Draft:10/30/2013

Spillovers, Coordination Failure, and the Consequences of Fragmentation in Rural India

Andrew Foster

Department of Economics, Brown University

Hari Nagarajan

Institute for Rural Management Anand

Abstract

In this paper we take advantage of a 5-year panel data set from two villages in rural India to examine the potential role of spillovers and coordination failures as a source of inefficiency due to fragmentation. We first develop a model based on a simple characterization of the spatial structure of land fragments that illustrates how spillover effects impact input use and productivity. Our results suggest that spillover effects are, for the most part negative and that there are significant failures in coordination. The estimates also suggest that, consistent with the model, the surprising absence of a diversification effect may be attributed to higher levels of input variance arise in farmers with more fragmented holdings.

I. Introduction

In the last decade concerns about rising food prices and a perceived need to transition workers from low productivity work in agriculture to higher value-added activities in manufacturing and services has led to renewed interest in the structure of production and sales of agricultural commodities. It has been argued, for example, that scale economies associated with agricultural mechanization may have changed the relative profitability of large and small farms and thus raised the benefits of either actual or effective consolidation of landholdings (Foster and Rosenzweig 2011). While the study of land size in part involve an analysis of the relative productivity of farmers with different levels of average holdings, it also involves understanding the differences in productivity among farmers with more and less fragmented holdings. Given the perception that failures in land, labor and credit markets are an important reason for this fragmentation, there is some question as to whether the process of development in rural areas may have already altered the relative costs and benefits of managing farms compose of multiple non-contiguous fragments.

From the perspective of an individual farmer, the usual arguments favoring fewer land fragments include reduced travel times, less boundary waste, the feasibility of larger-scale productive investment, and the ease of labor supervision. In short, when operating across multiple land fragments the farm enterprise will have a higher proportion of family and hired labor devoted to moving people and inputs across the fragments of the farm (Lipton, 2009). It is thus not surprising to find that early empirical research such as Bardhan (1973) used land fragmentation as an indicator of productive inefficiency. Others have argued, however, that the persistence of the institution suggests that there must also be some benefits to fragmentation of land holdings (McCloskey 1975). For example, fragmented land may facilitate diversification thereby reducing year-to-year variation in profitability. In wet years, when lowland fragments are waterlogged highland fragments may yield relatively favorable yields, while only lowland fragments may receive adequate moisture during dry years (Lipton 2009). Different fragments also might be differentially suited to different types of crops thus leading to protection from the consequences of fluctuations in the price of any particular crop.

Because fragmentation is a farmer-level measure, empirical studies of the consequences of fragmentation have typically made use of data aggregated to the level of the farmer. The approach pioneered by Bardhan (1973) consideres the consequences of total (considering all fragments) area on aggregate yields and on labor input per acre and finds differences in returns to scale in different region. The analysis also includes the number of fragments as a measure of fragmentation. Bardhan argues, as noted, that fragmentation is an index of inefficiency and the empirical results seem to support that view, with fragmentation in many cases leading to lower outputs and higher inputs in terms of labor. He also notes the need for analysis that uses better measures of fragmentation, inclusive of information on fragment distances, which suggests a particular interest in the role of travel time between fragments as a source of this efficiency.

A significant limitation of examining fragmentation aggregated to the level of the farmer is that, at that level, three potentially important determinants of yields—

fragmentation, total land area, and average fragment size[1]--are mutually collinear. That is, it is impossible to identify separately the effects of fragmentation given total land area from the effects of fragment size given total land area. Thus a world in which there are fragment level increasing returns to scale might show evidence of positive effects of total area on yields given the total number of fragments (because fragments are larger if total area is higher given the number of fragments) and a negative effect of the number of fragments given total area (because average fragment size will be smaller). These effects can, however, also be generated by a model in which, for example, travel time between fragments reduces total output and in which there are positive returns to scale in terms of overall cultivated area. To distinguish these different models one needs to make use of fragment level data.

One possible source of both fragment-level returns to scale and overall effects of fragmentation given total area that has not been received rigorous empirical or theoretical attention is coordination failures across fragment boundaries.[2] The idea here is that input use diffuses across space. For example, application of pesticides in a particular location of a farm will not only affect that particular location but also other nearby areas. These spillovers might be positive if the pesticide diffuses or, conversely, negative if the pesticide stays in one place and the pests look for unprotected crops. Similarly, water applied in one area may literally "spill over" into other nearby areas. If these spillovers

[1] As discussed in detail below, it is helpful to distinguish two uses of the word plot--subdivisions—parcels of land that are registered, bought, rented and sold--and fragments—contiguous sets of subdivisions that are farmed as a single unit. We will use the word plot interchangeably with fragment and specifically refer to a subdivision when the alternate meeting is needed.

[2] The notion that coordination failures between proximate fragments may be a significant source of fragmentation inefficiency is not itself new. Heston and Kumar (1983), in their class article discussing the persistence of land fragmentation in South Asia, citing McCloskey (1975), note that "costs of travel and enforced coordination" with neighbors could amount to about 7% of output.

occur within a particular farm they present no particular problem as a single farmer can internalize these effects. However, in the absence of a mechanism for coordination among farmers with proximate fragments, applications of inputs will generally be inefficient and this inefficiency will be greater when fragments are small for a given rate of input diffusion. There is also qualitative evidence from the farmers in our study area that suggests that coordination among neighbors may be an important issue. One farmer with 6.5 acres, for example, argues “There is much benefit [to land consolidation]. For drainage and irrigation. One applies fertilizer in one's field and water must be held there for a time. If there is another person's land adjacent your field will be drained. There are problems. For instance my younger brother's land was adjacent to mine and he sold it to someone else. And he drains my freshly fertilized field.”

While this is the first analysis, of which we are aware, to test directly for spillover effects that arise from the diffusion of agricultural inputs among neighboring farmers, there is a related body of work that focuses on learning about new technologies as a source of spillovers. There is, in particular, a significant and growing literature, reviewed by Foster and Rosenzweig (2011) on the adoption of agricultural technologies (and other technologies in rural areas) showing that experience by some households influences the behavior of other households through the acquisition of knowledge. A subset of this literature also examines the questions of whether these external effects are internalized by the community (e.g., Foster and Rosenzweig 1996). In contrast, however, to the learning literature in which spillover effects diminish as experience with the new technology accumulates, input spillovers, if present, are likely to persist. As a result input spillovers that are not internalized may create not only short but long term returns to scale.

Moreover, in contrast to learning externalities, which are always non-negative, it is possible that input spillovers can create negative externalities.

Perhaps even more so than in the learning literature, the empirical analysis of spillover effects and coordination failures in inputs is especially challenging from the perspective of demands on the data. Not only must one observe fragment level information for each farmer, but one must have fragment level information on the farmers with fragments that are proximate to the his fragments. This requirement is unlikely to be met by data sets with only a sample of village households. Instead one either needs to purposively sample households with adjacent fragments or, perhaps more straightforwardly, have a census of all farmers in particular region.

