205 604

FARM SIZE AND LAND USE INTENSITY IN INDIAN AGRICULTURE

1991

India - Agriculture

Rajan K. Sampath
Professor
Department of Agricultural and Resource Economics
Colorado State University
Fort Collins, CO 80523 USA

AAEA 1991

#20544

Farm Size and Land Use Efficiency In Indian Agriculture

I. Introduction: Recently, Bhalla and Roy (1988) subjected one of the popular stylized facts of agriculture in developing countries, namely, the strong inverse relationship between farm size and land productivity, to a closer scrutiny and found that once proper account is made of exogenous land quality variables, the inverse relationship is observed to weaken, and in many cases, to disappear. They argued that "exclusion of land quality, a variable negatively correlated with farm size results in the coefficient of land being biased downward. Thus, the stylized fact of a negative relationship between farm size and farm productivity may in large part have been due to omission of soil quality variables from the estimated equations." In this context, they pointed out also that "if . . .land quality differences are a major (and perhaps sole) explanation of the inverse relation then a complete re-evaluation of the farm size productivity debate is called for." This note is written in this spirit and throws some doubts on the validity of one of the popular hypotheses often tested by researchers in the context of the discussion of the inverse relationship between farm size and land productivity, namely, the inverse relationship between farm size and land use intensity [Cline, (1970), Bharadwaj, (1974), Berry and Cline (1979) and references therein].

II. Data Source: The main body of farm sizewise/statewise, cross-section data for India comes from a survey conducted by the National Sample Survey (NSS) and published in the two NSS reports on survey results in the October 1984 and July-October 1985 issues of "Sarvekshana" Journal of the NSS organization. The data reported in these publications are based on a careful survey on the use of irrigation in rural areas. The reference period for the survey was the agricultural year July 1975 - June 1976. The survey covered the whole rural India except for a few small states and union territories.

In the Central Sample, 8,304 villages were involved and the total number of households actually surveyed was 88,046. The publications mentioned above present farm sizewise and statewise data on the variables used in this study. One advantage of this data set is that it covers entire India and the period of the survey is neither the pre-green revolution period as the early farm management studies were upon which previous studies were based [Sen (1964), Bharadwaj (1974), and references therein] nor the initial green revolution period, such as that of the Berry and Cline study but rather at the matured state of the green revolution.

III. <u>Farm size and Land use Intensity</u>: We study the relationship between farm size and land use intensity in terms of two different definitions of land use, namely, Net Area Sown (N) and Gross cropped area (G). N and G variables measure the extensive and intensive use of land, respectively. Net area sown (NAS) measures the extensive use of land better than the land area operated (LAO) because LAO includes in its definition lands that are cultivable but lying unused such as the permanent fallows, cultivable waste land, etc. that produce no crops and consequently no income. NAS by excluding fallow land estimates properly the extensive use of land that produces crop and consequently income for the farmer. Since our interest is in estimating the relationship between farm size (which is measured in terms of size of land operated) and land use, the appropriate definition of extensive use of land is NAS. The gross cropped area (GCA) measures the intensity of land use and differs from the NAS in that it counts as twice and thrice all lands that are double and triple cropped, respectively. Of course, these two variables NAS and GCA are used as proxies to measure the intensity of land use because an ideal definition of land use intensity would go beyond the spatial aspect to include cultivation practices, use of fertilizer, etc. Finally, we define the size of farm in terms of the size of operational holding of the farm. In defining these three variables, we followed the current practice in the literature [Cline (1970), Bharadwaj (1974), Sampath and Gopinath (1978), Berry and Cline (1979), and references therein)].

