Barriers to Farm Profitability in India: Mechanization, Scale and Credit Markets

Andrew D. Foster and Mark R. Rosenzweig

September 2010
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 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. Figure 1 provides the cumulative distribution of land ownership, based on the Indian Census of 2001. 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 242 villages in 17 of the major states of India, which we describe and employ extensively below. Figure 2 plots 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, supervisory labor, and own equipment use. By this measure, which
does not take into account the likelihood that larger farmers have lower credit costs,
landownership and farm productivity are strongly positively associated. Figure 3 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 about 12 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. Figure 4 shows that not only do per-acre supervision costs decline with
owned farm size, but above 12 acres, total supervision costs decline. These patterns appears 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
Indeed, the data indicate that total labor costs per acre monotonically fall with ownership
holdings, as seen in Figure 5, which plots total labor costs per acre by owned land size.

Of course, Figures 3 through 5 merely describe associations between scale, labor use and
profitability. It is possible that within India larger farms are located where land is higher quality,
where farmers are better-educated, where credit markets operate more effectively, or where
agricultural conditions generally are more favorable to agriculture. Moreover, land holdings 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 and land ownership and most have examined only particular market constraints and
not the interactions between them in terms of their affects on farm productivity.

¹This presumption also presupposes that family members do not require supervision to
work hard. We show that both hired and family labor require the same amount of supervision
time.
In this paper, we examine theoretically and empirically whether farm scale and lack of mechanization are important proximate and causal barriers to farm productivity and profitability, with particular attention to both the problem of eliciting labor effort and the role of credit markets in an environment with stochastic output. We look at this issue in India, where property rights are reasonably well-established, and where there exists labor, credit and input rental markets, inclusive of those for land and mechanized inputs. We also examine the barriers to mechanization. In order to illuminate the role of returns to scale in mechanization and access to credit in a relatively tractable structure we first develop a model in which there are constant returns to scale in land, total work (work done using labor and/or equipment) and other variable inputs, and allow the level and composition of work to be determined by the relative productivity and costs of different sources of work. Farmers are freely able to rent in or rent out capital equipment and face constant-price variable inputs; however, labor effort depends on the amount of supervision. The model provides conditions under which larger farmers will be more efficient and use more machinery while small farmers will use larger amounts of labor and be less efficient.

We then augment the model in two ways. First, we incorporate credit market imperfections, allowing for differences in access to credit by small and large farmers based on their ownership of land. This augmented model provides predictions on how the returns to capital and variable inputs and the extent of mechanization (owned equipment) vary with land ownership. Second, we introduce risk and dynamics, incorporating productivity shocks, savings, and the persistence of soil nutrients across seasons, to identify how both the ownership of mechanized assets and land affect input returns and profitability in a risky environment.

The model departs from the traditional literature that focuses on scale issues and/or mechanization in that it distinguishes between the capacity and quantity of mechanized implements and builds in the realistic property that higher-capacity mechanized implements require more physical space. For example, a row-crop cultivator that handles eight rows at a time will be approximately four times as productive as a two-row cultivator but will need a greater

\[^2\text{Our last round of data indicate that land records were computerized in 80\% of the villages.}\]
area at the end of each row to turn around. This approach to modeling the relationship between the scale of operation, mechanization and profitability differs in particular from a more conventional approach based on investment indivisibilities and, implicitly, fixed-capacity machinery. For example in the presence of indivisibilities a farmer may wish to purchase a given machine only if he plans to use that item often enough, thus creating a relationship between the level of use and the productivity of machinery given area. But there are problems with this alternative approach. At least in its most simple form, an indivisible machinery model does not reconcile easily with evidence we present below that there are scale economies across plots for a given farmer at a given time and that investment in machinery tends to rise with scale rather than jump abruptly and then stay relatively stable. Moreover, machinery is indeed scalable (e.g., tractors have different horse power, row-cultivators vary in terms of number of rows). There is also significant advantage in terms of tractability of working with a model in which scale economies only arise through land area.

Besides providing a coherent framework for understanding the interactions among the size of owned landholdings, ownership of mechanized inputs, credit and labor market imperfections, and agricultural efficiency, the model also provides tests that enable identification of the distinct roles of technical scale economies and credit barriers in shaping the relationship between assets returns and per-acre efficiency by own land size. In the absence of a feasible way of experimentally varying ownership holdings or farm scale, key to the empirical 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, as we show, can identify pure scale effects, net of measured plot

---

3 Note also that to the extent that there is a loss at the end of each row one will lose a certain amount of space per row but will not lose space based on the number of rows. In this case what would matter in terms of scale economies is the difference between capacity and the length of a row, which in the case of a square plot would be the square root of area.
One shortcoming of our methods is that, because they impound all time-invariant characteristics of farmers into a fixed effect, we cannot identify whether, for example, the low level of schooling of farmers in India is also a barrier to mechanization and farm efficiency. 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 net of all differences in plot and farm characteristics that do not time vary. To obtain causal effects of landownership on profitability and investments, we exploit the fact in the nine-year period 1999-2008, a fraction of 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.\(^4\)

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 change 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 ownership helps overcome credit constraints on both investment and variable input use. 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 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

\(^4\)One shortcoming of our methods is that, because they impound all time-invariant characteristics of farmers into a fixed effect, we cannot identify whether, for example, the low level of schooling of farmers in India is also a barrier to mechanization and farm efficiency.
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. This suggests that industrialization may not only augment economic growth but also raise agricultural productivity to the extent that the exit of people from agriculture to industry is accompanied by land consolidation, as those exiting sell their land. Any growth-augmenting agenda that has as its aim the achievement of a more productive agricultural sector therefore needs to focus on barriers to agricultural exit and land consolidation, especially given the inherent, pervasive and so far intractable problems of credit and insurance markets.

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

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

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, is\(^5\) must be supervised in order to be productive, and (iii) that machinery varies by capacity. These assumptions are embodied in the following function:

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

\(^5\)We provide empirical evidence consistent with this assumption in the empirical section of this paper.
For an interior solution we require $v < 1$, $v < 2\delta$.

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

where $l_s$ denotes supervisory labor, $l_m$ denotes manual labor, $l_s / l_m$ is a constant-returns labor-services production function, $q$ denotes the capacity of each machine, and $k$ denotes the number of machines. Note that the advantage of large farms with respect to higher-capacity equipment is embodied in the function $\phi(a)$ - increasing the capacity of a machine augments productivity more the larger is $a$.

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 $c_q q^\nu$, where $\nu < 1$. We also assume that there is a perfect rental market for machines. The cost of production is

$$(3) \quad p_f f + c_q q^\nu + w_m l_m + w_s l_s,$$

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

$$(4) \quad \tilde{c}(e, a) = \min p_f f + c_q q^\nu + w_m l_m + w_s l_s \quad \text{subject to (2)}.$$

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

$$(5) \quad q = \phi(a) \frac{1 - \nu}{2 - \nu}.$$

Expression (5) indicates that optimal capacity is determined only by area and the parameter $\nu$ and, in particular, is not sensitive to the required total work. Larger operations will use higher-capacity equipment, but an increase in the elasticity of the machinery price with respect to capacity ($\nu$)
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, where \( \omega \), depends on area.

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. We may write the solution to the cost minimization problem as

\[
\tilde{c}(e, a) = c(a)e
\]

and the conditional factor demands as, for example,

\[
\tilde{k}^e(e, a) = k^e(a)e
\]

Implicit differentiation yields

\[
k^e(a) = \frac{k^e(a)\phi'(a)(-2(1 - \delta) + \omega(2 - \nu)(l^m(a)(l^e / l^m)))^\delta}{\phi(a)(1 - \delta)}
\]

which is positive for \( \delta \) 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 \( \delta \) 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

\[
l^e(a) = -\frac{l^e(a)\phi'(a)\phi(a)^{2\delta - 1}k^e(a)^{\delta}(2 - \nu)}{(1 - \delta)} < 0
\]

---

\( ^8 \)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 \( \omega \), depends on area.
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

$$\pi(a) = \max_{\mathbf{a}, \mathbf{f}} a \cdot g(e(a), f(a)) - c(a) \cdot \mathbf{e} - p_f f$$

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

$$\pi^*(a) = \max_{\mathbf{e}^*, \mathbf{f}^*} g(e^*, f^*) - c(a) \cdot \mathbf{e}^* - p_f f^*$$

The envelope condition implies

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

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

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

and fertilizer per unit area increases in area

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

if fertilizer and work are complementary, where $\epsilon_{ee}^*$ is the own price elasticity of demand for
work and $\varepsilon_{fe}$ the fertilizer-work cross-price elasticity. If fertilizer and work are substitutes, the fact that costs of work decline with area will result in substitution away from fertilizer.

The number of machines $k$ per unit area will be increasing in area, for $\delta$ sufficiently close to 1, because (i) there will be an overall expansion of work (15) and (ii) $k$ is increasing in total work. In particular,

$$\frac{dk^*}{da} = \frac{k^*}{k^e} (l_m^e (\alpha)(1-\delta) \frac{k^e}{l_m^e} \varepsilon_{ae} + k^e (\alpha))$$

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).

