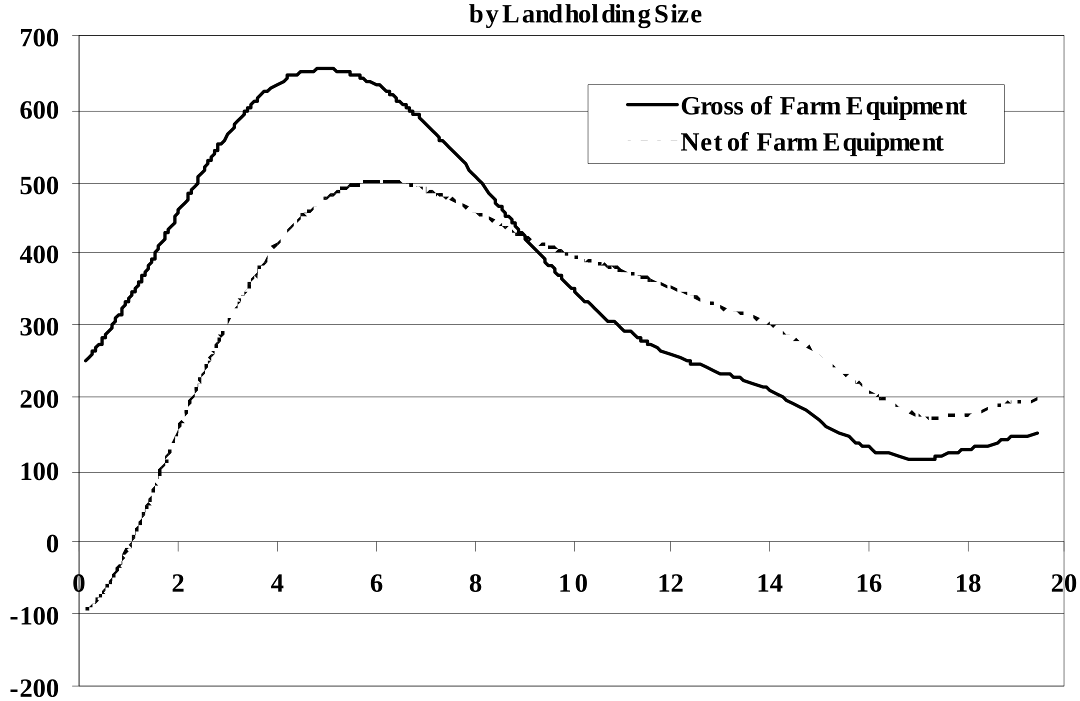


Figure 5. Locally-weighted FE-IV Estimates of the Returns to Capital Equipment Value (one sd Confidence Bounds), by Landholding Size

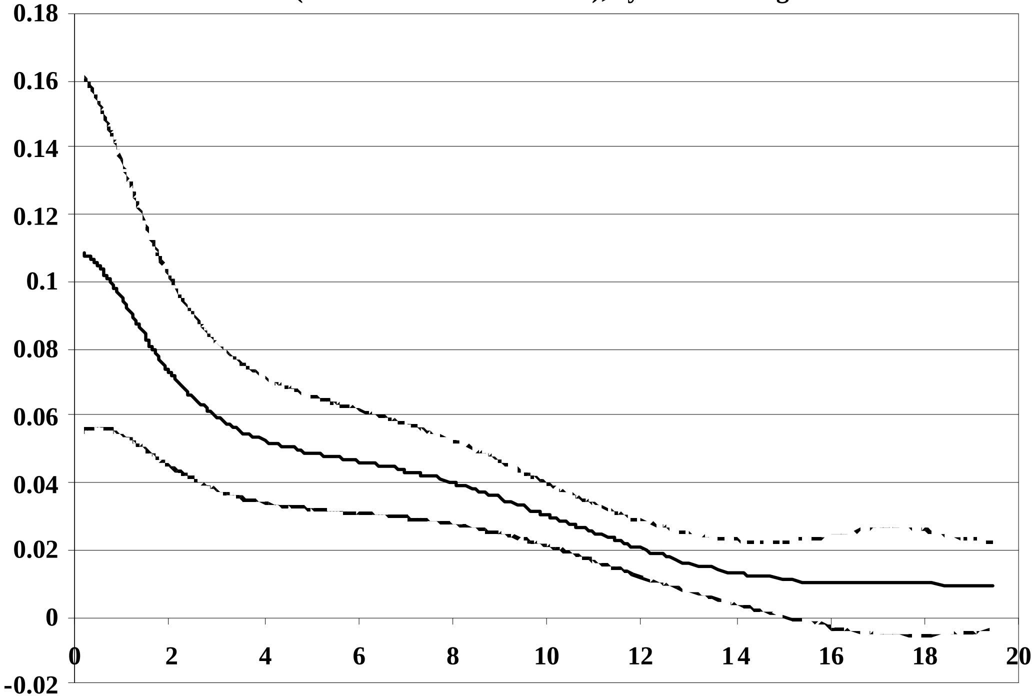


Figure 6. Within-Village Relationship Between the Probability of Leasing In and Leasing out Land, by Ownership Holdings, 2006 (N=119,349)