In this paper we take advantage of a 5-year panel data set from two villages in rural India to examine the potential role of spillovers and coordination failures as a source of inefficiency due to fragmentation. We first develop a model based on a simple characterization of the spatial structure of fragment that illustrates how spillover effects impact input use and productivity. We then test the implications of the model using a series of reduced form and structural input and yield equations. Identification for the structural equations comes from the fact that, given the structure and equilibrium concept used in the model, the characteristics of the non-adjacent fragments of the owner of an adjacent fragment only affect input use on a given fragment through the inputs on the adjacent fragment. Our results suggest that spillover effects are present. The estimates also suggest that the surprising absence of a diversification effect (fragmented holding exhibit greater total variance than unfragmented holdings given total area) may be attributed in part to a spillover effect.

II. Theory

In order to incorporate the effects of input and, in particular, to capture the impacts of spillover effects from the input use on proximate fragments it is helpful to construct an explicitly spatial model that distinguishes between the actual amount of a given input applied at a particular point and the effective input at that point, which adjusts for the spillover effects caused by inputs from neighboring points. The model as constructed yields a series of testable predictions about the relationships between fragment size, inputs of neighboring fragments, and the degree of fragmentation of the holdings of the farmer of a given fragment and of the those farming fragments that abut that given fragment.

We assume that output per acre at a point (x, z) depends not on actual inputs at that point but on the effective input I^* according to a concave quadratic production function

$$g(I^*) = g_1 I^* - \frac{1}{2} g_2 (I^*)^2 . \quad (1)$$

The effective input on any given fragment of land, in turn, depends on the amount of that input applied to a particular point (the direct effect) as well as the weighted (by proximity) average amount of input applied to neighboring points (the spillover effect).

In particular, consider a map in the form of a Cartesian plane placed on a wall and let fragments be rectangles of varying area A_i arrayed along an infinitely wide ribbon of fixed height h on the map (say bounded at top and bottom by roads)[3] so that each fragment has two proximate fragments managed by other farmers. Thus on the map, the

[3] Describing this as a map on the wall rather than an actual representation of the land viewed from above facilitates discussion of the different dimensions as height and width rather than the more ambiguous length and width. .

left and right boundaries of fragment *i* can be denoted x_i^L and $x_i^R = x_i^L + A_i / h$,

respectively, and the top and bottom boundaries can be denoted $z_i^T = h$ and $z_i^B = 0$.

Now suppose that input use does not vary in the vertical direction so that for any point (*x*,z) it can be characterized by some function $\tilde{I}(x)$. Further, suppose that the relative spillover impact of applying input use at a point *x* on different neighboring points can be described by a function $\alpha f(\alpha u)$ that is symmetric about zero.[4] Then the effective input at a point (*x*,z) is

(3) $$I^*(x,y) = \phi_1 \tilde{I}(x) + \phi_2 \int_{x'=-\infty}^{\infty} \tilde{I}(x')\alpha f(\alpha(x'-x))dx'$$

where ϕ_1 and ϕ_2 with $\phi_1 + \phi_2 = 1$ may be thought of, for positive spillovers, as relative weights of direct and spillover effects on effective inputs. That the weights sum to one ensures that effective spillovers equal actual inputs if inputs do vary across neighboring fragments—that is if $\tilde{I}(x) = \tilde{I}$ then $I^*(x,y) = \tilde{I}$. We assume $\phi_1 \in [0,2)$. That is inputs are defined such that the direct effect must be nonnegative but to capture moderate negative spillovers we allow $-1 < \phi_2 < 0$. In terms of pesticides, for example, $\phi_2 > 0$ implies that pesticides diffuse from one location to another but $\phi_2 < 0$ implies pesticide use in one fragment moves pests to a neighboring fragment.

[4] The function $\alpha f(\alpha u)$ is essentially a density function with mean zero and *f*(0)=1. For positive spillovers it thus may be thought of as the probability that a molecule of input deposited at point x jumps horizontally to point $x+u$. We disregard the possibility of diffusion in the vertical direction given that the strategic effects of interest only arise from interactions with neighbors who are, by assumption, only to the left and right.

Now suppose that each farmer applies informally uniformly on his fragment[5]. Let I_i denote inputs per unit land on fragment *i* and let the input use of all farmers on the fragments to the left and all farmers to the right of him be denoted by J_i^L and J_i^R, respectively. Then the effective input on fragment *i* on points within the range $x \in [x_i^L, x_i^R]$ and $z \in [0, h]$ is as

(4) $$I^*(x,z) = \phi_1 I_i + \phi_2 (I_i + (J_i^L - I_i)F(-\alpha(x - x_i^L)) + (J_i^R - I_i)F(-\alpha(x_i^R - x)))$$

where $F(u) = \int_{u=-\infty}^{u} f(u')du'$. For small α diffusion is high and even large differences in input use result in only small differences in the spillover effect across neighboring fragments; however, even if $\alpha = 0$ effective input use will jump at the boundary as long as $\phi_1 > 0$. As α gets larger diffusion decreases. As $\alpha \to \infty$, $F(\alpha x)$ becomes a step function and thus effective input use is equal to actual input use regardless of whether spillovers are positive or negative.

Intuition for equation (4) may be gained through the examination of Figures 1 and 2. These figures plot effective inputs for a particular set of parameter values and inputs. In this case it is assumed that the fragment located in the interval [0,2] on the horizontal axes has a per-acre input level of 3. For Figure 1 the input use of the left and right proximate fragments are 1 and 2, respectively, and for Figure 2 the inputs uses are 4 and 5, respectively. Three alternatives are considered for each figure: negative spillovers $(\phi_1 = 3/2, \phi_2 = -1/2)$, positive spillovers $(\phi_1 = 1/2, \phi_2 = 1/2)$, and zero

[5] In principle a farmer could do better by allowing his input to vary so that effective input does not vary within a fragment. Such variation would be hard for a farmer to implement in practice and, in the context, of our model would be a second order effect in the parameter α and thus would not change the main implications derived.

spillovers $(\phi_1 = 1, \phi_2 = 0)$. We set $\alpha = 4$ and use a logistic function for F (u). In Figure 1 it is evident that when one's neighbors use lower levels of inputs and there are positive spillovers, effective input use is low, relative to no spillovers given input use. The opposite is true when there are negative spillovers. In Figure 2 we examine the alternate case in which neighbors are using higher input levels that the individual of interest. In this case, for example, negative spillovers reduce effective inputs.