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 shown in (10). 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,

$$\frac{dl_m^*}{da} = \frac{dl_m^{s*}}{da} (1 + \varepsilon_{ae} (1-\delta))$$

The expression for supervisory labor is the same except that the subscript $m$ is replaced with an $s$.

**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
\[
(19) \quad \rho(a,b^*) = (1 + r(a,b^*))b^* 
\]
where the interest rate \( r \) is increasing in \( b^* \) and decreasing in owned land. 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.\(^9\) 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

\(^9\)In principle, a similar argument may be made for family labor. A farmer with less area
for a given family labor may have lower need to finance hired labor inputs given area and thus
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. The limitation of this
argument is that family labor and dependents of those family workers 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 (a) with food shares at 60-80% it is unlikely that the liquidity effects of family
labor will be substantial and (b) loans to smaller farmers may be otherwise more expensive due
to collateral requirements and/or the relatively high transaction costs per rupee loaned from the
perspective of the lender.
The farmer’s maximization problem with credit market imperfections can thus be restated as
\begin{equation}
\pi^*(a) = \max g(e^*, f^*) - \rho(a, b^*) - (1 + r_0)\alpha^*
\end{equation}
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
\begin{equation}
\frac{d\pi^*}{da} = -\rho_a - \rho_b c'(a)e^*
\end{equation}
where \( \rho_b = \frac{\partial \rho}{\partial b^*} = \frac{dr}{db^*} b^* + 1 + r(a, b^*) > 1 \). Thus profits per acre rise with the size of owned landholdings. The existence of credit market imperfections, 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 as in (14). 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 (22) is relevant to the question of whether land consolidation will improve true profitability per acre. We now 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. The profit function in terms of estimated profits is given by
\begin{equation}
\hat{\pi}^* = g(e^*, f^*) - c(a)e^* - p_f f^*
\end{equation}
where the inputs are determined by programming problem (21). In this case we have

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

where \( \rho_b > 1 \) and the second term in parentheses is positive. Estimated profits also increase with owned landholdings. Comparing (24) 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 \( \rho_c < 0 \) and (24) would be positive if \( \rho_c < 0 \).

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

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

while the marginal return to estimated profits is

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

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 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,

\[ \frac{de^*}{do^*} = \frac{df^*}{do^*} = 0 \]

\[ ^{10} \text{These conditions coincide in the case in which the interest rate is proportional to borrowing per acre.} \]
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, (24) 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).

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 (19), financial intermediaries differentially 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. However, 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 \( \hat{\varepsilon} \). We define the value function recursively as follows:

\[
(28) \quad v(n_t^*, h_t^*) = \max_{\theta_t} \left( \theta_t + g(f_t^* + n_t^*) - r(f_t^* - h_t^*) - h_{t+1}^* + \beta v(n_{t+1}^*, h_{t+1}^*) \right)
\]

where \( \beta < 1 \) is the discount factor and

\[
(29) \quad h_{t+1}^* = h_t^* + \lambda (\theta_t - E(\theta_t))
\]

where \( \lambda \) 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 \( \theta_t \),

\[
(30) \quad n_{t+1}^* = n_t^* + f_t^* - \alpha \theta_t
\]

The idea is that more rapidly growing plants, for example, will deplete the soil of nutrients relatively quickly. For example, if \( \hat{\varepsilon} \) is rainfall, more nutrients are used if rainfall and soil nutrients are complements. The production and credit functions are

\[
(31) \quad g(x) = g_1 x - g_2 x^2
\]
and

\( r(x) = x + r_2x^2 \)

where \( x \) are the respective arguments in (28) and for notational simplicity we set the fertilizer price to one. All of the parameters in (31) and (32) 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):

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

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

\[
\frac{d\hat{\pi}_1^*}{d\theta_0} = -\alpha \frac{g_2 + g_1r_2 - \beta \nu_m}{g_2 + r_2 - \beta \nu_m} + \lambda \frac{(1 - \beta)(g_1 - 1)r_2}{g_2 + (1 + \beta)r_2 - \beta \nu_m}
\]

where \( \nu_m \) is the second derivative from the value function, \( g_1 > 1 \) and interior maximum.

The two key parameters in (34) are \( \alpha \) and \( \beta \), 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 th 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 Kashmir 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 oversampling of high-income households, however, 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 142 villages in the panel survey. Appendix Figure A provides the distribution of own
landholdings in the set of sampled villages in comparison to that from the Census of 2001 reported
in Figure 1. 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. 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 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 as well as all input and output prices. Cross-plot, within-farmer estimates can be biased if plots vary by characteristics that affect productivity. The survey includes seven plot characteristics. 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. 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. Thus, in addition to the panel of households over the 1999-2007-9 interval, there is within-crop year panel data of up to three periods on farm plots.

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 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). The 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, land is primarily 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. Figure 6 provides the fraction of land leased from immediate relatives, others and ‘landlord.’ Outside of West Bengal, 72% of land is rented from family members, and 28% from others. Only a negligible fraction of households report renting land from a landlord. In contrast, in West Bengal, 26% of farmers rent from landlords, and only 7% from family. These data suggest that, with perhaps the exception of West Bengal, reform of tenancy laws may not play a major role in overcoming the disadvantages of small farms.

A key feature of the 2007-8 data, as noted, is that it includes information on supervision time at the plot level. We can thus directly test an assumption of the model that all manual labor, whether family or hired, requires supervision. We estimate a supervision cost equation across plots for the same farmer in a given season:
We could include family size in the specification, but family size may be endogenous; on farms where supervision is particularly advantageous, or hiring supervisory labor is difficult, more family members may be in place.

\[
I_{ijt} = a_{ijt} + a_{t}f_{ijt} + a_{h}l_{ijt} + a_{A}A_{ijt} + X_{ijt}a_{i} + e_{ijt},
\]

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, \( a_{ijt} \) = farmer/season fixed effect, \( X_{ij} \) 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 \( a_{f} = a_{h} > 0 \), that an increase in hired or family labor equally increases supervision time. Note that the farmer/season fixed effect picks up all market prices and all farm-level characteristics in a given season.

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 (35). 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 (35). 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 supervision 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 supervision costs equally. This result is also robust to the inclusion of plot characteristics. The difference in the

---

\(^{11}\) We could include family size in the specification, but family size may be endogenous; on farms where supervision is particularly advantageous, or hiring supervisory labor is difficult, more family members may be in place.
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. Instead, as the model suggests and as the descriptive statistics confirm, mechanization also reduces the need for manual labor and thus supervision costs.

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 variation in the size of plots for a given farmer affects plot-specific per-acre profitability, 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

\[
\delta_{ijt} = b_{0j} + b_{A}A_{ijt} + b_{N}N_{ijt} + \sum_{a}X_{a}a_{j} + u_{ijt},
\]

where \(\delta_{ijt}\) = 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. 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 (\(b_{A}>0\)).

The first column of Table 2 reports the estimates of equation (36) 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 - the larger the size of the plot, given the farmer’s ownership holdings, capabilities, preferences, and family size, the higher are profits per acre. And, if other plots are on average smaller or total cultivated area on the farm is greater, the plot is also more profitable.

Are these profit estimates consistent with scale effects associated with mechanization? The third and fourth columns report estimates of equation (37) with per-acre profits replaced by a dummy variable taking on the value of one if a tractor is used on the plot. 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. 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 thus absorbs farmer characteristics and input prices faced by the farm:
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.

Table 3 presents estimates of (37) 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. Again estimates are shown with and without the plot characteristics. We can reject the hypothesis 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. 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 7 presents smoothed local-area estimates of the effects of fertilizer on profits by land ownership from (6), 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 direct effect of owned land and machinery on per-acre profitability we need to allow for the possibility that landownership is 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

\[
\hat{p}_{ijt} = c_{0ijt} + c_A A_{ijt} + c_f f_{ijt} + X_{ijt} \alpha_t + \hat{\epsilon}_{ijt},
\]

where \( f_{ijt} \) = plot-specific fertilizer expenditures per acre and \( \hat{\epsilon}_{ijt} \) is an iid error. Profit-maximization implies that \( c_f = 0 \).

\( \hat{\epsilon}_{ijt} \) is a sum of a plot-specific random error and a fixed effect that is uncorrelated with the errors in the production function. The estimates presented in Table 3 distinguish between the effects of landholding size on the fixed effect and on the error term. Where there is a positive correlation between \( \epsilon \) and \( f \), the estimates for \( f \) will be inconsistent. That is, the estimates of \( f \) will not reflect the true marginal effects of fertilizer use.