An operational holding consists of two types of land, namely, irrigated and unirrigated that have significantly differing impacts on land use and productivity and usually different farms have different proportions of their holdings irrigated. Thus, while representing the impact of farm size on land use intensity, it is very critical to take into account this distinction between irrigated (I) and unirrigated land (U). Otherwise the analyses will be seriously flawed. By adding linearly I and U to define total area (T), previous studies have implicitly assumed that one unit of land with irrigation facilities has the same potential for cropping intensity as one unit of land with no irrigation facilities at all, which clearly is not so. Availability of irrigation is a prime determinant of land use in its both intensive and extensive forms. In the absence of irrigation, Indian rainfall conditions are not conducive for double and triple cropping of land in a year. In arid and semi-arid regions, irrigation helps in bringing much of cultivable wasteland under plough. Thus irrigation makes possible higher extensive and intensive use of land in three ways: (1) it helps bring new lands that are otherwise fallow/barren under cultivation and thereby increase the net sown area of a farm; (2) it makes possible growing of crops during the dry season and thereby increasing the land use intensity; and (3) it makes possible growing of shorter duration crops (which cannot be grown without adequate and dependable water supply to the crop that irrigation makes possible) and thereby makes multiple cropping possible [GOI (1972), GOI (1976), Seckler and Sampath (1985), Pal (1985), Dhawan (1988)]. Thus linear aggregation of I and U leads to aggregation (misspecification) bias in the estimation of parameters and testing of hypotheses.

Our objective in this paper is to estimate the relationship between land use and farm size in terms of two approaches. The first approach ignores the distinction between I and U and the second approach takes the distinction explicitly into account in the analysis. This way we will be able to show the problems that misspecification creates in the study of farm size and land use.

IV. Functional Forms: The relationship between farm size (in terms of total land operated) and land use (in terms of G and N) can be represented by one of the following functional forms:

$$N \text{ or } G = a_0 + a_1 T \qquad \text{...(1a)}$$

$$N \text{ or } G = b_0 + b_1 T + b_2 T^2 \qquad \text{...(1b)}$$

$$N \text{ or } G = A\, T^a \qquad \text{...(1c)}$$

or taking the distribution between irrigated and unirrigated land into account, we can restate the above relations as

$$N \text{ or } G = \alpha_0 + \alpha_1 I + \alpha_2 U \qquad \text{...(2a)}$$

$$N \text{ or } G + \beta_o + \beta_1 I + \beta_2 I^2 + \beta_3 U + \beta_4 U^2 \qquad ...(2b)$$

$$N \text{ or } G = B I^{\alpha} U^{\beta} \qquad ...(2c)$$

The conventional procedure is to choose that functional form that explains the variations in N or G the best. If 1a or 2a turns out to be the best equation, then that would indicate a size neutral impact on land use since according to these equations every unit increase in T or I or U leads to a constant increase in N or G. If 1b or 2b turns out to be the best equation, then it would indicate economies or diseconomies of scale depending on whether b_2, β_2 and β_4 are negative or positive respectively; and finally if 1c or 2c turns out to be the best equation, then it would indicate diseconomies or economies of scale depending upon whether "a" and $(\alpha + \beta)$ are less than or greater than unity.

The implications of diseconomies of scale for land reform are fairly obvious. In all the above-mentioned cases except 2c, any redistribution of land from large to small farms would improve land use. For example, let us represent the (G) function of two farms, one small (P) and the other large (R) by

$$G_P = b_o + b_1 T_1 - b_2 T_1^2 \qquad ...(3)$$

and

$$G_R = b_o + b_1 T_2 - b_2 T_2^2 \qquad ...(4)$$

If we transfer, t (>o) amount of land from R farm to P farm, then total G will change from G_1 to G_2 where

$$G_1 = 2b_o + b_1 (T_1 + T_2) - b_2 (T_1^2 + T_2^2) \qquad ...(5)$$

and $$G_2 = 2b_o + b_1 [(T_1 + t) + (T_2 - t)] - b_2 [T_1 + t)^2 + (T_2 - t)^2]$$
$$= 2b_o + b_1 [(T_1 + T_2) - b_2 (T_1^2 + T_2^2 + 2t^2 + 2 T_1 t - 2 T_2 t)] \qquad ...(6)$$