Estimation of (4), if possible, would say a great deal about the importance of spillover effects and thus of the possibility that fragmentation affects profitability due to the lack of coordination across farms. In practice, of course, one does not generally observe levels of input, effective inputs, or yields at a particular point in space. Instead one observes aggregates at the level of the fragment. One thus must relate inputs for a fragment, I_i, as well as for that of neighboring fragments J_i^L and J_i^R, to average yields for that fragment

(5) $$y_i{*} = \frac{1}{A_i} \int_{x=x_i^L}^{x_i^R} \int_{z=0}^{h} g(I{*}(x,z))\,dx\,dz\,.$$

We now consider a maximization problem across the N_k fragments of farmer k:

(7) $$\max \sum_{i=1}^{N_k} \int_{x=x_{ik}^L}^{x_{ik}^R} \int_{z=0}^{h} g(I{*}(x,z))dxdz$$

subject to the constraint

(8) $$\sum_{i=1}^{N_k} s_{ik}\Delta_k I_i = \Gamma_k \Delta_k$$

Where Γ_k denotes total inputs available to farmer *k* per unit cultivated area. Initially we consider choices condition on the shadow price of inputs (the Lagrange multiplier on the constraint) μ_k, and thus the analysis applies equally to the case where inputs may be purchased in any quantity that is desired at a market price μ_k. Differentiating, solving, and allowing for possible other fragment characteristics d_{ik} (such as distance of the plot from the home) that may enter the problem, for example, through heterogeneity in the production function parameters as in $g_1(d_{ik})$ and $g_2(d_{ik})$, we obtain equations for inputs and yields by farmer *k* on fragment *i* that condition on neighboring farmer inputs and the shadow price of inputs for *k*.

(8) $$I_{ik} = f^1(s_{ik}, J_{ik}, d_{ik}, \mu_k, \Delta_k)$$

(9) $$y_{ik} = f^2(s_{ik}, J_{ik}, d_{ik}, \mu_k, \Delta_k)$$

where $J_{ik} = (J_{ik}^L + J_{ik}^R)/2$ [6] and $s_{ik} = A_{ik}/\Delta_k$ is the share fragment *i* is of total land $\Delta_k = \sum_{i=1}^{N_k} A_{ik}$ of farmer k.

To maintain analytic tractability it is convenient to construct linear approximations to (8) and (9) around $\alpha = 0$ [7]. Thus, for example,

(10) $$I_{ik} = \Bigg(\frac{g_1(d_{ik}) - g_1(d_{ik})\phi_2 - g_2(d_{ik})\phi_2 J_{ik} + g_2(d_{ik})\phi_2^{\,2} J_{ik} - \mu_k}{g_2(d_{ik})(1-\phi_2)^2} + \frac{\left(-g_1(d_{ik})\phi_2 + g_2(d_{ik})\phi_2 J_{ik} - g_2(d_{ik}) J_{ik} + g_1(d_{ik}) - 2\mu_k\right)\phi_2 s_{ik}\Delta_k}{hg_2(d_{ik})(-1+\phi_2)(1-\phi_2)^2}\alpha + O(\alpha^2)\Bigg)$$

The order-α^2 comparative statics (we omit the $O(\alpha^2)$ term) of interest for (10) are thus:

---

[6] We will employ throughout the assumption that only average neighborhood characteristics enter the decision rules rather than entering them as separate arguments. This assumption is, of course, quite useful empirically and may be justified to the extent that these expressions are approximately linear in inputs.

[7] The expansion around a small value may seem a bit unrealistic given the effects of many inputs is likely to be quite local. However, it provides a substantial advantage in terms of tractability and seems to capture the main forces we wish to illuminate. Moreover, a high rate of diffusion does not necessarily imply anything about whether there are spillovers; that is governed by the parameter ϕ_2.

(11) $$\frac{dI_{ik}}{d\mu_k} = -\frac{1}{g_2\phi_1^2} + 2\frac{\phi_2 s_{ik}\Delta_k}{\phi_1^3 h g_2}\alpha$$

(12) $$\frac{dI_{ik}}{ds_{ik}} = \frac{\phi_2\Delta_k\left(g_2 J_{ik}\phi_1 - g_1\phi_1 + 2\mu\right)}{\phi_1^3 h g_2}\alpha$$

and

(13) $$\frac{dI_{ik}}{dJ_{ik}} = -\frac{\phi_2}{\phi_1} + \frac{\phi_2 s_{ik}\Delta_k}{\phi_1^2 h}\alpha$$

,

where we have dropped the dependence of the g_i on d_{ik} for notational convenience.

Assuming α is sufficiently small, (11) provides the expected result that an increase in the shadow price of inputs results in lower input use. Expression (12) shows that the effect of an increase in fragment size on inputs per acre is zero in the absence of a spillover effect ($\phi_2 = 0$). Otherwise, if input levels across neighbors are sufficiently close then the term in parenthesis will be positive for negative spillovers and negative for positive spillovers. [8] Thus an increase in fragment size will generally increase input use if spillovers are negative. Finally the effects of neighbors' inputs on own inputs to fragment *i* will be positive for small α if spillovers are negative and negative if spillovers are positive. In short negative spillovers create a kind of arms race—with coordination both farmers could lower input costs with relatively minor consequences for effective inputs and thus output. Conversely, with positive spillovers, farmers free ride on the inputs of their neighbors leading overall to under provision of inputs.

Similarly one can obtain approximate (to order α^2) estimates of comparative statics for (9):

---

[8] If $J_{ik} = I_{ik}$ then the first order condition dictates $g_1 - g_2 J_{ik} = \mu_k$ so $\phi_1 g_1 - \phi_1 g_2 J_{ik} < 2\mu_k$ for $\phi_1 < 2$.

(12) $$\frac{dy_{ik}^*}{d\mu_k} = -\frac{\mu}{g_2\phi_1^2} + 2\frac{\phi_2 s_{ik}\Delta_k\mu_k}{\phi_1^3 h g_2}\alpha$$

(13) $$\frac{dy_{ik}^*}{ds_{ik}} = \frac{\phi_2\Delta_k\mu_k^2}{\phi_1^3 h g_2}\alpha$$

(14) $$\frac{dy_{ik}^*}{dJ_{ik}} = 0$$

Not unexpectedly an increase in the shadow price of inputs lowers yields. Consistent with the effects on inputs, an increase in fragment size in (13) has no effect on yield in the absence of spillovers. With negative spillovers an increase in size decreases yields and with positive spillovers an increase in size increases yields. Finally, we see that to order α^2 there is no effect of neighbors' inputs on yields. In the case of negative spillovers one raises one's own inputs sufficiently to offset the negative consequences of the neighbors' input use on yields. With positive spillovers one reduces ones inputs to fully offsets the benefits received from one's neighbors. Of course the offset effect does influence profitability because one pays for one's inputs but not for one's neighbors inputs so profits will be increasing in neighbors' inputs with positive spillovers and decreasing in neighbors' inputs with negative spillovers.

At this point it is helpful to consider the implications of the model for the debate about returns to scale in agriculture. Intuition would suggest that with spillovers larger farmers will be more efficient because they can better internalize spillover effects. This will be true regardless of whether there are positive or negative spillovers. But in fact larger famers have higher yields with positive spillovers and lower yields with negative spillovers. Thus the relationship between yields and land sizes is not necessarily a good gauge of efficiency. Nor, in a sense, is individual profitability. In the presence of positive spillovers a small farmer with large neighbors will be extremely profitable because his inputs are in effect being financed by his neighbors. To evaluate efficiency one needs to think about input use from the perspective of the economy as a whole.