\( \hat{\epsilon}_{ijt} \) is an iid error. Profit-maximization implies that \( c_f = 0 \).

\[
\hat{p}_{ijt} = c_{0ijt} + c_A A_{ijt} + c_f f_{ijt} + X_{ijt} \alpha_t + \hat{\epsilon}_{ijt},
\]

where \( f_{ijt} \) = plot-specific fertilizer expenditures per acre and \( \hat{\epsilon}_{ijt} \) is an iid error. Profit-maximization implies that \( c_f = 0 \).

\( \hat{\epsilon}_{ijt} \) is a sum of a plot-specific random error and a fixed effect that is uncorrelated with the errors in the production function. The estimates presented in Table 3 distinguish between the effects of landholding size on the fixed effect and on the error term. Where there is a positive correlation between \( \epsilon \) and \( f \), the estimates for \( f \) will be inconsistent. That is, the estimates of \( f \) will not reflect the true marginal effects of fertilizer use.

To estimate the direct effect of owned land and machinery on per-acre profitability we need to allow for the possibility that landownership is 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

\[
\hat{p}_{ijt} = c_{0ijt} + c_A A_{ijt} + c_f f_{ijt} + X_{ijt} \alpha_t + \hat{\epsilon}_{ijt},
\]

where \( f_{ijt} \) = plot-specific fertilizer expenditures per acre and \( \hat{\epsilon}_{ijt} \) is an iid error. Profit-maximization implies that \( c_f = 0 \).

Table 3 presents estimates of (37) 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. Again estimates are shown with and without the plot characteristics. We can reject the hypothesis 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. 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 7 presents smoothed local-area estimates of the effects of fertilizer on profits by land ownership from (6), 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 direct effect of owned land and machinery on per-acre profitability we need to allow for the possibility that landownership is 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

\[
\hat{p}_{ijt} = c_{0ijt} + c_A A_{ijt} + c_f f_{ijt} + X_{ijt} \alpha_t + \hat{\epsilon}_{ijt},
\]

where \( f_{ijt} \) = plot-specific fertilizer expenditures per acre and \( \hat{\epsilon}_{ijt} \) is an iid error. Profit-maximization implies that \( c_f = 0 \).

\( \hat{\epsilon}_{ijt} \) is a sum of a plot-specific random error and a fixed effect that is uncorrelated with the errors in the production function. The estimates presented in Table 3 distinguish between the effects of landholding size on the fixed effect and on the error term. Where there is a positive correlation between \( \epsilon \) and \( f \), the estimates for \( f \) will be inconsistent. That is, the estimates of \( f \) will not reflect the true marginal effects of fertilizer use.

\[
\hat{p}_{ijt} = c_{0ijt} + c_A A_{ijt} + c_f f_{ijt} + X_{ijt} \alpha_t + \hat{\epsilon}_{ijt},
\]

where \( f_{ijt} \) = plot-specific fertilizer expenditures per acre and \( \hat{\epsilon}_{ijt} \) is an iid error. Profit-maximization implies that \( c_f = 0 \).

\( \hat{\epsilon}_{ijt} \) is a sum of a plot-specific random error and a fixed effect that is uncorrelated with the errors in the production function. The estimates presented in Table 3 distinguish between the effects of landholding size on the fixed effect and on the error term. Where there is a positive correlation between \( \epsilon \) and \( f \), the estimates for \( f \) will be inconsistent. That is, the estimates of \( f \) will not reflect the true marginal effects of fertilizer use.
\( \hat{\alpha}_{ijt} \) is an iid error. Controlling for farm machinery (mechanization), we expect that the coefficient \( d_i > 0 \) if there are scale effects and also that machinery has a positive marginal return (\( d_k > 0 \)), 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 (38) across survey years to eliminate the farmer fixed effect, we get

\[
\frac{\bar{\alpha}}{dt} = \bar{\alpha}d_0 + d_A \bar{\alpha}A_j + d_k \bar{\alpha}k_j + \bar{\alpha}d_{ijt},
\]

where \( \bar{\alpha} \) is the intertemporal difference operator. In (39), even if the errors \( \hat{\alpha}_{ijt} \) are iid, investments in capital assets such as land or equipment will be affected by prior profit shocks in a world in 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 \( \bar{\alpha}A_{ijt} \) high when \( \bar{\alpha}d_{ijt} \) is low. That is, even if the contemporaneous \( \text{cov}(\hat{\alpha}_{ijt}, A_{ijt}) = 0 \), because assets are measured prior to the profit shock, \( \text{cov}(\bar{\alpha}d_{ijt}, A_{ijt}) \neq 0 \). 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. 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, 2003) contributed to predicting household splits. These include the age of the head in 1999, an indicator of whether the head in 2007-8 had brothers, and a measure of the educational inequality among the claimants to the head’s land in 1999.

Table 4 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 .002 level of significance. Indeed, post-1999 inheritance of land is a significant predictor of the change in landholdings over the period along with the head’s age 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 variants of (39) in which we omit capital equipment in order to estimate the unconditional relationship between landownership size and profitability gross of mechanization. The first two columns of Table 5 report estimates of the two variants of the per-acre profit function (39) 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 5 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 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 8 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.

Figure 9 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. As expected if credit costs decline with land size, the marginal returns to capital decline with land ownership size - 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, and thus employ too much labor, given our findings from the plot level data that labor and equipment are substitutes.

6. Farm Size and Equipment Investment and Rental

Figure 9 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 both 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 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.\footnote{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.} 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.

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:

\begin{equation}
K_{jt} = e_{0t} + e_{Ajt} + e_{kjt} + e_{Bjt} + e_{BAjt} \cdot B_{jt} + i_{jt} + \zeta_{ijt},
\end{equation}

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 ($i_{jt}$), 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 (40) are presented in Table 6; the first-stage estimates are presented in Table 7. 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 6 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 6 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.

This analysis makes use of the 2007-8 household data that includes information on three consecutive planting seasons. We estimate first the relationship between lagged profits and lagged fertilizer use and current profits at the farm level. We include village and season dummies to capture variation in wages and prices, which also influence profitability, as well as farmer specific effects that are constant across seasons. Because we incorporate lagged variables in the profit equation and a farmer fixed-effect estimation is restricted to the second and third seasons. Table 8 presents the within-farm estimates of the effects of farm-level lagged profits and fertilizer use on current profitability per acre, stratified by landholdings.

While there is significant evidence of positive scale economies in terms of profitability, particularly among small landholding households, we see a negative and significant effect of lagged farm profits on current profits in each of the three landholding groups. For the lowest
landholding group a 1000 Rupee increase in farm profits per acre is associated with a 266 Rupee decrease in a subsequent period. This negative coefficient is in principle attributable to two sources, the first is that an increases in production in one period result in greater nutrient extraction and thus lower productivity and/or higher fertilizer costs, both of which lower profitability in the second period. The fact that the lagged fertilizer coefficient is positive in each case is also supportive of this point–greater usage of fertilizer in the past increases current soil nutrition.\(^{14}\) Of course, another potential source of this negative coefficient is the fact that the difference in profits between period t and t+1 is negatively correlated with the differenced residual from the profit equation as argued above. However, the lagged profit coefficient, as noted, also reflects credit constraints. It is thus interesting that the lagged profit shocks get progressively more negative as land size increases. This may reflect the fact that credit costs are lower for larger farmers.

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, which allows us to separate the effect of a crop shock on liquidity from the effect of the shock on soil nutrients. Essentially one can augment the dynamic model by letting cash on hand depend on the unanticipated deviations in the across-plot average shock so that

\[ h_{t+1}^* = h_t^* + \delta(\theta_t - E_t \hat{\theta}_t) \]

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 9, are strongly consistent with the notion that liquidity shocks importantly determine input use and thus affect profitability among small farmers. In particular, conditioning on the lagged profits and fertilizer use on a given plot, a 1000 Rupee increase in

\[^{14}\text{Note that estimation of fertilizer effect is complicated by the fact that we cannot control for nutrient status. But using the fact that fertilizer responds to nutrient status, we may show that unconditional on soil nutrient status lagged fertilizer use will positively affect current profits:}

\[
\frac{d \hat{\pi}}{df} = \frac{(g_2 + g_1 r_2 - \beta \nu_m) r_2}{(g_2 + g_1 r_2 - \beta \nu_m)(g_2 - \beta \nu_m)} > 0
\]
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. 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 consolidation, one should expect to see transfers of land in the form of leasing and or sales that on net move from smaller to larger farms. Smaller farmers should be more likely to sell land and reverse tenancy should be the norm. As we have noted land sales are simply too scarce to characterize patterns. Moreover, our data suggest that, just as for tenancy, land sales are within family - in our data 95% of land sales are from parent to child.