It can be easily shown that

$$G_2 > G_1 \qquad ...(7)$$

whenever

$$t < (T_2 - T_1) \qquad ...(8)$$

which will be easily met since it only requires that the amount of land that is transferred be such that it does not make the small farm become larger than the large farm after the transfer. We can call this economic status preserving transfer condition. Similar is the result in the case of equation 2b so long as the amount of land transferred from the large to the small farm is such that the large farm still has more of both I and U than small farm after the transfer, or transfer of land is done only from that category of land in which the large farm has more land than the small farm. In the case of equation 1c, any transfer of land from large to small farm [$t < (T_R - T_P)$] will improve overall G so long as "a" is less than unity. But the case of equation 2c is a bit more complicated since the relationship is multiplicative. Given the small and large farm G functions as

$$G_P = A I_p^{\alpha 1} U_P^{\alpha 2} \qquad ...(9)$$

and

$$G_R = A I_R^{\alpha 1} U_R^{\alpha 2} \qquad ...(10)$$

any transfer of I or U from large to small farms will lead to increase in G only if

$$\frac{dG_P}{dI_p} > \frac{dG_R}{dI_R} \qquad ...(11)$$

or

$$\frac{dG_P}{dU_p} > \frac{dG_R}{dU_R} \qquad ...(12)$$

Simplifying (11) and (12) results in the following conditions:

$$I_p > \left(\frac{U_R}{U_P}\right)^{\frac{\alpha 2}{\alpha 1 - 1}} I_R \qquad ...(13)$$

and

$$U_P > \left(\frac{I_R}{I_P}\right)^{\frac{\alpha_2}{\alpha_2 - 1}} U_R \qquad \text{...(14)}$$

If the above conditions are violated, then, even if $o < \alpha_1, \alpha_2 < 1$, any transfer of land from large to small could result in reduction in total gross cropped area resulting in lower land use.

The purpose of the above discussion is to show clearly how critical it is to arrive at the proper functional form so that correct inferences could be drawn with regard to land use across farm size groups and the implications of government land reform policies for land use. Further, even if land redistribution leads to improvement in land use efficiency, the question arises as to whether that is the optimal policy for the government to follow since according to equations 2b and 2c, there is another politically and administratively more easier way to improve land use and reduce inequity simultaneously through irrigation distribution. In almost all the LDCs surface irrigation potential constitutes more than half of the total irrigation potential and it is totally under government control and ownership for its development and distribution. To the extent irrigation improves the extensive and intensive use of land, the government can improve land use and equity through its irrigation distribution policy.

V. Results and Discussion:

Table 1
Farm Size and Net Area Sown

The estimated values of regression parameters and their standard errors

Equ. No.	Dep. Var.	Constant term	T	T^2	I	I^2	U	U^2	$\bar{R}^2$/(F)
1	N	0.1097	0.8243*						.947
			(0.0175)						(2212)
2	N	0.0427	0.8700*	-0.003					.946
			(0.0663)	(0.0042)					(1102)
3	N	0.0049			1.0440*		0.6878*		.963
					(0.0325)		(0.0232)		(1622)
4	N	-0.1071			1.1325*	-0.008	0.7427*	-0.0062	.963
					(0.0965)	(0.0070)	(0.0848)	(0.0063)	(828)
5	ln N	-0.3633	0.2245*						0.7188
			(0.0126)						(320)
6	ln N	-0.3882			0.2767*		0.1920*		0.7242
					(0.276)		(0.0197)		(166)
			ln T		ln I		ln U		
7	ln N	-0.1096	0.9688*						.987
			(0.0099)						(9665)
8	ln N	0.6393			0.7219*		0.3125*		.987
					(0.0188)		(0.0117)		(4631)
					ln (I/U)				
9	ln (N/U)	0.6411			0.6801*				.965
					(0.0117)				(3401)

Note: Total number of observations is 127. * = Significant at 99% confidence level.

Table 1 presents the regression results pertaining to nine equations relating farm size with net area sown. Since the first four equations have the same dependent variable, we can choose among them the best equation solely on the basis of $\bar{R}^2$ [Rao and Miller (1972)]. Under this criterion, variations in N are better explained by variations in I and U simultaneously than by variations in T (=I+U) alone as shown by their respective $\bar{R}^2$. Also, since variations in I and U are not strictly in proportion to increase in farm size (T) and to the extent the two types of land lead to differential impacts on net sown area, as we will show shortly that this in fact is so, the parameters estimated by equation 1 and 2 will have aggregation (misspecification) bias and as such they are not reliable estimates.