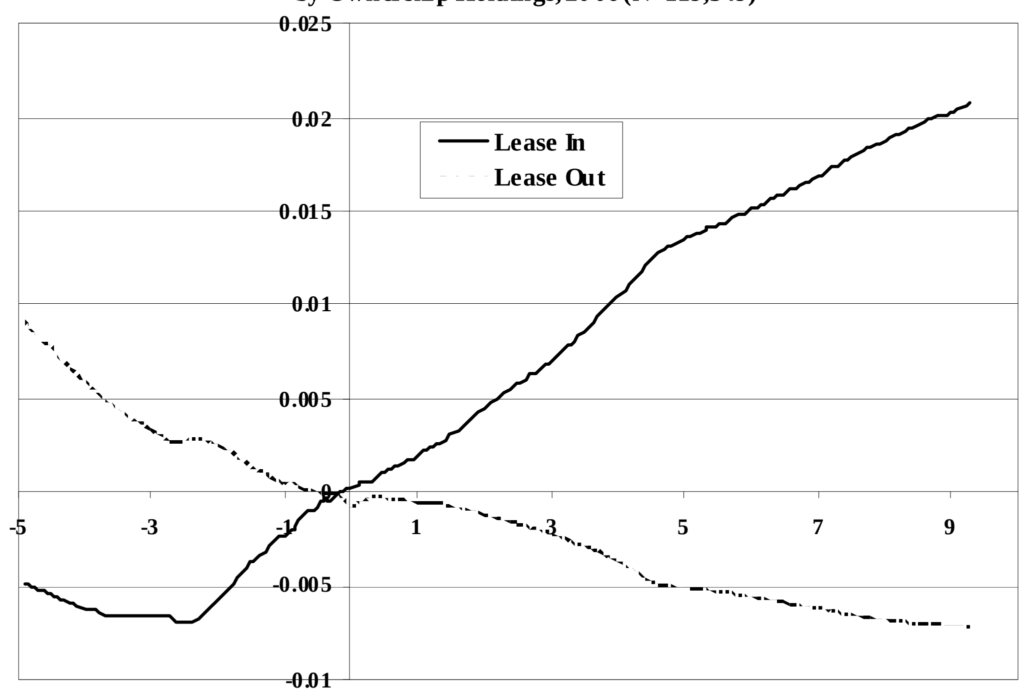


Figure 7. Locally-weighted FE-IV Estimates of the Effect of Landholding Size on Per-Acre Total Labor Costs, by Landholding Size

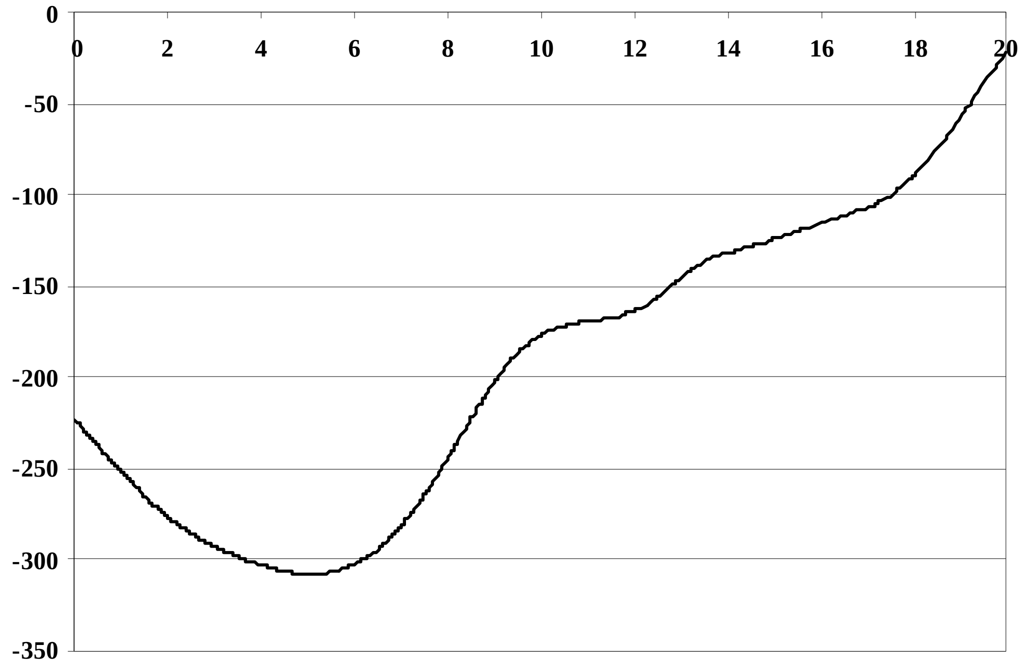


Figure 8. Fraction of the Agricultural Labor Force that is Surplus, by Minimum Acreage Farm Size Floor