Leasing is more prevalent, although some of the identified scale effects, particularly those associated with credit markets, cannot be exploited through leasing. But, as also noted, leasing is also quite rare with less than 3 percent of households reporting leasing in or leasing out in a given year. Given that scale economies arise in part from contiguous land and that most leasing happens within family, perhaps for reasons of moral hazard associated with land upkeep, the opportunities for productive trade appear to be small. Nonetheless, the 2006 village listing data gives a large enough sample size to look at 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, gross and net of village fixed effects, are plotted in Figures 11 and 12. The first figure, which does not account for village effects, yields a somewhat confusing picture.
One sees, in particular, that although leasing in and leasing out are both quite rare, the leasing in probability is more than twice as high as the leasing out probability. The problem is that Figure 11 combines both differences in leasing behavior by land size within a village and differences across villages in average land size that may be correlated with overall levels of leasing. Indeed, our model suggests that in areas with relatively large plots on average there should less need for leasing to capture unrealized scale economies and thus the rental market may be inactive. As a result one might observe a decline in leasing in and/or leasing out with land size even though within a village, of course, leasing in and leasing out must balance.

Figure 12, which removes village effects, provides a more consistent picture and strongly supports the hypothesis that leasing goes in the direction of capturing scale economies. 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 is profitable.

10. Conclusion

In this paper we have used new panel data describing Indian farms to examine the question of whether the size of landownership holdings matters for farm efficiency and profitability, exploiting both panel information on profits at the plot level and the consequences of division of landholdings due to households splits. Our empirical results indicate that larger farms are more efficient, given the resource costs of farming. On farms with larger owned landholdings there is more mechanization, less labor use per acre inclusive of supervisory labor, and, most importantly, higher 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 in terms
of input efficiency.

These findings imply that farms in India are too small and under-mechanized. Consolidation of landholdings would not only raise farm productivity but also lower labor use per acre as farms adopted mechanized inputs. This in turn implies that industrialization that led to the exit of many workers from the agricultural sector, if that were accompanied by land consolidation, would result in a more efficient agricultural sector. The consequences of labor exit and land consolidation on a large scale, given our estimates, would have to be examined in a general-equilibrium context in order to appropriately assess these effects.

As with any set of findings that suggest that there are profitable opportunities from altering the allocation of resources, the question is why farms in India remain small. This question is also beyond the scope of this paper, but our findings suggest that a serious research program meant to discover how to improve agricultural efficiency in India, and perhaps other countries where property rights area reasonably in place, might be directed to examining the question of why there are so many people in agriculture farming at a small scale and under-exploiting the efficiency of mechanization in the presence of globally increasing returns to farm size.
Figure 1. The Cumulative Distribution of Owned Landholdings (Acres) in India (2001 Census)
Figure 2. Mechanization and Owned Landholdings (Acres), 2007-2008

- Tractor
- Plow
- Thresher

<1 1 2 3 4 5 6 7 8 9 10+
Figure 3. Profits per Acre and Profits per Acre less Supervision Costs, by Owned Landholdings, 2007-8
Figure 4. Total Supervision Costs and Supervision Costs per Acre by Owned Landholdings Size
Figure 5. Total Labor Costs per Acre, by Owned Landholdings, 2007-8
Figure 6. Source of Leased-in Land: India and West Bengal, 2006

- **Immediate Family Member**
- **Other**
- **Landlord**

- **India, except West Bengal**
- **West Bengal**
Figure 7. 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 8. Locally-weighted FE-IV Estimates of the Returns to Capital Equipment Value (and .95 Confidence Intervals), By Landholding Size
Figure 9. 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 10. Locally-weighted Within-Plot Estimates:
The Effects of Lagged Farm Profits on Plot-Specific Profits per Acre (one sd confidence bounds), by Landholding Size
Figure 11. Relationship Between the Probability of Leasing In and Leasing out Land, by Ownership Holdings without Village Fixed Effects, 2006 (N=119,349)
Figure 12. Within-Village Relationship Between the Probability of Leasing In and Leasing out Land, by Ownership Holdings, 2006 (N=119,349)
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

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

Absolute value of asymptotic t-ratios in parentheses. <sup>a</sup>Specification includes season dummy variables; clustered t-ratios at the farm level. <sup>b</sup>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

<table>
<thead>
<tr>
<th></th>
<th>Profits per Acre</th>
<th>Any Tractor Services Used</th>
<th>Total Labor Costs per Acre</th>
</tr>
</thead>
<tbody>
<tr>
<td>Plot area</td>
<td>145.4 (2.34)</td>
<td>.00403 (1.99)</td>
<td>-.00404 (1.98)</td>
</tr>
<tr>
<td></td>
<td>157.0 (2.51)</td>
<td>.00405 (1.98)</td>
<td>-.00404 (2.74)</td>
</tr>
<tr>
<td>Area of other plots</td>
<td>118.7 (1.95)</td>
<td>.00333 (1.77)</td>
<td>-.00351 (1.76)</td>
</tr>
<tr>
<td></td>
<td>130.8 (2.14)</td>
<td>.00336 (1.76)</td>
<td>-.00351 (1.45)</td>
</tr>
<tr>
<td>Total number of plots</td>
<td>482.3 (3.15)</td>
<td>-.0333 (6.68)</td>
<td>-123.2 (1.27)</td>
</tr>
<tr>
<td></td>
<td>473.8 (3.06)</td>
<td>-.0330 (6.53)</td>
<td>-123.2 (1.27)</td>
</tr>
<tr>
<td>Include soil characteristics?*</td>
<td>N Y N Y N Y</td>
<td>N Y</td>
<td>N Y</td>
</tr>
<tr>
<td>Number of plot observations</td>
<td>14,290 14,290</td>
<td>14,290 14,290</td>
<td>14,290 14,290</td>
</tr>
<tr>
<td>Number of farmers</td>
<td>4,130 4,130</td>
<td>4,130 4,130</td>
<td>4,130 4,130</td>
</tr>
</tbody>
</table>

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

<table>
<thead>
<tr>
<th>Owned landholdings</th>
<th>&lt; 4 acres</th>
<th>4-10 acres</th>
<th>10+ acres</th>
</tr>
</thead>
<tbody>
<tr>
<td>Fertilizer use this season</td>
<td>1.49 (3.73)</td>
<td>1.46 (3.89)</td>
<td>3.23 (2.25)</td>
</tr>
<tr>
<td>Plot area</td>
<td>29.9 (1.02)</td>
<td>30.4 (0.97)</td>
<td>901.9 (3.03)</td>
</tr>
<tr>
<td>Include soil characteristics?</td>
<td>N</td>
<td>Y</td>
<td>N</td>
</tr>
<tr>
<td>Number of plot observations</td>
<td>4,008</td>
<td>4,008</td>
<td>1,935</td>
</tr>
<tr>
<td>Number of farmers</td>
<td>1,939</td>
<td>1,939</td>
<td>851</td>
</tr>
</tbody>
</table>

Absolute value of asymptotic t-ratios in parentheses clustered at the village level. Soil characteristics include measures of depth, salinity, percolation and drainage; five soil colors (red, black, grey, yellow, brown, off-white); and five soil types (gravel, sandy, loam, clay, and hard clay). All specifications include plot distance and fertilizer used in the prior period.
Table 4
Panel Data (1999-2008) First-Stage Farmer FE Estimates: Owned Landholdings and Value of Farm Equipment

<table>
<thead>
<tr>
<th>Variable:</th>
<th>Own Landholdings (acres)</th>
<th>Farm equipment x 10^3</th>
</tr>
</thead>
<tbody>
<tr>
<td>Inherited land (acres) after 1999</td>
<td>.193</td>
<td>.881</td>
</tr>
<tr>
<td></td>
<td>(2.69)</td>
<td>(1.23)</td>
</tr>
<tr>
<td>Value of owned inherited mechanized assets in 1999 x 10^3</td>
<td>-.00458</td>
<td>.103</td>
</tr>
<tr>
<td></td>
<td>(0.34)</td>
<td>(1.22)</td>
</tr>
<tr>
<td>Value of owned inherited non-mechanized assets in 1999 x 10^3</td>
<td>-.0746</td>
<td>2.86</td>
</tr>
<tr>
<td></td>
<td>(2.10)</td>
<td>(3.26)</td>
</tr>
<tr>
<td>Value of assets inherited after 1999 x 10^3</td>
<td>.263</td>
<td>-7.45</td>
</tr>
<tr>
<td></td>
<td>(0.41)</td>
<td>(1.06)</td>
</tr>
<tr>
<td>Standard deviation of the schooling of family claimants in 1999</td>
<td>-.0724</td>
<td>2.245</td>
</tr>
<tr>
<td></td>
<td>(1.04)</td>
<td>(2.87)</td>
</tr>
<tr>
<td>Head’s age in 1999</td>
<td>-.0247</td>
<td>-.280</td>
</tr>
<tr>
<td></td>
<td>(2.24)</td>
<td>(1.65)</td>
</tr>
<tr>
<td>Whether respondent has brothers</td>
<td>-.237</td>
<td>1.40</td>
</tr>
<tr>
<td></td>
<td>(0.48)</td>
<td>(0.33)</td>
</tr>
<tr>
<td>Number of observations</td>
<td>3,994</td>
<td>3,524</td>
</tr>
<tr>
<td>Number of farmers</td>
<td>1,749</td>
<td>1,749</td>
</tr>
</tbody>
</table>