In particular, consider a symmetric Nash equilibrium of a large group of farmers with interlaced fragments where all fragments are of identical size who face an identical input price. Determining optimal I_i given $\overline{J}_i$ and then substituting $\overline{J}_i = I_i$, solving and substituting back into profits yields profits per unit area as a function of area, π_i, for which

(15) $$\frac{d\pi_i}{ds_i} = 2\frac{\phi_2{}^2 \Delta_k \mu_k{}^2 \alpha}{\left(2-\phi_2\right)^3 g_2}$$

Thus, there are positive scale economies as long as there are spillovers and regardless of whether spillovers are positive or negative.

As noted, the analysis to this point has conditioned on the shadow price of inputs. As should be clear from the maximization problem defined by (6), (7) and (8), the first order condition for inputs conditional on the input price and the inputs of neighboring fragments does not depend on the size of other fragments. But it is also instructive to consider the case in which inputs are constrained so that μ_k is endogenous. The shadow price of inputs in that case will be determined in particular by the distribution of fragments owned by a particular farmer as well as the distribution of input use of neighbors of those fragments, which itself is determined by the distribution of fragments of those neighbors, *ad infinitum*. In particular, to order α^2 we can solve for μ_k, yielding an equation of the form

(16) $$\mu_k = f^3(\eta_k, \Delta_k, \overline{d}_k, \overline{J}_k)$$

($\bar{J}_k$ denotes the average neighbors' inputs across all the fragments of farmer *k* and so forth) where:

(17) $$\eta_k = \sum_{i=1}^{N_k} s_i^{\,2} \,,$$

which is the Herfindahl-Hirschman Index (HHI) typically used to measure market concentration. In particular,

(18) $$\frac{d\mu_k}{d\eta_k} = \frac{\phi_2 \Delta_k \alpha\left(g_1 - (2-\phi_2)\Gamma_k g_2 + (1-\phi_2)\bar{J} g_2\right)}{2h^2} \,.$$

Intuition for this result may be obtained by considering a thought experiment for a farmer with two fragments of different sizes. Now suppose the larger fragment is made larger and smaller fragment in such a way that total area is the same. Then to first order (12) tells us that input use per area on the large fragment, given negative spillovers, will decrease as much as the input use per area on the small fragment rises. But this same absolute change in input use per area is multiplied by area so total input use and thus total input use per total area at given price declines. To satisfy the constraint (8) given fixed inputs the shadow price of inputs must fall.

Substitution of (16) into (9) and (10) provides input decision rules that depend on neighbors' inputs and the set of characteristics of farmer k

(19) $$I_{ik} = f^{\,4}(s_{ik}, J_{ik}, d_{ik}, \eta_k, \Delta_k, \bar{d}_k, \bar{J}_k)$$

(20) $$y_{ik} = f^{\,5}(s_{ik}, J_{ik}, d_{ik}, \eta_k, \Delta_k, \bar{d}_k, \bar{J}_k)$$

The shadow price of inputs is also increasing in average (across *k*'s farmers) neighbors' inputs given negative spillovers:

(18) $$\frac{d\mu_k}{d\bar{J}_k} = -\frac{\phi_2 g_2\left(2(1-\phi_2)h - (2-\phi_2)\eta_k \alpha \Delta_k\right)}{4h^2} \,.$$

As neighbors' input use goes up one must compensate by increasing one's own input for a given price. Thus if input supply is fixed the shadow price must rise. Consequently an increase in the average (across *k*'s plots) of neighbors' average input use ($\overline{J}_k^\dagger$), should decrease input use on fragment *i* net of neighbor's (of *i*) average input use, given negative spillovers. This result arises because the higher average input use by neighbors of the other fragments of farmer *k* increases input use on those fragments and this in turn raises the shadow price of inputs for farmer *k*.

Further, noting that the inputs of neighbors' farmers are themselves a result of input decisions made by that neighbor, which themselves are affected by the input decisions on *i*, we can in principle solve simultaneously to obtain reduced form input demand equations that depend on both own and neighbors' attributes. E.g., if $s_{ik}^\dagger$ denotes the average of the fraction of the land of neighbors to i that is in the respective adjacent plots and so forth then

$$(21) \qquad I_{ik} = f^{6}(s_{ik}, d_{ik}, \eta_k, \Delta_k, \overline{d}_k, s_{ik}^\dagger, d_{ik}^\dagger, \eta_{ik}^\dagger, \Delta_{ik}^\dagger, \overline{s}_k^\dagger, \overline{d}_k^\dagger, \overline{\eta}_k^\dagger, \overline{\Delta}_k^\dagger)$$

and

$$(22) \qquad y_{ik} = f^{7}(s_{ik}, d_{ik}, \eta_k, \Delta_k, \overline{d}_k, s_{ik}^\dagger, d_{ik}^\dagger, \eta_{ik}^\dagger, \Delta_{ik}^\dagger, \overline{s}_k^\dagger, \overline{d}_k^\dagger, \overline{\eta}_k^\dagger, \overline{\Delta}_k^\dagger)$$

That is input use and yield per acre on fragment *i* by farmer *k* depend on the characteristics of that fragment and farmer, the characteristics of the fragments and farmers that are adjacent to *i*, and the characteristics of the fragments and farmers that are adjacent to all of *k*'s other plots. The true reduced form, of course, should in principle also include information on the neighbors of the neighbors and so forth; however in the empirical analysis below we limit to this first level of recursion.

The previous comparative statics yield predictions about how a change in the HHI controlling for fragment area on a particular fragment affects input use on that fragment from (21) and (22). In particular for fixed inputs, total area, and area on a particular fragment, an increase in *k*'s HHI with negative spillovers increases input use on that fragment. Moreover given negative spillovers an increase in the HHI of neighbors to *k*'s *i*th fragment will increase use on the neighbors' fragments (J_{ik}) and this will in turn increase input use (I_{ik}) on fragment *i* through (13). Thus both the own (η_k) and average neighbors' HHI ($\eta_{ik}^{\dagger}$) will lead to an increase in input use on a given fragment in the presence of negative spillovers, and conversely for positive spillovers.

III. Data

The data requirements for an analysis of this type are, as noted, extensive. One must have access to longitudinal data on all the fragments of each farmer in the sample, as well as all the fragments of those farmers who have land proximate to the fragments of farmer i. The data for this study were collected over a 5-year period beginning 1995 in 2 contiguous villages situated in Tamil Nadu, India that we denote as N and T. For measuring spillovers or attributing efficiency losses or gains to spillovers one must have space covering data. That is, there must be input and output information for contiguous fragments and preferably for all farmers within a village. Such data were collected from 137 farmers (cultivators) in N and 83 farmers (cultivators) in T, for the period 1995-6 to 1999-2000. The survey included detailed questions on inputs, outputs, land quality, irrigation management, apart from those that related to household welfare such as demographics, incomes, consumption, and, asset ownership. The summary statistics are given in Tables 1 and 2. Table 1 provides means and standard deviations of the key variables in the data set. The average farmer has 4.3 fragments with a standard deviation of 4.03. The HHI averages .58. Thus there is variation in fragment size within farmer as well as the number of fragments that is critical for assessing the difference between fragment and scale effects. There is also substantial variation in the number of fragment neighbors, with the average number of neighbors being 1.8 and the standard deviation being 2.0. The characteristics of households and details of input use in a more disaggregated manner are provided in Table 2. Yield per acre turns out to be higher for larger farmers compared to small and marginal farmers and input varies across farm sizes. There is evidence of increased mechanization by the large farmers as shown by significantly less use of labor especially during harvesting. Small and marginal farmers as expected use labor intensive methods of cultivation.