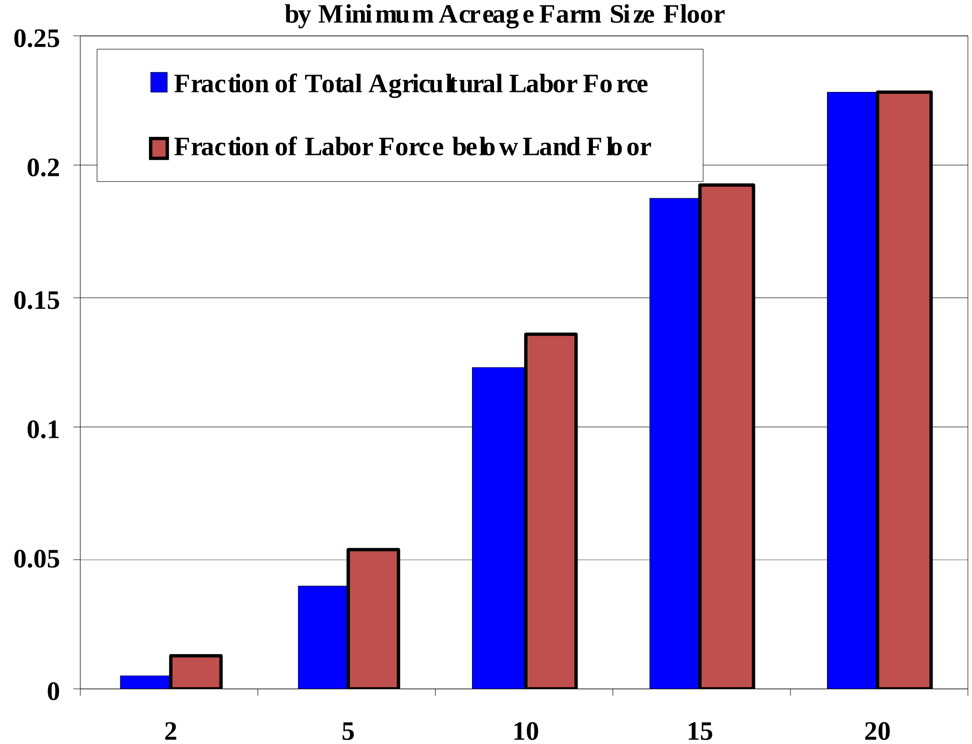


Table 1. Within-Village and Within-Farmer and Season Plot Level Estimates (2007-2008):
Effects of the Use of Hired and Family Labor on Supervision Costs, by Estimation Procedure

Estimation procedure:	Village Fixed-Effects[a]		Farmer-Season Fixed-Effects	
Hired labor costs	.0402 (3.05)	.0383 (2.94)	.0387 (3.46)	.0399 (3.37)
Family labor costs, less supervision time	.134 (4.11)	.140 (4.23)	.0410 (1.57)	.0375 (1.44)
Plot area	.00407 (1.91)	.00409 (1.92)	3.58 (1.51)	3.35 (1.37)
Owned landholdings	2.74 (0.62)	2.80 (0.62)	-	-
Plot characteristics included[b]	N	Y	N	Y
Number of observations	18,484	18,201	18,484	18,201
Number of farmer-seasons	8,685	8,587	8,685	8,587

Absolute value of asymptotic t-ratios in parentheses. [a]Specification includes season dummy variables; clustered t-ratios at the farm level. [a]Plot characteristics include measures of depth, salinity, percolation and drainage; five soil colors (red, black, grey, yellow, brown, off-white); five soil types (gravel, sandy, loam, clay, and hard clay), and distance from the household residence.

Table 2. Within-Farmer, Plot-Level Estimates Across Three Seasons (2007-8):
Effects of Plot Size on Plot-Specific Profits, Labor Costs, and Fertilizer Use per Acre and Use of Tractor Services

	Profits per Acre		Any Tractor Services Used		Total Labor Costs per Acre	
Plot area	145.4	157.0	.00403	.00404	-107.2	-106.9
	(2.34)	(2.51)	(1.99)	(1.98)	(2.74)	(2.73)
Area of other plots	118.7	130.8	.00333	.00351	-55.3	-55.4
	(1.95)	(2.14)	(1.77)	(1.76)	(1.45)	(1.45)
Total number of plots	482.3	473.8	-.0333	-.0330	123.2	123.2
	(3.15)	(3.06)	(6.68)	(6.53)	(1.27)	(1.27)
Include soil characteristics?[a]	N	Y	N	Y	N	Y
Number of plot observations	14,290	14,290	14,290	14,290	14,290	14,290
Number of farmers	4,130	4,130	4,130	4,130	4,130	4,130

Absolute value of asymptotic t-ratios in parentheses. [a]Soil characteristics include depth, salinity, percolation and drainage; five soil colors (red, black, grey, yellow, brown, off-white); five soil types (gravel, sandy, loam, clay, and hard clay). All specifications include season*state dummy variables, plot distance.