Anderson-Rubin Wald test of weak instruments $\hat{\chi}^2(7)$, $p$-value

| Anderson-Rubin Wald test of weak instruments $\hat{\chi}^2(7)$ | 22.32 |
| $p$-value | .0022 |

Absolute value of asymptotic t-ratios in parentheses clustered at the household level.
Table 5
Panel Data Estimates (1999-2008): Effects of Own Landholdings and Own Farm Equipment on Profits per Acre, by Estimation Procedure

<table>
<thead>
<tr>
<th>Estimation procedure:</th>
<th>Village Fixed-Effects&lt;sup&gt;a&lt;/sup&gt;</th>
<th>Farmer Fixed-Effects</th>
<th>Farmer Fixed-Effects IV&lt;sup&gt;b&lt;/sup&gt;</th>
</tr>
</thead>
<tbody>
<tr>
<td>Owned landholdings</td>
<td>13.1 (2.77)</td>
<td>8.35 (0.48)</td>
<td>389.6 (1.99)</td>
</tr>
<tr>
<td></td>
<td></td>
<td>3.53 (0.20)</td>
<td>250.3 (1.37)</td>
</tr>
<tr>
<td>Value of farm equipment</td>
<td>- (1.16)</td>
<td>- (3.26)</td>
<td>- (2.14)</td>
</tr>
<tr>
<td></td>
<td>0.00746 (1.16)</td>
<td>0.0114 (3.26)</td>
<td>0.0347 (2.14)</td>
</tr>
<tr>
<td>Number of observations</td>
<td>3,994</td>
<td>3,524</td>
<td>3,524</td>
</tr>
<tr>
<td>Number of farmers</td>
<td>2,138</td>
<td>1,749</td>
<td>1,749</td>
</tr>
<tr>
<td>Kleinberger-Paap underidentification test statistic</td>
<td>(4) 13.4, .0093</td>
<td>(6) 16.5, .0113</td>
<td></td>
</tr>
<tr>
<td>Hansen J overidentification test statistic</td>
<td>(3) 0.59, .898</td>
<td>(5) 5.47, .361</td>
<td></td>
</tr>
</tbody>
</table>

Absolute value of asymptotic t-ratios in parentheses. <sup>a</sup>Specification includes year=2008 dummy; clustered t-ratios at the household level. <sup>b</sup>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 head’s age in 1999, and owned asset values in 1999.
Table 6
Retrospective Panel Data Estimates (2008): Effects of Own Landholdings and Own Farm Equipment on Investment in Farm Equipment and Equipment Rental, by Estimation Procedure

<table>
<thead>
<tr>
<th>Dependent variable</th>
<th>Equipment Investment</th>
<th>Equipment Hire Expenditure</th>
</tr>
</thead>
<tbody>
<tr>
<td>Estimation procedure</td>
<td>FE-Farmer</td>
<td>FE-Farmer IV&lt;sup&gt;b&lt;/sup&gt;</td>
</tr>
<tr>
<td>Owned landholdings</td>
<td>16.3 (0.05)</td>
<td>663.8 (2.15)</td>
</tr>
<tr>
<td>Landholdings x bank</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>Value of owned farm equipment</td>
<td>-0.0843 (1.30)</td>
<td>-0.909 (8.67)</td>
</tr>
<tr>
<td>Bank within 10 Km</td>
<td>3524 (2.27)</td>
<td>1820 (0.61)</td>
</tr>
<tr>
<td>Number of farmers</td>
<td>3,522</td>
<td>3,522</td>
</tr>
</tbody>
</table>

Absolute value of asymptotic t-ratios in parentheses. <sup>a</sup>Specification includes year=2008 dummy; clustered t-ratios at the village level. <sup>b</sup>Instruments include land inherited after 1999, assets inherited after 1999, and the presence of a bank within 10 km in 1999.
Table 7
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

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

Absolute value of asymptotic t-ratios in parentheses clustered at the farm level. Specifications include season*vector dummy variables.
Within-Farmer Estimates Across Three Seasons (2007-8):
Effects of Previous-Period Profit Shocks on Current Profits per Acre, by Owned Landholding Size

<table>
<thead>
<tr>
<th>Farm size:</th>
<th>Owned Landholdings&lt;4</th>
<th>Owned Landholdings&gt;=4, &lt;10</th>
<th>Owned Landholdings&gt;=10</th>
</tr>
</thead>
<tbody>
<tr>
<td>Farm profits per acre, previous season</td>
<td>-.266</td>
<td>-.503</td>
<td>-.797</td>
</tr>
<tr>
<td></td>
<td>(4.33)</td>
<td>(6.53)</td>
<td>(5.63)</td>
</tr>
<tr>
<td>Fertilizer use, previous season (value per acre)</td>
<td>1.42</td>
<td>.104</td>
<td>3.63</td>
</tr>
<tr>
<td></td>
<td>(3.62)</td>
<td>(1.82)</td>
<td>(4.18)</td>
</tr>
<tr>
<td>Total cultivated area, this season</td>
<td>2494.5</td>
<td>215.5</td>
<td>301.6</td>
</tr>
<tr>
<td></td>
<td>(2.42)</td>
<td>(0.53)</td>
<td>(2.17)</td>
</tr>
<tr>
<td>Number of farmers</td>
<td>2,176</td>
<td>1,061</td>
<td>580</td>
</tr>
</tbody>
</table>

Absolute value of asymptotic t-ratios in parentheses clustered at the farm level. Specifications include season*village dummy variables.
# Table 9
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

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

Absolute value of asymptotic t-ratios in parentheses clustered at the farm level. Specifications include season*vector dummy variables.
Barriers to Farm Profitability in India: Mechanization, Scale and Credit Markets

Andrew D. Foster and Mark R. Rosenzweig

September 2010

Viewed across countries of the world, the most productive agricultural areas are:

1. Family farmed
2. Large scale
3. Mechanized

Indian agriculture is

1. Family farmed, but
2. Very small scale: 80% of farms are less than two acres in size and 95% are less than five acres (Indian Census, 2001)
3. Not mechanized: Less than 3% of farms between 2 and 4 acres own a tractor; but 30% of farms 10+ do (Our national rural survey, 2008)
Descriptively, in India we observe a negative relationship between profitability per acre and landholdings, with and without taking into account supervision costs. Drop in supervision and total labor used per-acre as well.

The question posed in this paper is whether increasing the scale and mechanizing Indian agriculture is key to improving agricultural productivity.

Prior literature focusing on barriers to high productivity agriculture in low-income countries:

1. Many studies on scale economies. Key distinction is between output per acre and profitability. Mixed evidence, unclear definitions of meaning of scale economies, unclear about nature of markets, ownership versus leased operational holdings.

   Biggest *empirical* problem: scale is endogenous, correlated with unobservable farmer and land attributes

   Unlikely to carry out RCT varying owned land.
2. Focus on market imperfections as barriers to agricultural productivity in low-income countries:
   A. Credit markets   B. Insurance markets   C. Labor markets

   But, these market problems are not confined to poor countries:
   All farmers everywhere are credit-constrained;
   Few farmers purchase formal insurance in any country;
   And farmers in developed countries face the same problems of worker shirking and employ the same labor arrangements as in poor countries.

   Problems of credit, insurance, shirking have not been solved anywhere

3. Focus on technology

   But Indian farmers using in many cases the same seeds as in developed countries (Bt cotton, HYV wheat)

4. Focus on behavioral issues impeding efficient resource use

   But the brains of Indian farmers are hard-wired the same as humans elsewhere

5. Property rights

   This is a big issue in many countries, but in India, property rights are reasonably well-established: titling, computerized land records

   The key difference between high-productivity agriculture in developed countries and a low-income country like India would thus appear to be scale and mechanization

   Put another way, are there too many people employed in agriculture, in the sense that removing workers and consolidating landholdings to fewer farmers would increase farm efficiency?

   “Surplus” labor revisited? but not via old and discredited micro mechanisms of compensatory labor supply or nutrition-efficiency wages

This paper:

1. Set out a model to clarify the interrelationships among mechanization, scale economies, credit market constraints, agency issues, lack of insurance, and profitability.

2. Key feature is that scale economies (convexities) arise solely because higher-capacity, mechanized farm implements require more space to operate effectively (two-row versus 8-row harvester); farmers can choose optimally number and capacity of equipment (machines are scalable - not an indivisibility issue).

3. Market setting: Labor requires supervision (shirking); credit costs depend on amount borrowed and on collateral - owned land. But farmers can freely rent equipment and purchase and sell labor services at constant prices.

4. Tests established distinguishing operational scale effects and size of owned landholdings on profitability and role of credit market imperfections.