One of the innovations of this paper is the structure of the data. Because testing of the hypothesis necessitates that the input and output data be collected for the entire space, we first identify the subdivisions owned or leased in. These subdivisions are revenue units and not necessarily units of cultivation. The next step is to use the information on subdivisions to identify the cultivating units. Such cultivating units are referred to as fragments and all of the input and output data are collected at the fragment level. In Figure 3 we present details of a particular set of subdivisions and fragments. This graph is a representation of one part (section 168) from N village. It is further divided into subdivision such as 168-1, 168-2, 168-4A, 168-5E etc. If a farmer, for example, owned 168-1, 168-2, and, 168-5E, then the most likely method of cultivation is to treat 168-1 and 168-2 as a single unit for farming purposes. Hence this farmer own three subdivisions but cultivates two fragments. The input and output data in this study are collected at the fragment rather than the subdivision level.

The spatial distribution of the data for one of the villages is shown in Figures 4 and 5. Figure 4 shows the overall coverage and how the different fragments are divided up into cultivating units. Shade differences among neighboring cultivating units indicate they are managed by different farmers but each shade may represent multiple farmers. It is evident that although the survey includes all fragments that are farmed by members of the village, not all fragments were surveyed due to the fact that some farms are farmed by people not in the village and others were left fallow. Figure 5 provides an illustration of a subset of the multiple-fragment farmers in N. Fragments are evidently quite dispersed and it is evident that the different fragments of a given farmer have different neighbors. It is also notable that there is substantial variation within farmer in the size of fragments and that the relatively large fragments tend to be cultivated in one area. Table 2 provides a largely descriptive evidence of the fact that larger farmers benefit from greater yields but at the level of fragments, size has an adverse impact on yield. Farmers with larger fragments also tend to use more family labor.

IV. Estimates

We proceed by first by examining simple descriptive regressions of input use per acre by type and yields at the level of the fragment as a function of the HHI, fragment area and total area. Standards errors are grouped at the farmer level. When not controlling for land areas, greater concentration appears to result in a lower per acre non-labor and labor inputs. Controlling for fragment and total area, however, provides a ore mixed picture. A one standard deviation increase in the HHI, for example, controlling for total and fragment area results in a .18 standard deviation increase in inputs per acre. Conversely measures of labor are negatively and significantly related to the HHI, with substantial negative effects. On net there is no significant effect on yield. Fragment sizes operate in the reverse with larger fragments being correlated with lower non-labor inputs and yield and higher family and hired labor inputs. Total area on the other hand is associated with higher non-labor inputs and labor but lower family inputs. At the very least these results suggest that returns to scale are not simple at either the fragment or farm level. Not only are there distinct effects of different measures of plot area but different inputs are differentially affected by these measures.

Table 4 provides random and fixed effects estimates of the reduced form equations () and ().Note that averages across the different plots are constant for a given farmer and are thus excluded from the fixed effects specifications. Results are overall quite consistent across specifications and suggest the presence on net of negative spillovers. First, in column (1) of the estimates an increase in area substantially lowers input use per acre, with an elasticity of -.787. Given an elasticity of negative one would suggest that total inputs per plot are fixed and this magnitude seems overly high and suggests the possible presence of measurement error in area or, more plausibly, an underreporting of differences in total input use across small and large plots. However, taken literally, this figure suggests given (11) that spillovers are negative. Small farmers must compensate for the input use of their neighbors by using more input use themselves but larger farmers are largely protected from effects that only operate at the boundary. There is a negative effect on input use on the maximum distance of a particular plot from other plots but a compensating positive effect of average distance. These distance effects, which are not primarily related to scale economies, may reflect either differences in the cost of transportation or differences in the relative returns to inputs on more distant plots. The former seems relatively transparent. With regard to the latter if, for example, more distant plots are less likely to be depleted one may use less fertilizer on those plots. In either case the fact that the compensating effect works in the other direction is also plausible in the presence of input constraints. That is the fewer inputs one uses on other plots the more inputs one can use on a particular plot.

The HHI effect is positive, with an increase in consolidation of .1 yielding a 7.3 percent increase in inputs in the RE specification. Given () and (), this result is again consistent with negative spillovers. Given total area, if other plots are more concentrated there is less need for inputs to compensate for neighboring effects on other plots and thus there is an increase in available inputs for the fragment of interest.

The effects of characteristics of adjacent fragments operate in principle through the inputs of neighbors. In general one would expect to see, for negative spillovers, that the coefficients on adjacent plots (the second panel of 4) for input equations have the same sign for fragments adjacent to the fragment of interest as those for that fragment itself (the first panel) due to the “arms race” effect. The characteristics of fragments adjacent to the farmers other fragments (the third panel) should have the opposite effect due to the effect on constrained inputs. For positive spillovers the opposite would be true. While adjacent area and HHI are not significantly different from zero in the random effects specifications the distance and average distance effects are significant and indeed of the same sign as their respective coefficients in the first panel. The distance results also holds for the fixed effects specification but in this case adjacent area is positively related to input use, which is not expected in the presence of negative spillovers. The third panel, for which only random effects estimates are possible, matches up well with the negative spillovers model. For example, if farmer *k*'s plots other than *i* are adjacent to larger plots then input use on plot *i* by farmer *k* increases.

The negative effects of fragment *i* area in Table 4 are also consistent with negative spillovers as are the positive effects of the HHI. As there are no direct scale economies greater input use driven by scale or by that shadow price of inputs generates greater yields. The distance effects, however, operate in the opposite direction suggesting that these effects are having direct effects on the production parameters. This result suggests, for example, that more distance fragments are getting fewer inputs at least in part because they are more productive.

The effects of the adjacent fragment's HHI in the second panel also supports negative spillovers. Greater concentration of HHI for farmers of adjacent plots should reduce inputs on those plots and this in turn should reduce inputs on those plots and
[to be completed]

Table 5 and 6 provide estimates of Equations (19) and (20) equations. In particular, we note that these expressions condition on the inputs of both the plots that are adjacent to fragment i of farmer k and the inputs of fragments that are adjacent to k's other fragments. Given that these inputs are determined simultaneously with own inputs one needs in principle to find a source of variation in neighbors' inputs. But neighbors fragment characteristics are excluded from (19) and (20) and are likely to be correlated instruments and thus may serve as appropriate instruments.

The effect of the inputs on neighboring plots are also positive. Controlling for own area, a 10 percent increase in neighbor's inputs results in a 1.7 percent increase in own inputs in all four specifications. This positive effect is exactly what is predicted by () in the presence of negative spillovers—one needs to compensate for the neighbors' activities by using more inputs on one's own fragments. The average inputs across other plots should in principle be negative because greater demand for inputs on other plots should raise the cost of inputs on a particular plot but the effect is essentially zero. This result in itself might fit with the idea that inputs are freely available at a given price, but it is hard to square that conclusion with the HHI effect noted above.