Table 3. Within-Farmer, Within-Season Plot-Level Estimates (2007-8):
Effects of Plot-Specific Fertilizer Use on Plot-Specific Profits, by Owned Landholdings

Owned landholdings	< 4 acres		4-10 acres		10+ acres	
Fertilizer use this season	1.49	1.46	3.23	3.24	.182	.166
	(3.73)	(3.89)	(2.25)	(2.28)	(0.20)	(0.19)
Plot area	29.9	30.4	901.9	747.4	39.3	1.16
	(1.02)	(0.97)	(3.03)	(2.45)	(0.48)	(0.02)
Include soil characteristics?	N	Y	N	Y	N	Y
Number of plot observations	4,008	4,008	1,935	1,935	1,173	1,173
Number of farmers	1,939	1,939	851	851	464	464

Absolute value of asymptotic t-ratios in parentheses clustered at the village level. All specifications include plot distance, fertilizer used in the prior period.

Table 4. Panel Data Estimates (1999-2008): Effects of Own Landholdings and Own Farm Equipment on Profits per Acre, by Estimation Procedure

Estimation procedure:	Village Fixed-Effects[a]		Farmer Fixed-Effects		Farmer Fixed-Effects IV[b]	
Owned landholdings	13.1 (2.77)	13.1 (2.80)	8.35 (0.48)	3.53 (0.20)	369.6 (1.98)	256.9 (1.45)
Value of farm equipment	-	.00746 (1.16)	-	.0114 (3.26)	-	.0351 (2.02)
Number of observations	3,994	3,994	3,524	3,524	3,524	3,524
Number of farmers	2,138	2,138	1,749	1,749	1,749	1,749
Kleinberger-Paap underidentification test statistic χ^2(df), *p*-value					(5) 12.6, .0271	(7) 13.6, .0509
Hansen J overidentification test statistic χ^2(df), *p*-value					(4) 0.44 .979	(6) 5.59 .471

Absolute value of asymptotic t-ratios in parentheses. [a]Specification includes year=2008 dummy; clustered t-ratios at the household level. [b]Instruments include land inherited after 1999, assets inherited after 1999, whether the current head has brothers, the standard deviation of the schooling of inheritance claimants, the father's age in 1999, whether the father is co-resident in 1999, and owned asset values in 1999.

Table 5. Retrospective Panel Data Estimates (2008): Effects of Own Landholdings and Own Farm Equipment on Investment in Farm Equipment and Equipment Rental, by Estimation Procedure

Dependent variable	Equipment Investment			Equipment Hire Expenditure		
Estimation procedure	FE-Farmer	FE-Farmer IV[b]	FE-Farmer IV[b]	FE-Farmer	FE-Farmer IV[b]	FE-Farmer IV[b]
Owned landholdings	16.3 (0.05)	663.8 (2.15)	-400.7 (1.06)	122.7 (1.69)	196.3 (2.01)	201.2 (1.90)
Landholdings x bank	-	-	1654 (2.50)	-	-	-6.29 (0.04)
Value of owned farm equipment	-.0843 (1.30)	-.909 (8.67)	-.906 (8.54)	-.0187 (3.51)	-.0500 (1.16)	-.0500 (1.16)
Bank within 10 Km	3524 (2.27)	1820 (0.61)	-2920 (0.84)	-.319.8 (0.94)	279.6 (0.53)	-258.0 (0.29)
Number of farmers	3,522	3,522	3,522	1,833	1,833	1,833

Absolute value of asymptotic t-ratios in parentheses. [a]Specification includes year=2008 dummy; clustered t-ratios at the village level. [b]Instruments include land inherited after 1999, assets inherited after 1999, and the presence of a bank within 10 km in 1999.

Table 6. Within-Plot Estimates Across Three Seasons (2007-8):
Effects of Previous-Period *Farm-Level* Profit Shocks on *Plot-Level* Current Fertilizer Value per Acre, by Owned Landholding Size

Farm size:	Owned Landholdings<4	Owned Landholdings>=4, <10	Owned Landholdings>=10
Farm profits per acre, all other plots, previous season	.00706 (1.32)	.00687 (0.98)	.00350 (0.48)
Farm profits per acre, this plot, previous season	.00182 (0.34)	.00655 (0.82)	.00792 (1.02)
Fertilizer use, this plot, previous season (value per acre)	-.310 (3.97)	-.485 (4.44)	-.543 (5.25)
Number of plot observations	6,068	3,258	1,919
Number of farmers	1,351	678	311

Absolute value of asymptotic t-ratios in parentheses clustered at the farm level. Specifications include season*village dummy variables.