5. Dynamics introduced to test role relationship between ownership scale and profitability given lack of insurance and credit constraints

6. Empirical application exploits unique data:

   A. Panel data at the plot level on all inputs, outputs and costs (3 seasons)
   1. Within-farmer/season cross plots: Permits identification of pure scale effects on profitability and input use: across plots for the same farmer, so that credit and input costs are held constant (plot characteristics matter)
   2. Within-plots across seasons: perfectly controls for plot and farmer characteristics; distinction between plot and farm level shocks on plot-specific profitability

   B. Longer-term panel at the farm level (9-year interval)
   Identifies effects of changing owned landholdings: exploit family division, inheritance of land
   For same family, observe what happens when land is broken up at death of head: IV, FE approach

7. Results: Yes, increasing own landholdings increases significantly profits per acre and mechanization and lowers labor intensity; scale economies and imperfect credit
2. Theory

A. Technical scale economies, cultivated land area and mechanization

Two inputs to production: Total work to be done $e$, by either machinery or labor

Fertilizer $f$

Output per acre $y$ in a given crop cycle is CRS in acreage $a$, and given by

\[ y = \sigma g(e / a, f / a) \]

The work technology is:

\[ e = \left( \omega_m(l_m, l_s) \right)^{\delta} + \omega_s((\phi(a) - q)qk)^{\delta} \right)^{1 / \delta} \]

where

- $l_m = \text{manual labor service}, \left( \frac{l_m}{l_s} \right)$
- $l_s = \text{constant-returns function of supervision}$

\[ q = \text{machine capacity and } k = \text{the number of machines}. \]

Costs:

A. The price per day of a machine = $c_qq^\gamma$, where $\gamma < 1$. So the rental price increases non-linearly in capacity $q$.

B. Rental markets for machine and labor services are perfect.

A profit-maximizing farmer maximizes the value of (1) minus costs (3) subject to (2).

Optimal machine capacity depends only on area and the elasticity of the price schedule with respect to capacity - larger farms use higher capacity machines

\[ q = \phi(a) \frac{1 - \nu}{2 - \nu} \]

Can solve for the effect of land area, conditional on the amount of “work” to be done

Important features of the work technology:

A. Machinery varies by capacity.

B. The advantage of large farms with respect to higher-capacity equipment is embodied in the function $\phi(a)$ - increasing the capacity of a machine augments productivity more the larger is $a$.

C. Manual labor and machinery services are imperfect substitutes

D. Manual labor, regardless of the hired or family status of that labor, must be supervised in order to be productive, and (iii) that machinery varies by capacity.

A. More machines (of increasing capacity) per unit of work $e$, the larger is $a$.

\[ k^* \gamma(a) = \frac{k^*(a)\phi'(a) - 2\delta(1 - \delta) + \delta(2 - \nu)}{\phi(a)(1 - \delta)} \]

which is positive for sufficiently close to one.

B. Less labor per unit of work and lower costs, conditional on work

\[ \hat{\ell}_m(a) = -\frac{\ell_m(a)\phi'(a)\phi(a)^{2\delta - 1}}{(1 - \delta)} \kappa_1 \delta k^*(a)^{\delta} > 0 \]

\[ c'(a) = -c(a)\phi'(a)\phi(a)^{2\delta - 1} \kappa_1 \delta k^*(a)^{\delta} > 0 \]

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

Embed cost function in profit function, letting the superscript * denote per-acre quantities:

\[ \pi^*(a) = \max g(e^*, f^*) - c(a) e^* - p_f f^* \]

Results:

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

\[ \frac{d\pi^*}{da} = -c(a)e^* > 0 \]

B. Larger operational holdings will use inputs more intensively

\[ \frac{de^*}{da} = c(a) \frac{g_m}{g_m g_f - g_{df}} = c(a) e^* \frac{e^*}{c(a)} > 0 \]

C. Larger operational holdings will use more machinery per acre

\[ \frac{dk^*}{da} = k^* \left( \mu^*(a)(1 - \delta) \frac{k^* c_a}{1 - \mu^*} + k^*(a) \right) \]

Will larger operations use less labor per unit area?

A. There is an increase in work intensity as the increasing returns associated with machinery lower unit work costs.

B. There is also a decrease in the amount of labor per unit work.

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:

\[ \frac{dl_w^*}{da} - \frac{dl^*}{da} e^*(1 + \varepsilon_a(1 - \delta)) \]

where \( \varepsilon_a \) = price elasticity of work

B. Scale effects and credit market imperfections

Now introduce credit market imperfections.

Distinction between ownership and rental now important.

The farmer’s maximization problem with credit market imperfections:

\[ \pi^*(a) = \max g(e^*, f^*) - \rho(a, b^*) - (1 + r^*) o^* \]

where \( r^* \) = rate of return on savings.

Profits per acre rise more steeply with land if there are credit market imperfections:

\[ \frac{d\pi^*}{da} = -\rho_a - \rho c(a)e^* \]

A. There is a negative effect of owned area on interest rates given input use per acre, \( \rho_a < 0 \)

B. Savings in cost per unit of work associated with scale lower the amount borrowed \( b^* \), thus further lowering interest costs.

Note: if no technical scale economies \( (c^*(a) = 0) \), still rising profitability and if perfect credit markets, rising profitability if there are scale economies.

Is there a direct test of credit market imperfections?
The marginal return to capital is given by

\[ \frac{d\pi^*}{du^*} = -\varphi_3 + (1 + \varphi_2) = r(a, b^*) - \varphi_2 + \frac{dr(a, b^*)}{db^*} b^* > 0 \]

Perfect credit markets: \( r(a, b^*) = r_0 \)

Implications:

A. When borrowing costs are independent of land ownership and equal to the returns on savings, the marginal return to capital assets is zero.

Non-zero marginal return to machinery implies imperfect credit markets

B. Marginal returns to capital decline with land size: \( \frac{d\pi^*}{da^*} \) because debt \( b^* \) and interest rates \( r \) fall with \( a \)

C. Larger farms more likely to purchase equipment: face lower borrowing costs but same rental price as other farmers

C. Farm size, savings, no insurance and profits

Now, introduce stochastic output, lack of insurance and possibility of savings: dynamics

Before borrowing determined by demand for inputs

Now, also depends on cash on hand

Cash on hand depends on prior profits

But, prior profits also affect current decisions directly: soil nutrient effects

Two reason for looking at dynamics:

1. In empirical work, the existence of lagged profit effects has implications for econometrics (why farmer fixed effects estimates are biased)

2. Another direct test of credit market imperfections

3. A third advantage to larger land size via credit markets: immunity to shocks

We assume a stationary problem

State variables: soil nutrition \( n^* \), cash on hand \( h^* \)

Farmers choose fertilizer levels \( f^* \) prior to the realization of a shock \( \varv \)

The value function:

\[ v(n_1^*, h_1^*) = \max E_t (\theta_t + g(f_t^*, n_t^*) - r(f_t^*, h_t^*) - h_{t+1}^* + \beta v(n_{t+1}^*, h_{t+1}^*)) \]

where \( \beta < 1 \) is the discount factor

\[ h_{t+1}^* = h_t^* + \delta (\theta_t - E_t \theta_t) \]

where \( \delta \) denotes the extent to which unanticipated shocks are saved.

For \( \delta = 1 \) unanticipated shocks are fully saved as in the permanent income model

For \( \delta = 0 \) cash on hand is a constant

Soil nutrient dynamics:

\[ n_{t+1}^* = n_t^* + f_t^* - \alpha \theta_t \]

\( \alpha \) reflects the idea is that more rapidly growing plants will deplete the soil of nutrients relatively quickly (if \( \varv \) is rainfall, more nutrients are used if rainfall and soil nutrients are complements).

The production and credit functions are quadratic

\[ g(x) = g_1 x - g_2 x^2 \quad \text{and} \quad r(x) = x + r_1 x^2 \]

All of the parameters are positive:

- diminishing returns
- cost of credit \( r \) increases at a higher rate with the amount of credit

Profits gross of credit costs:

\[ \hat{\pi}_1^* = \theta_t + g(f_t^*, n_t^*) - f_t^* \]
The effect of a previous period shock on next-period’s profits is

\[
d\pi_t = -\alpha \frac{g_2 + g_1 r_2 - \beta v_{m}}{g_2 + (1 + \beta) g_1 - \beta v_{m}} \quad (33)
\]

where \( v_{m} \) is the second derivative from the value function, \( v_{m} > 0 \) for an interior maximum.