We now turn to Table 6, which carries out the same set of analyses for yields. As shown in equations () through () the signs on yields correspond with the signs on yields (because greater inputs per unit area, all else being equal, produce greater yields). The one exception is neighbors' inputs because, to a linear approximation, farmers fully offset the consequences of neighbors' inputs by increasing own inputs. As anticipated, the area and HHI effects in Table 6 correspond to those in Table 5, as should be expected under negative spillvoers. Larger farms have lower yields and farmers with more concentrated holdings have higher yields. The distance effects, however, work in the opposite direction from those in the inputs. The positive own distant effect seems consistent with the idea that the distance effects reflect differences in productivity of more distant plots rather than the idea than differences in the cost of transporting inputs. However, it is not clear that this should yield positive average distance effects.

The effects of neighbors' inputs on yields are of more direct interest and, interestingly, differ between the random and fixed effects specifications. In the random effects specifications neighbors' (to *i*) inputs have a significant positive effect on yields which offsets the negative effects of neighbors' across all plots. The latter effect is consistent with negative spillovers in the sense that an increase in neighbors' inputs on other plots but this is puzzling given that there was no evidence of such an effect in terms of measured inputs use. The former, however, is not consistent with the linear model. Interestingly, however, in the fixed effects specification the effect of inputs of neighbors to i is indeed zero as predicted by (). Overall, then, the yield equation seems to further support the idea of negative spillovers..

The final column in Table 6, incorporates further structure. In particular, it conditions on both own and neighbors inputs using a random effects specification and using the HHI to instrument own inputs (recall from () that HHI only enters through the input price. While as expected, own inputs increase yields and consistent with the idea that the negative effects of area on inputs is a consequence of the

V. Conclusion

This paper has shown that negative spillovers in input use are sufficiently important to lead to significant negative returns to scale in land. On its face this suggests that further consolidation of land would be unwarranted. But this conclusion would be misleading because decreasing returns in terms of input use generate positive effects of consolidation on a per fragment basis. On net it appears that consolidation would tend to not only increase average yields but also decrease yield variance.

Table 1:Means and Standard Deviations

Variable	Mean	Std. Dev.
HHI Index	0.576	0.324
Number of Fragments	4.294	4.030
Distance to Largest Fragment	2.220	3.702
Avg. Distance to Largest Fragment	2.220	2.542
Ln Cultivated Area (Fragment)	-0.316	1.089
Ln Cultivated Area (all Fragments)	0.975	1.132
Head Yrs Schooling	7.271	4.199
Head Female	0.168	0.374
Mean Height of Adults	5.241	0.301
Number of Adults	3.374	1.483
Mean Number of Kids	0.640	0.862
Number Neighbors	1.794	2.018
Log Value Family Labor	6.073	2.314
Log Value Hired Labor	8.208	0.874
Log Value of Inputs	7.194	1.030
Log Quantity Pesticide	3.817	1.056
Log Quantity Fertilizer	3.218	2.219
Log Yield	7.288	0.440

Table 2: Household Characteristics

Variable	Nelpathur	Thirunagari	Small farmers	Marginal farmers	Medium farmers	Large farmers
	Mean (Std. Dev.)					
Male headed households	0.78 (0.41)	0.89 (0.31)	0.81 (0.39)	0.87 (0.34)	0.80 (0.40)	0.88 (0.33)
Head age	56.66 (11.86)	49.00 (8.82)	52.49 (10.59)	53.24 (11.72)	60.47 (8.64)	63.25 (17.52)
Head education	6.66 (3.90)	6.73 (4.24)	6.16 (3.96)	7.61 (3.69)	7.47 (4.28)	9.25 (4.77)
Family size	3.93 (1.70)	3.53 (1.53)	3.48 (1.50)	4.27 (1.84)	4.75 (1.76)	4.23 (1.27)
No. of earners	1.75 (0.90)	1.77 (0.88)	1.72 (0.86)	1.89 (1.01)	1.87 (0.89)	1.50 (0.51)
Members(<=14)	0.65 (0.88)	0.59 (0.80)	0.55 (0.81)	0.81 (0.91)	0.85 (0.98)	0.55 (0.71)
Members (>14)	2.86 (1.54)	2.80 (1.18)	2.68 (1.31)	3.15 (1.56)	3.36 (1.78)	2.83 (0.84)
Average age of the household	37.36 (11.51)	34.79 (9.57)	36.88 (11.36)	33.59 (8.30)	38.32 (12.68)	39.15 (8.02)
Production & inputs per acre						
Yield (Kilograms)	1469.63 (361.38)	1695.80 (267.18)	1297.58 (284.15)	1226.98 (180.81)	1471.89 (252.81)	1634.10 (353.64)
Labor for harvesting (units)	7.62 (3.28)	8.51 (3.09)	8.93 (3.20)	6.18 (2.39)	6.16 (1.78)	4.71 (1.91)
Labor for land preparation(units)	3.48 (2.85)	3.24 (3.39)	3.68 (3.69)	2.68 (0.80)	2.86 (0.44)	3.11 (0.50)
Fertilizer & manure (Qty)	210.77 (415.03)	132.96 (40.60)	169.65 (210.27)	106.14 (68.08)	170.36 (96.83)	237.54 (584.84)
Bullocks & Machinery (frequency)	17.59 (29.08)	10.97 (9.08)	15.73 (26.24)	15.82 (19.23)	10.84 (8.57)	4.41 (1.38)
Irrigation (including labor cost) (Rs)	48.40 (33.72)	56.68 (27.84)	15.11 (6.14)	36.19 (13.57)	15.93 (7.52)	62.60 (32.04)
Other labor(Rs)	96.63 (37.85)	117.61 (36.62)	113.49 (41.89)	95.95 (21.95)	72.11 (14.13)	66.92 (10.77)

Table 3: Fragmentation, Landholding and Area

	Inputs	Inputs	Hired Labor	Hired Labor	Family Labor	Family Labor	Yield	Yield
HHI	-24.329	83.992	-178.882	-116.379	-0.276	-0.873	-0.324	0.173
	[2.79]	[5.59]	[1.59]	[1.23]	[5.13]	[9.01]	[2.11]	[1.08]
Ln Fragment Area		-38.05		59.628		0.105		-0.137
		[5.23]		[1.39]		[6.15]		[2.61]
Ln Total Area		34.343		51.758		-0.23		0.174
		[5.09]		[2.12]		[9.03]		[2.57]
Constant	28.521	-79.179	155.652	88.057	0.403	1.004	0.789	0.293
	[3.84]	[5.26]	[1.69]	[1.23]	[8.11]	[10.15]	[6.26]	[2.08]
Observations	2135	2135	2140	2140	2140	2140	2085	2085
R-squared	0.01	0.13	0	0.01	0.02	0.08	0.01	0.02