Table 7. Within-Plot Estimates Across Three Seasons (2007-8):
Effects of Previous-Period *Farm-Level* Profit Shocks on *Plot-Level* Current Profits per Acre, by Owned Landholding Size

Farm size:	Owned Landholdings<4	Owned Landholdings>=4, <10	Owned Landholdings>=10
Farm profits per acre, all other plots, previous season	.140 (2.04)	.0628 (0.67)	.0366 (0.17)
Farm profits per acre, this plot, previous season	-.456 (5.60)	-.504 (4.78)	-.540 (2.88)
Fertilizer use, this plot, previous season (value per acre)	1.54 (2.61)	.789 (1.97)	1.87 (1.22)
Number of plot observations	6,068	3,258	1,919
Number of farmers	1,351	678	311

Absolute value of asymptotic t-ratios in parentheses clustered at the farm level. Specifications include season*village dummy variables.

Appendix A

Scale Effects on Per-acre Profits and Inputs

We may write the solution to the cost minimization problem as

(A1) $$\tilde{c}(e,a) = c(a)e$$

and the conditional factor demands as, for example,

(A2) $$\tilde{k}^c(e,a) = k^c(a)e$$

Implicit differentiation yields

(A3) $$k^c{}'(a) = \frac{k^c(a)\phi'(a)(-2(1-\delta)+\omega_l(2-\nu)(l_m^c(a)l(ls/lm)))^\delta}{\phi(a)(1-\delta)},$$

which is positive for δ sufficiently close to one. That is, for a given work level *e*, when machinery is sufficiently substitutable for labor the number of machines, of increasing capacity, increase as scale increases. The ambiguity in terms of quantity of machinery arises for lower δ even when machinery and labor are substitutes because higher-capacity machinery can produce more work in less time.

Supervisory, manual labor, and the shadow price of work, for a given level of work, all decline with land area because an increasing share of work is supplied by machinery

(A4) $$l_s^c{}'(a) = -\frac{l_s^c(a)\phi'(a)\phi(a)^{2\delta-1}\kappa_1{}^\delta k^c(a)^\delta(2-\nu)}{(1-\delta)} < 0$$

(A5) $$l_m^c{}'(a) = -\frac{l_m^c(a)\phi'(a)\phi(a)^{2\delta-1}\kappa_1{}^\delta k^c(a)^\delta(2-\nu)}{(1-\delta)} < 0$$

(A6) $$c'(a) = -c(a)\phi'(a)\phi(a)^{2\delta-1}\kappa_1{}^\delta k^c(a)^\delta(2-\nu) < 0,$$

where $\kappa_1 = (1-\nu)/(2-\nu)^2$.

To actually determine how much work is done and the total use of labor and machinery we now embed the cost-minimization problem in a profit-maximizing one. In particular, let

(A7) $$\pi(a) = \max a\cdot g(e/a, f/a) - c(a)\cdot e - p_f f$$

or, letting the superscript * denote per-acre quantities:

(A8) $$\pi^*(a) = \max g(e^*, f^*) - c(a)\cdot e^* - p_f f^*$$

The envelope condition implies

(A9) $$\frac{d\pi^*}{da} = -c'(a)e^* > 0 .$$

Profits per acre increase with area, because the cost of work per unit area decreases. Larger operations are more profitable on a per-acre basis. Similarly, larger operational holdings will use inputs more intensively, as per-acre work increases in unit area

(A10) $$\frac{de^*}{da} = c'(a)\frac{g_{ee}}{g_{ee}g_{ff} - g_{ef}^{\,2}} = c'(a)\varepsilon_{ec}\frac{e^*}{c(a)} > 0$$

and fertilizer per unit area increases in area

(A11) $$\frac{df^*}{da} = -c'(a)\frac{g_{ef}}{g_{ee}g_{ff} - g_{ef}^{\,2}} = c'(a)\varepsilon_{fc}\frac{f^*}{c(a)}$$

if fertilizer and work are complementary, where ε_{ec} is the own price elasticity of demand for work and ε_{fc} is the fertilizer-work cross-price elasticity.

The number of machines *k* per unit area will be increasing in area, for δ sufficiently close to 1, because (i) there will be an overall expansion of work (A9) and (ii) *k* is increasing in total work. In particular,

(A12) $$\frac{dk^*}{da} = \frac{k^*}{k^c}\left(l_m^c{}'(a)(1-\delta)\frac{k^c}{l_m^c}\varepsilon_{ec} + k^c{}'(a)\right).$$

Whether total expenditures on machinery will rise for $\delta < 1$ as land size increases depends on whether the pricing of machinery is sufficiently elastic to capacity. Regardless of whether the number of machines used per unit area increases or decreases, whether a farmer owns a machine of a given capacity or greater is rising in area as indicated by (5).