Two effects of opposite sign (empirical challenge to separate them):

If \( \delta = 0 \), liquidity \( h \) does not depend on unanticipated income shocks
then, only a nutrient carryover or depletion effect, which is negative

If \( \alpha = 0 \), no nutrient carryover,
then a positive shock in period 0 induces higher savings and lowers credit
so that profits are higher

Note: there is no effect of lagged profits if borrowing does not effect credit
or if borrowing is very small: large farmers

Data

Sources:
NCAER REDS 2007-8 round
NCAER REDS 1999 round

Represents rural households in 242 villages in 17 major states of India
(excludes Assam and J & K)
Fifth and sixth rounds of panel survey going back to 1968-69
Originally meant to be representative of entire rural population of India
Oversampled areas conducive to green revolution success higher-income
households

Provide comprehensive inputs and outputs for farming: inclusive of family labor,
supervisory labor, use of own equipment, all costs
Village-level information on institutions and facilities
Histories of family linkages: trace out break-up of households across generations

Four types of information used in our analysis

A. Panel information at the plot level over three seasons from the REDS
2007-8: 10, 947 plots, 2/3 observed at least twice
B. Retrospective histories of equipment and land acquisitions from survey
date in 2007-8 back to 1999: 4, 961 farmers
C. Panel of farm households from the combined 2007-8 and 1999 rounds:
2,848 panel farm households
D. Listing information in 2006 for all households in all villages: 119, 284
households

Use data, guided by model, to estimate effects of land size, mechanization

on profitability on mechanization on returns to mechanization

Key challenges of empirical work:

A. Heterogeneity in land, farmers
B. Endogeneity of land and equipment ownership
C. Identification of pure scale from credit-market effects
Key features of methods:

1. Exploit plot level panel, distinguish plot- from farm-level effects
2. Exploit change in landownership over time based on family break-up:
   - Less than 3% of farmers purchased (sold) land in 9-year period
   - Over 19% of farmers inherited land in 9-year period 1999-2008
3. Allow estimated effects to vary by land size, as theory suggests (local linear approximations by land ownership size)

First, data show that two assumptions of model are not unrealistic:

1. Rental of land unimportant, confined to immediate family members
2. No difference between family and hired labor in terms of demands on supervision

### 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

<table>
<thead>
<tr>
<th>Estimation procedure:</th>
<th>Village Fixed-Effects</th>
<th>Farmer-Season Fixed-Effects</th>
</tr>
</thead>
<tbody>
<tr>
<td>Hired labor costs</td>
<td>0.0402 (3.05)</td>
<td>0.0387 (2.94)</td>
</tr>
<tr>
<td></td>
<td>0.399 (3.37)</td>
<td></td>
</tr>
<tr>
<td>Family labor costs, less supervision time</td>
<td>0.134 (4.11)</td>
<td>0.140 (4.23)</td>
</tr>
<tr>
<td></td>
<td>0.0410 (1.57)</td>
<td>0.0375 (1.44)</td>
</tr>
<tr>
<td>Plot area</td>
<td>0.00407 (1.91)</td>
<td>0.00409 (1.92)</td>
</tr>
<tr>
<td></td>
<td>3.58 (1.51)</td>
<td>3.35 (1.37)</td>
</tr>
<tr>
<td>Owned landholdings</td>
<td>2.74 (0.62)</td>
<td>2.80 (0.62)</td>
</tr>
<tr>
<td></td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>Plot characteristics included</td>
<td>N Y</td>
<td>N Y</td>
</tr>
<tr>
<td>Number of observations</td>
<td>18,484</td>
<td>18,201</td>
</tr>
<tr>
<td>Number of farmer-seasons</td>
<td>8,685</td>
<td>8,587</td>
</tr>
</tbody>
</table>
| Absolute value of asymptotic t-ratios in parentheses. Specification includes season dummy variables; clustered t-ratios at the farm level. 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.

Road map of empirical applications:

A. Estimates of the pure effects of scale and returns to fertilizer use by owned land:
   - exploits within-farm plot differences in size (8 measures of plot characteristics) on profits per acre, use of mechanized inputs, labor intensity (farmer FE)

B. Estimates of mechanization and landownership size on profits per acre:
   - exploits household panel; uses IV, farmer fixed-effects for causal effects

C. Estimates of the determinants of mechanization and role of credit:
   - Exploits retrospective history of equipment acquisitions; uses IV, farmer fixed-effects

D. Estimates of lagged profit shocks, by landholding size, on current profits distinguishing soil-nutrient from liquidity/credit effects
   - Exploits panel of plots across three seasons; (plot and farmer fixed effects)

E. Final section: are we as smart as farmers? Does leasing behavior of Indian farmers reflect the land size effects we find (who rents from whom)
   - Exploits listing data: observe entire market
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

<table>
<thead>
<tr>
<th></th>
<th>Profits per Acre</th>
<th>Any Tractor Services Used</th>
<th>Total Labor Costs per Acre</th>
</tr>
</thead>
<tbody>
<tr>
<td>Plot area</td>
<td>145.4</td>
<td>157.0</td>
<td>-107.2</td>
</tr>
<tr>
<td></td>
<td>(2.34)</td>
<td>(2.51)</td>
<td>(2.74)</td>
</tr>
<tr>
<td>Area of other plots</td>
<td>118.7</td>
<td>130.8</td>
<td>-55.3</td>
</tr>
<tr>
<td></td>
<td>(1.95)</td>
<td>(1.77)</td>
<td>(1.45)</td>
</tr>
<tr>
<td>Total number of plots</td>
<td>482.3</td>
<td>473.8</td>
<td>123.2</td>
</tr>
<tr>
<td></td>
<td>(3.15)</td>
<td>(6.68)</td>
<td>(1.27)</td>
</tr>
<tr>
<td>Include soil characteristics?</td>
<td>N</td>
<td>Y</td>
<td>N</td>
</tr>
<tr>
<td>Number of plot observations</td>
<td>14,290</td>
<td>14,290</td>
<td>14,290</td>
</tr>
<tr>
<td>Number of farmers</td>
<td>4,130</td>
<td>4,130</td>
<td>4,130</td>
</tr>
</tbody>
</table>

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

<table>
<thead>
<tr>
<th>Owned landholdings</th>
<th>&lt; 4 acres</th>
<th>4-10 acres</th>
<th>10+ acres</th>
</tr>
</thead>
<tbody>
<tr>
<td>Fertilizer use this season</td>
<td>1.49</td>
<td>1.46</td>
<td>1.49</td>
</tr>
<tr>
<td></td>
<td>(3.73)</td>
<td>(3.89)</td>
<td>(3.89)</td>
</tr>
<tr>
<td>Plot area</td>
<td>29.9</td>
<td>30.4</td>
<td>30.0</td>
</tr>
<tr>
<td></td>
<td>(1.02)</td>
<td>(0.97)</td>
<td>(0.97)</td>
</tr>
<tr>
<td>Include soil characteristics?</td>
<td>N</td>
<td>Y</td>
<td>N</td>
</tr>
<tr>
<td>Number of plot observations</td>
<td>4,008</td>
<td>4,008</td>
<td>1,173</td>
</tr>
<tr>
<td>Number of farmers</td>
<td>1,939</td>
<td>1,939</td>
<td>1,173</td>
</tr>
</tbody>
</table>

Absolute value of asymptotic t-ratios in parentheses clustered at the village level. Soil characteristics include measures of depth, salinity, percolation and drainage; five soil colors (red, black, grey, yellow, brown, off-white); and five soil types (gravel, sandy, loam, clay, and hard clay). All specifications include plot distance and fertilizer used in the prior period.
Landownership, profits and the returns to mechanization

The profit equation we seek to estimate is

\[ \delta_t = \delta_0 + \delta_A A_t + \delta_i i_t + \delta_{it}, \]

where \( t \) is survey year

\[ k = \text{value of all farm machinery} \]

\[ i = \text{unobservable household fixed effect: land quality, farmer ability} \]

\[ \epsilon_i = \text{iid error} \]

Problem \( \text{cov}(A_t, i_t) \neq 0 \)

Standard fixup - exploit panel data, farmer fixed effect removed

With two observations (1999, 2007-8):

\[ \tilde{\delta}_t = \tilde{\delta}_0 + \delta_i \tilde{A}_t + \delta_{it} \tilde{A}_t k + \tilde{\epsilon}_t, \]

where \( \tilde{A} \) is the intertemporal difference operator.

But, our model says that lagged profit shocks will affect current profits, thus making possible acquisition of land or capital: \( \tilde{\delta}_t \) high when \( \tilde{\epsilon}_t \) is low.

Even if the contemporaneous \( \text{cov}(\tilde{\delta}_t, A_t) = 0 \) (assets are measured prior to the shock) \( \text{cov}(\tilde{\delta}_t, \tilde{A}_t) < 0 \) FE technique introduces a negative bias, where credit is constrained

We use IV to obtain consistent estimates of \( \delta_i \) and \( \delta_i \)

Over the nine-year interval between surveys 19.9% of farms divided and farmers inherited land.

The instruments we use to predict the change in landholdings and equipment of a farmer between 1999 and 2007-8 are:

A. The value of owned mechanized and non-mechanized assets inherited prior to 1999


C. The age of the head in 1999, an indicator of whether the head in 2007-8 had brothers, and a measure of the educational inequality among the claimants to the head’s land in 1999 (F&R, 2003).