Robust t statistics in brackets

Table 4: Random (RE) and Fixed Effects (FE) Estimates of Reduced Form Input and Yield Equations Incorporating Own and Neighbors' (N) Fragment Characteristics

	Ln Input	Ln Yield	Ln Input	Ln Yield
	RE	RE	FE	FE
Main Fragment i Variables				
Ln Area	-0.787	-0.037	-0.975	-0.047
	[52.62]	[4.02]	[68.15]	[3.68]
Dist Max	-0.096	0.011	-0.109	0.01
	[21.38]	[3.37]	[28.84]	[3.04]
HHI	0.726	0.278		
	[7.75]	[7.24]		
Av Dist	0.12	-0.013		
	[9.98]	[2.35]		
Variables for Neighbors of Fragment i				
Ln N Area	0.061	-0.067	0.132	-0.098
	[1.21]	[1.79]	[3.10]	[2.57]
N Dist Max	-0.017	0	-0.014	-0.002
	[1.81]	[0.07]	[1.80]	[0.26]
N HHI	0.025	0.42	-0.002	0.267
	[0.23]	[5.79]	[0.02]	[3.03]
N Av Dist	0.071	0.032	0.047	0.014
	[4.90]	[3.09]	[3.71]	[1.26]
Variables for Neighbors of all Fragments				
Av Ln N Area	0.354	0.202		
	[3.25]	[3.56]		
Av N Dist	0.023	0.01		
	[1.27]	[0.97]		
Av N HHI	-0.161	-0.264		
	[1.28]	[3.59]		
Av N Av Dist	-0.067	-0.06		
	[2.53]	[4.23]		
Observations	2140	2140	2140	2140
Groups	1050	1050	1050	1050
R-squared	0.84	0.04		

Absolute value of z statistics in brackets

Also includes controls for number of neighbors, soil characteristics, canal water, and farmer characteristics

Table 5: Least Squares (LS) and Instrumented (IV) Random (RE) and Fixed Effects (FE) Estimates of Input Decision Rules Incorporating Own Fragment Characteristics and Neighbors' (N) Inputs

	LS-RE	IV-RE	LS-FE	IV-FE
Main Fragment i Variables				
Ln Area	-0.789	-0.841	-0.974	-0.97
	[54.25]	[62.23]	[69.16]	[67.89]
Dist Max	-0.097	-0.1	-0.108	-0.109
	[22.22]	[25.80]	[29.50]	[29.37]
HHI	0.708	0.777		
	[8.01]	[7.84]		
Av Dist	0.121	0.123		
	[10.32]	[9.71]		
Inputs of Neighbors of Fragment i and All Fragments				
Ln N Input*	0.169	0.167	0.092	0.17
	[7.92]	[4.36]	[4.28]	[3.57]
Av Ln N Input*	0.003	0.006		
	[0.36]	[0.66]		
Other Fragment i/Farmer k Variables				
No N	1.098	1.102	0.628	1.181
	[7.04]	[4.05]	[3.97]	[3.47]
Numb N	-0.024	-0.027	-0.031	-0.03
	[2.66]	[3.19]	[3.58]	[3.45]
Canal	0.078	0.049	-0.068	-0.069
	[1.60]	[1.05]	[1.26]	[1.28]
Sandy	0.189	0.181	0.168	0.161
	[4.66]	[4.59]	[3.60]	[3.42]
Head Ed	0.014	0.014		
	[2.62]	[2.40]		
Av Adult Ht	0.098	0.09		
	[1.58]	[1.35]		
Constant	4.589	4.598	6.565	6.018
	[11.88]	[9.98]	[40.84]	[17.77]
Observations	2140	2140	2140	2140
Number of group(v3n year)	1050	1050	1050	1050
R-squared			0.84	

Absolute value of z statistics in brackets
*Endogenous variable instrumented with N's Area, Distance, HHI, Av Distance

Table 6: Least Squares (LS) and Instrumented (IV) Random (RE) and Fixed Effects (FE) Estimates of Yield Equations and Productions Function Incorporating Own Fragment Characteristics and Neighbors' Fragment Inputs

	Ln Yield LS-RE	Ln Yield IV-RE	Ln Yield LS-FE	Ln Yield IV-FE	Ln Yield LS-RE
Main Fragment i Variables					
Ln Area	-0.025	-0.032	-0.046	-0.048	0.031
	[2.80]	[3.49]	[3.62]	[3.74]	[2.95]
Dist Max	0.009	0.009	0.008	0.008	0.017
	[2.74]	[2.70]	[2.48]	[2.52]	[4.84]
HHI	0.354	0.331			
	[9.88]	[8.60]			
Av Dist	-0.011	-0.013			-0.044
	[2.02]	[2.27]			[8.64]
Inputs on Fragment i, Neighbors of i and All Neighbors					
Ln Input*					0.066
					[6.06]
Ln N Input*	0.053	0.089	0.015	-0.033	0.003
	[4.03]	[4.13]	[0.76]	[0.78]	[0.58]
Av Ln N Input*	-0.02	-0.016			-0.026
	[3.51]	[2.60]			[4.59]
Other Fragment/Farmer Variables					
No N	0.408	0.672	0.127	-0.214	
	[4.34]	[4.44]	[0.89]	[0.70]	
Numb N	0.013	0.011	0.009	0.009	
	[2.38]	[2.03]	[1.19]	[1.11]	
Canal	-0.191	-0.189	-0.117	-0.116	-0.218
	[6.58]	[6.54]	[2.42]	[2.40]	[7.37]
Sandy	-0.042	-0.054	-0.077	-0.072	-0.089
	[1.90]	[2.35]	[1.83]	[1.71]	[3.96]
Head Ed	0	-0.001			-0.003
	[0.02]	[0.24]			[1.19]
Av Adult Ht	0.081	0.081			0.04
	[2.66]	[2.62]			[1.28]
Constant	6.521	6.271	7.249	7.587	6.851
	[33.23]	[27.38]	[50.15]	[25.00]	[37.61]
Observations	2140	2140	2140	2140	2140
Number of group(v3n year)	1050	1050	1050	1050	1050
R-squared			0.03		

Figure 1:
Simulated Geographical Pattern of Inputs when Neighbor Use Lower Inputs