Appendix B

Testing for Perfect Capital Markets using Profits that Ignore Interest Costs

The profit function in terms of estimated profits, profits that do not include interest costs, is given by

(B1) $$\hat{\pi}^* = g(e^*, f^*) - c(a)e^* - p_f f^*,$$

where the inputs are determined by programming problem (21) and * indicates per-acre. In this case we have

(B2) $$\frac{d\hat{\pi}^*}{da} = (\rho_b - 1)(c(a)\frac{de^*}{da} + p_f \frac{df^*}{da}) - c'(a)e^*$$

where $\rho_b > 1$ and the second term in parentheses is positive. Estimated profits also increase with owned landholdings. Comparing (A2) to (14) indicates that the gradient in estimated profits, as with that of true profits, is steeper than would be the case in the absence of credit market effects. In the case in which there are no technical scale economies associated with machinery so $c'(a) = 0$, (14) would be zero but (22) would be positive if $\rho_a < 0$ and (A2) would be positive if $\rho_{ab} < 0$.[23]

A direct test of credit market constraints can be obtained by examining the returns to owned capital assets using true or estimated profits. The marginal return to capital in terms of true profits is given by

(B3) $$\frac{d\pi^*}{do^*} = -\rho_b + (1 + r_0) = r(a, b^*) - r_0 + \frac{dr(a, b^*)}{db^*} b^* > 0,$$

while the marginal return to estimated profits is

(B4) $$\frac{d\hat{\pi}^*}{do^*} = (\rho_b - 1)(c(a)\frac{de^*}{do^*} + p_f \frac{df^*}{do^*}).$$

The observed marginal returns to capital assets in the presence of credit constraints evidently differ depending on how profits are computed. However, it is easily established that when $r(a, b^*) = r_0$, that is when borrowing costs are independent of land ownership and equal to the

[23] These conditions coincide in the case in which the interest rate is proportional to borrowing per acre.

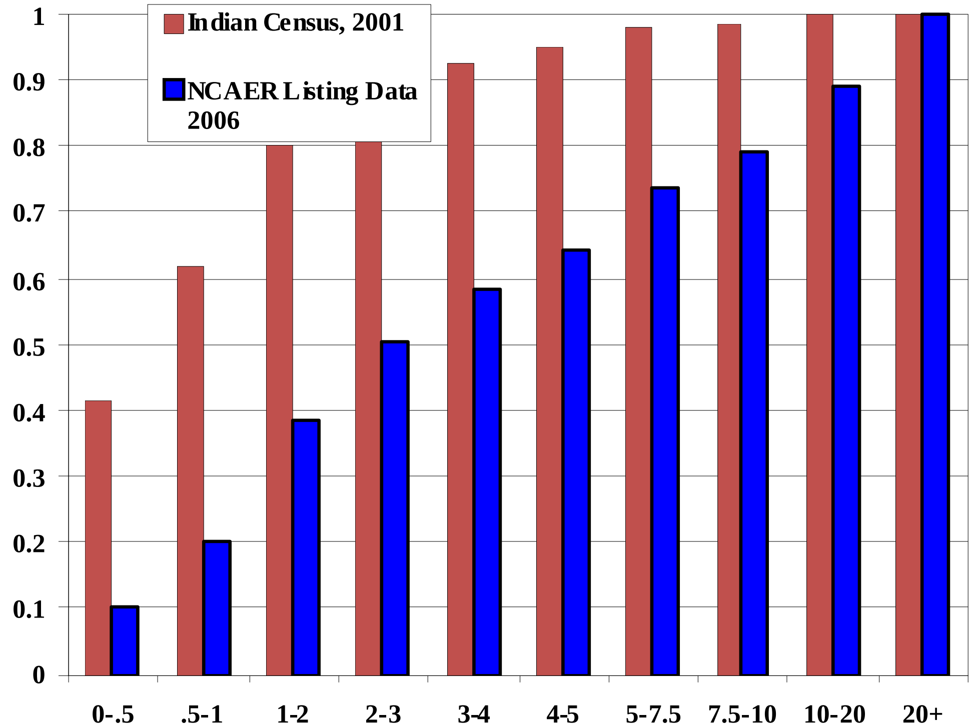
returns on savings, the marginal return to capital assets is zero for either measure of profits. This is because variation in owned machinery at the margin has no effects on the use of production inputs,