<table>
<thead>
<tr>
<th>Table 4</th>
<th>Panel Data (1999-2008) First-Stage FE Farmer Estimates: Owned Landholdings and Value of Farm Equipment</th>
</tr>
</thead>
<tbody>
<tr>
<td>Variable:</td>
<td>Own Landholdings (acres)</td>
</tr>
<tr>
<td>Inherited land (acres) after 1999</td>
<td>.193</td>
</tr>
<tr>
<td>Value of owned inherited mechanized assets in 1999 x 10^{-3}</td>
<td>-.00458</td>
</tr>
<tr>
<td>Value of owned inherited non-mechanized assets in 1999 x 10^{-3}</td>
<td>-.0746</td>
</tr>
<tr>
<td>Value of assets inherited after 1999 x 10^{-3}</td>
<td>.263</td>
</tr>
<tr>
<td>Standard deviation of the schooling of family claimants in 1999</td>
<td>-.0724</td>
</tr>
<tr>
<td>Head’s age in 1999</td>
<td>-.0247</td>
</tr>
<tr>
<td>Whether respondent has brothers</td>
<td>-.237</td>
</tr>
<tr>
<td>Number of observations</td>
<td>3,994</td>
</tr>
<tr>
<td>Anderson-Rubin Wald test of weak instruments: ( \hat{\chi}^2(?) ), ( p )-value</td>
<td>22.32</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Table 5</th>
<th>Panel Data Estimates (1999-2008): Effects of Own Landholdings and Own Farm Equipment on Profits per Acre, by Estimation Procedure</th>
</tr>
</thead>
<tbody>
<tr>
<td>Owned landholdings</td>
<td>13.1</td>
</tr>
<tr>
<td>Value of farm equipment</td>
<td>(2.77)</td>
</tr>
<tr>
<td>Number of observations</td>
<td>3,994</td>
</tr>
<tr>
<td>Number of farmers</td>
<td>2,138</td>
</tr>
<tr>
<td>Kleinberger-Paap underidentification test statistic ( \hat{\chi}^2(df), p )-value</td>
<td>(4) 13.4</td>
</tr>
<tr>
<td>Hansen J overidentification test statistic ( \hat{\chi}^2(df), p )-value</td>
<td>(3) 0.59</td>
</tr>
</tbody>
</table>

Absolute value of asymptotic t-ratios in parentheses. ‘Specification includes year=2008 dummy; clustered t-ratios at the household level.’ 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 head’s age in 1999, and owned asset values in 1999.
Table 6
Retrospective Panel Data Estimates (2008): Effects of Own Landholdings and Own Farm Equipment on Investment in Farm Equipment and Equipment Rental, by Estimation Procedure

<table>
<thead>
<tr>
<th>Dependent variable</th>
<th>Equipment Investment</th>
<th>Equipment Hire Expenditure</th>
</tr>
</thead>
<tbody>
<tr>
<td>Estimation procedure</td>
<td>FE-Farmer</td>
<td>FE-Farmer IV</td>
</tr>
<tr>
<td>Owned landholdings</td>
<td>16.3</td>
<td>663.8</td>
</tr>
<tr>
<td>(0.05)</td>
<td>(2.15)</td>
<td>(1.06)</td>
</tr>
<tr>
<td>Landholdings x bank</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>Value of owned farm equipment</td>
<td>-0.0843</td>
<td>-0.909</td>
</tr>
<tr>
<td>(2.50)</td>
<td>(0.39)</td>
<td>(0.94)</td>
</tr>
<tr>
<td>Bank within 10 km</td>
<td>3524</td>
<td>1820</td>
</tr>
<tr>
<td>(2.27)</td>
<td>(0.61)</td>
<td>(0.84)</td>
</tr>
<tr>
<td>Number of farmers</td>
<td>3,522</td>
<td>3,522</td>
</tr>
</tbody>
</table>

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

Table 7
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

<table>
<thead>
<tr>
<th>Dependent variable/Instrument</th>
<th>Owned Landholdings</th>
<th>Own Farm Equipment</th>
<th>Bank &lt; 10 km of the Village</th>
<th>Own Farm Equipment x Bank Proximity</th>
</tr>
</thead>
<tbody>
<tr>
<td>Inherited landholdings between 1999 and 2008</td>
<td>.938</td>
<td>-.142.6</td>
<td>-.0191</td>
<td>.0432</td>
</tr>
<tr>
<td>(18.6)</td>
<td>(0.40)</td>
<td>(3.65)</td>
<td>(0.57)</td>
<td></td>
</tr>
<tr>
<td>Inherited farm assets between 1999 and 2008 x 10^{-3}</td>
<td>-.00105</td>
<td>.546</td>
<td>.000973</td>
<td>.00229</td>
</tr>
<tr>
<td>(2.61)</td>
<td>(5.31)</td>
<td>(2.48)</td>
<td>(1.78)</td>
<td></td>
</tr>
<tr>
<td>Bank within 10 km of the village in 1999</td>
<td>-.118</td>
<td>-.841.8</td>
<td>-.918</td>
<td>-.3.53</td>
</tr>
<tr>
<td>(0.92)</td>
<td>(0.32)</td>
<td>(10.7)</td>
<td>(6.92)</td>
<td></td>
</tr>
<tr>
<td>Inherited landholdings x bank proximity</td>
<td>.04365</td>
<td>1499</td>
<td>.0182</td>
<td>.904</td>
</tr>
<tr>
<td>(0.85)</td>
<td>(2.22)</td>
<td>(2.92)</td>
<td>(10.9)</td>
<td></td>
</tr>
<tr>
<td>Number of farmers</td>
<td>3,522</td>
<td>3,522</td>
<td>3,522</td>
<td>3,522</td>
</tr>
</tbody>
</table>

Absolute value of asymptotic t-ratios in parentheses clustered at the farm level. Specifications include season*village dummy variables.
Table 8
Within-Farmer Estimates Across Three Seasons (2007-8):
Effects of Previous-Period Profit Shocks on Current Profits per Acre, by Owned Landholding Size

<table>
<thead>
<tr>
<th>Farm size:</th>
<th>Owned Landholdings&lt;4</th>
<th>Owned Landholdings&gt;=4, &lt;10</th>
<th>Owned Landholdings&gt;=10</th>
</tr>
</thead>
<tbody>
<tr>
<td>Farm profits per acre, previous season</td>
<td>-0.266 (4.33)</td>
<td>-0.503 (6.53)</td>
<td>-0.797 (5.63)</td>
</tr>
<tr>
<td>Fertilizer use, previous season (value per acre)</td>
<td>1.42 (3.62)</td>
<td>0.104 (1.82)</td>
<td>3.63 (4.18)</td>
</tr>
<tr>
<td>Total cultivated area, this season</td>
<td>2494.5 (2.42)</td>
<td>215.5 (0.53)</td>
<td>301.6 (2.17)</td>
</tr>
<tr>
<td>Number of farmers</td>
<td>2,176</td>
<td>1,061</td>
<td>580</td>
</tr>
</tbody>
</table>

Number of plot observations: 6,068 3,258 1,919
Number of farmers: 1,351 767 311

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

Table 9
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

<table>
<thead>
<tr>
<th>Farm size:</th>
<th>Owned Landholdings&lt;4</th>
<th>Owned Landholdings&gt;=4, &lt;10</th>
<th>Owned Landholdings&gt;=10</th>
</tr>
</thead>
<tbody>
<tr>
<td>Farm profits per acre, all other plots, previous season</td>
<td>0.140 (2.04)</td>
<td>0.0628 (0.67)</td>
<td>0.0366 (0.17)</td>
</tr>
<tr>
<td>Farm profits per acre, this plot, previous season</td>
<td>-0.456 (5.60)</td>
<td>-0.504 (4.78)</td>
<td>-0.540 (2.88)</td>
</tr>
<tr>
<td>Fertilizer use, this plot, previous season (value per acre)</td>
<td>1.54 (2.61)</td>
<td>0.789 (1.97)</td>
<td>1.87 (1.22)</td>
</tr>
<tr>
<td>Number of plot observations</td>
<td>6,068</td>
<td>3,258</td>
<td>1,919</td>
</tr>
<tr>
<td>Number of farmers</td>
<td>1,351</td>
<td>678</td>
<td>311</td>
</tr>
</tbody>
</table>

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

Figure 10. Locally-weighted Within-Plot Estimates:
The Effects of Lagged Farm Profits on Plot-Specific Profits per Acre (one sd confidence bounds), by Landholding Size

Figure 12. Within-Village Relationship Between the Probability of Leasing In and Leasing out Land, by Ownership Holdings, 2006 (N=119,349)
Conclusion

A. The principal proximate barriers to profitable, efficient agriculture in India is the small size of owned landholdings and lack of mechanization

Larger farms are more mechanized, have higher profits per acre, use labor less, more efficiently use fertilizer and equipment, and their incomes are less susceptible to profit shocks

B. Small landholdings lead to under-mechanization and higher labor costs due to technical scale economies, credit market imperfections and lack of insurance

C. Indian farmers attempt to increase scale through the land rental market, but that market is very limited (contiguity important) and does not mitigate credit issue

D. First step: key question is why are farms so small? Why are there too many farms?

E. Possible that attraction of people out of agriculture could make farming more efficient and productive if leads to land consolidation

“Out of agriculture” policy may be a pro agriculture policy