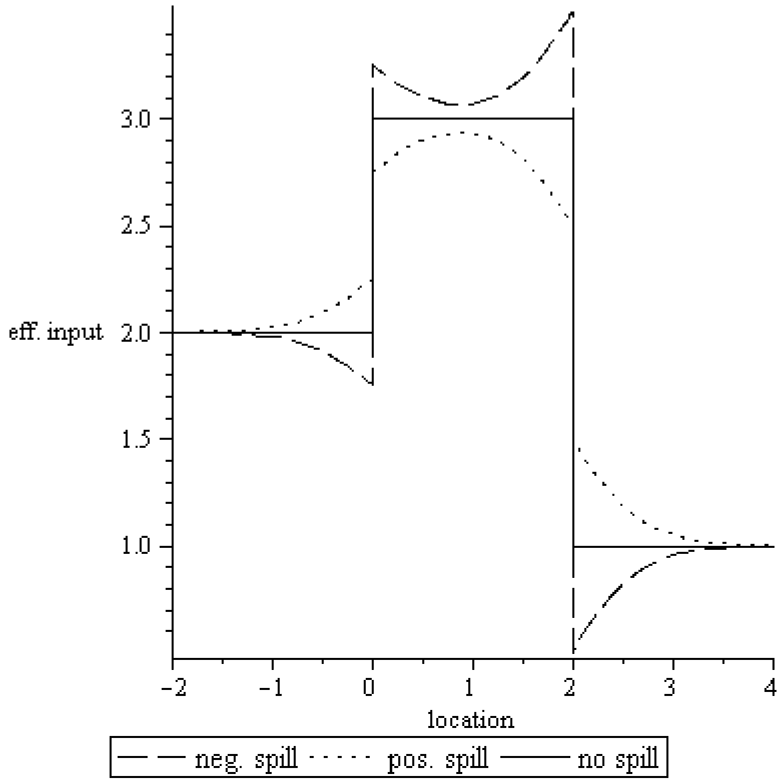

Figure 2:
Simulated Geographical Pattern of Effective Input Use When Neighbors Use Higher Inputs

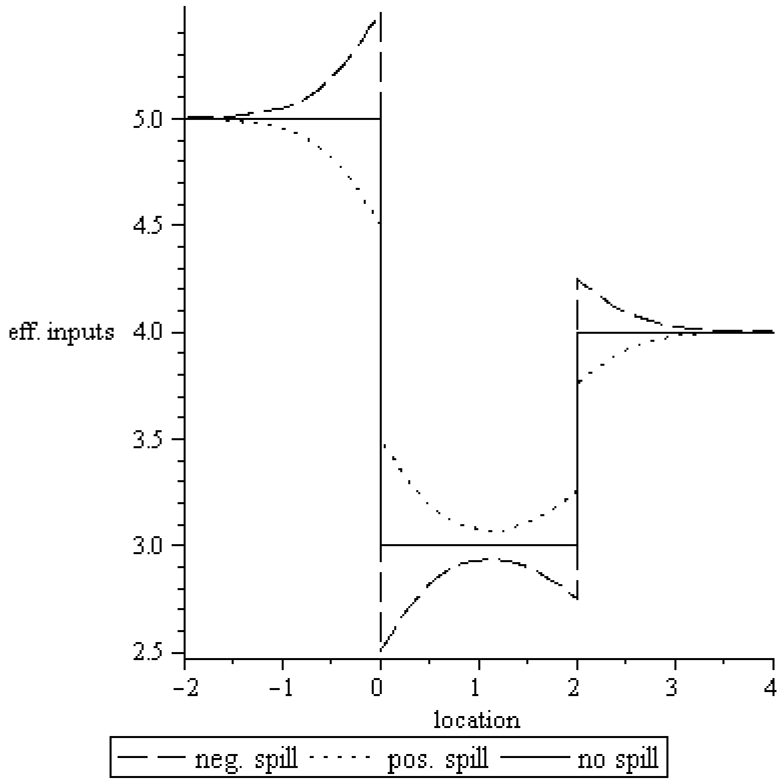

Figure 3: Subdivisions and Fragments

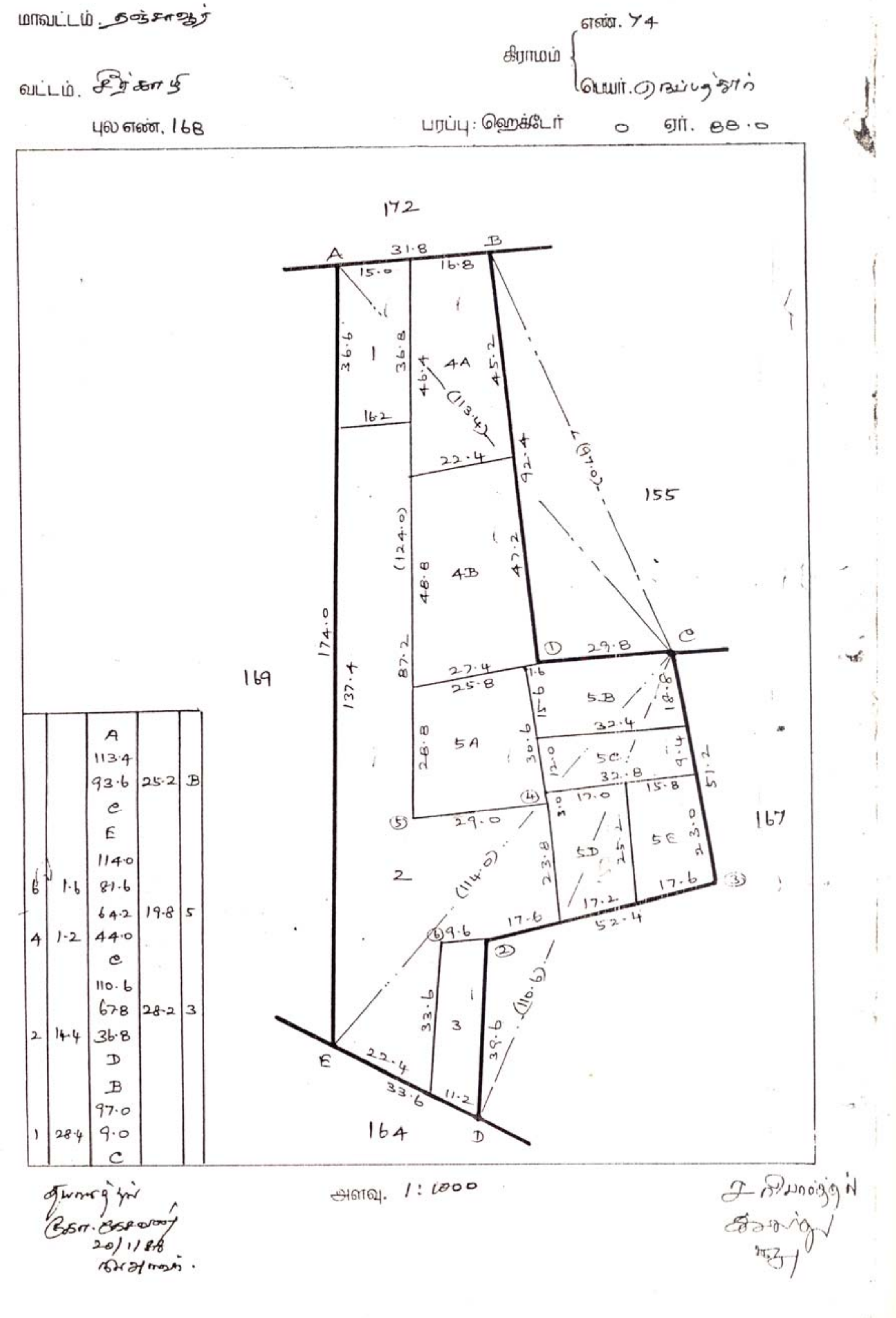

Figure 4: Cultivating Areas in Village N

Figure 5: Select Farmers with Multiple Cultivated Areas in Village N