(B5) $$\frac{de^*}{do^*} = \frac{df^*}{do^*} = 0 \ .$$

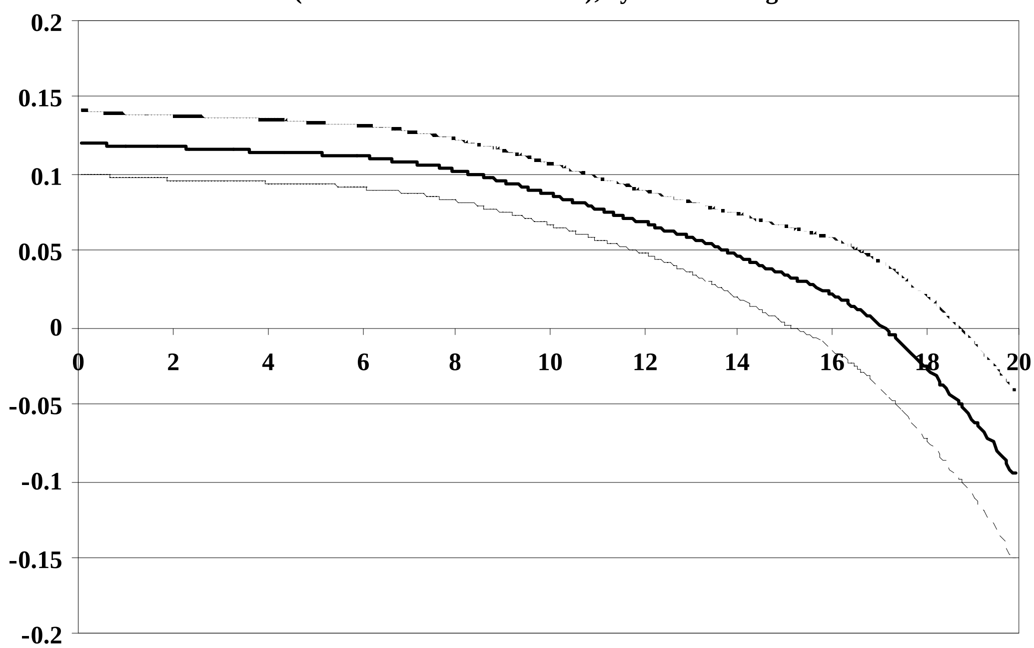
Therefore, the finding that there is a non-zero return, in terms of estimated profits, to owned capital assets would reject the hypothesis of perfect capital markets. The finding, moreover, that empirical profits rises less steeply with landholdings when credit costs are held constant than when they are not, (A2) compared with (14), would establish further that the lower per-acre profitability of smaller compared with larger landowners is due to disadvantages in the credit market, as depicted in (19).

Appendix Figures

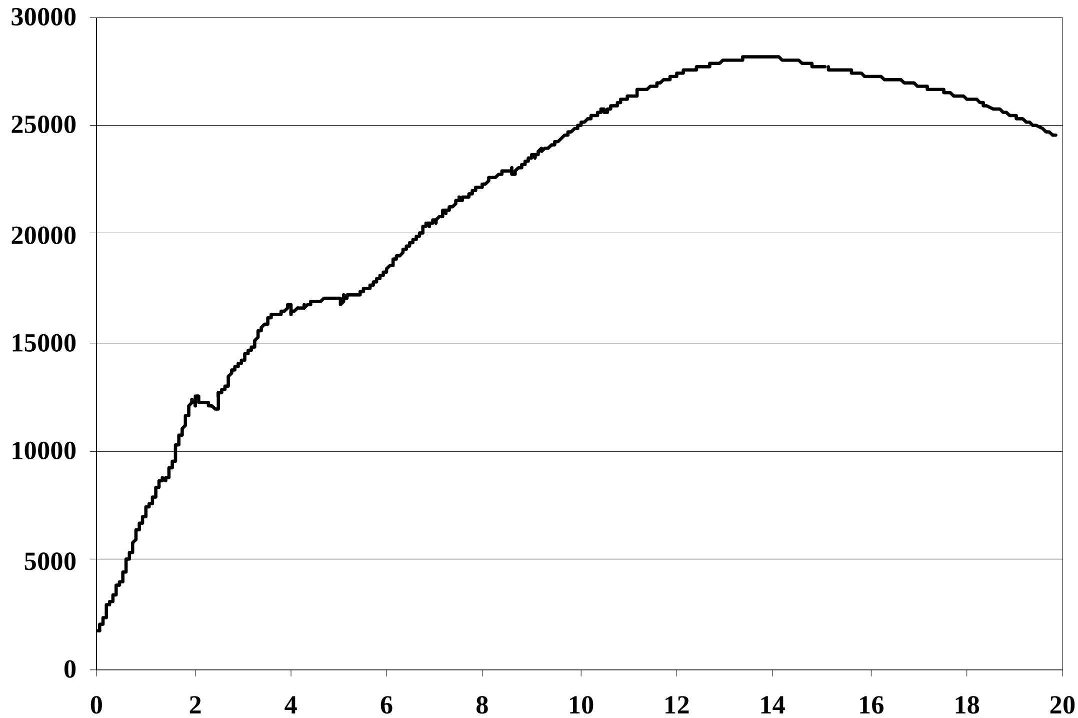
Appendix Figure A. Cumulative Distribution of Owned Landholdings (Acres), by Data Source



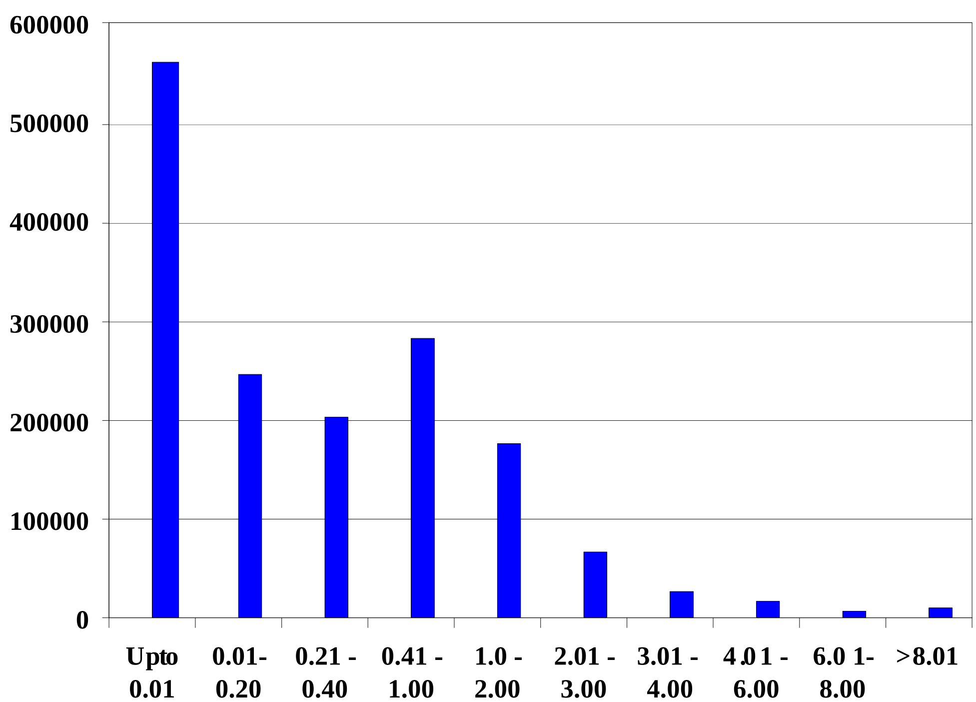
Appendix Figure B. Locally-weighted Within-Plot Estimates:
The Effects of Lagged Farm Profits on Plot-Specific Profits per Acre
(one sd Confidence Bounds), by Landholding Size



Appendix Figure C. Total Labor Costs and Landholding Size



Appendix Figure D. The Distribution of Owned Landholdings in India (July 2006 – June 2007, NSS): Number of Households (x1,000) in Intervals of Hectares



Appendix Table A
First-Stage Panel Data (1999-2008) Farmer FE Estimates: Owned Landholdings and Value of Farm Equipment

Variable:	Own Landholdings (acres)	Farm equipment x 10^{-3}
Inherited land (acres) after 1999	.187	.833
	(2.70)	(1.12)
Value of owned inherited mechanized assets in 1999 x 10^{-3}	-.00372	.111
	(0.27)	(1.35)
Value of owned inherited non-mechanized assets in 1999 x 10^{-3}	-.0753	2.85
	(2.15)	(3.22)
Value of assets inherited after 1999 x 10^{-3}	.200	-8.02
	(0.31)	(1.11)
Standard deviation of the schooling of family claimants in 1999	-.0870	1970
	(1.23)	(2.57)
Father's age in 1999	-.0127	-.124
	(1.07)	(0.91)
Father lives in separate household in 1999	1.09	8931
	(2.28)	(1.10)
Whether respondent has brothers	-.167	1789
	(0.35)	(0.44)
Number of observations	3,994	3,524
Number of farmers	1,749	1,749
Anderson-Rubin Wald joint test of weak instruments $\chi^2(8)$,	22.81	
p-value	.0036	

Absolute value of asymptotic t-ratios in parentheses clustered at the household level.

Appendix Table B
Retrospective Panel Data Estimates (2008) First Stage Fixed-Effects Farmer Estimates: Effects of Own Landholdings and Own Farm Equipment on Investment in Farm Equipment and Equipment Rental

Dependent variable/ Instrument	Owned Landholdings	Own Farm Equipment	Bank < 10 km of the Village	Own Farm Equipment x Bank Proximity
Inherited landholdings between 1999 and 2008	.938 (18.6)	-142.6 (0.40)	-.0191 (3.65)	.0432 (0.57)
Inherited farm assets between 1999 and 2008 x 10^{-3}	-.00105 (2.61)	.546 (5.31)	.000973 (2.48)	.00229 (1.78)
Bank within 10 km of the village in 1999	-.118 (0.92)	-841.8 (0.42)	-.918 (10.7)	-3.53 (6.92)
Inherited landholdings x bank proximity	.04365 (0.85)	1499 (2.22)	.0182 (2.92)	.904 (10.9)
Number of farmers	3,522	3,522	3,522	3,522

Absolute value of asymptotic t-ratios in parentheses clustered at the farm level. Specifications include season*village dummy variables.