Equations 5 and 6 provide estimates of semi-log equations and equations 7 and 8 provide estimates of double-log equations. Using the residual sum of squares (RSS) of linear equations 1 and 3 and semi-log equations 5 and 6, we tested the null hypothesis that irrigated and unirrigated lands have the same effect on N. The computed F statistic rejected the null hypothesis for both linear and semi-log equations at 99 percent and 95 percent confidence levels, respectively.

Equations 7 and 8 provide estimates of double-log regression equations. Double-log equations are clearly superior to semi-log equations 5 and 6 in terms of explanatory power ($\bar{R}^2$). Between 7 and 8, even though both the equations are equally good in terms of $\bar{R}^2$, because of aggregation bias involved in the estimation of equation 7, equation 8 should be preferred. Since logarithmic equations' $\bar{R}^2$ cannot be directly compared with linear equations' $\bar{R}^2$ as the dependent variables are different, we estimated the residual sum of squares (RSS) of the two equations using the standardized values of N (dividing N by its geometric mean). Once again, in terms of the minimum RSS criteria, we found the logarithmic equation as being the superior equation. To save space, we do not report details on this exercise.

According to equation 7, there exists diseconomies of scale since the estimated farm size elasticity parameter of N is statistically significantly less than unity at 99% confidence level. In contrast, according to equation 8, the homogeneity of the N function is equal to 1.0344 (=0.7219 + 0.3125). To test whether 1.0344 is greater than or equal to unity, we estimated equation 9 by imposing the linear homogeneity assumption on equation 8. The "F" test based on the RSS of equations 8 and 9 rejected the null hypothesis that equation 8 is linear homogeneous at 99% confidence level. In other words, 1.0344 is statistically significantly greater than unity indicating economies of scale. This inference is diametrically opposite to the inference we drew on the basis of equation 7 that the homogeneity of the function is less than unity indicating diseconomies of scale.

We know "a priori" that equation 7 is misspecified since it treats land with and without irrigation facilities as being equal in potential for land use which clearly is not so as we have shown with the help of the semi-log equations 5 and 6. Thus, this contradicts the earlier findings in the literature that the "extensive use" of land is negatively related to farm size (Bharadwaj, Berry and Cline, and references therein). The reason behind this finding is easy to see. By defining farm size strictly in terms of total land area, earlier studies assumed that the effect of lands with and without irrigation facilities on land use to be the same leading to aggregation bias in the estimated values of the farm size elasticity parameter resulting in the erroneous inference that there are diseconomies of scale in the "extensive" use of land. Since lands with and without irrigation facilities are neither equal in their effects on land use nor their distribution is proportional across farm size groups, it is desirable to define farm size in terms of multiple attributes and analyze the effect of size in terms of the homogeneity of the overall function rather than in terms of a single attribute such as land size alone.

Table 2 provides regression summary statistics relating to farm size and gross cropped area. As in Table 1, here too we have nine equations and the only difference is in the dependent variable, which is gross cropped area representing the "intensive use" of land. Once again comparing the first four equations we find our regression equations 3 and 4 in terms of I and U fare better in explanatory power than the conventional equations 1 and 2 in terms of T; and as we noted earlier, the parameters in equations 1 and 2 are not reliable since they are subject to misspecification bias. Adding quadratic terms do not improve $\bar{R}^2$ substantially for any of the equations, though in equation 4 the parameter for I^2 is statistically significant and positive indicating economies of scale in the use of irrigable land for cropping, but it is not numerically very high. Equations 3 and 4 clearly show the differential impacts of land with and without irrigation facilities on gross cropped area. According to equation 3, land with irrigation facilities contributes more than two times as much to G as land without irrigation facilities.

Table 2
Farm Size and Gross Cropped Area

The estimated values of the regression parameters and their standard errors

Equ. No.	Dep. Var.	Constant term	T	T^2	I	I^2	U	U^2	$\bar{R}^2$/(F)
1	G	0.3292	1.0207*						.887
			(0.0325)						(984)
2	G	0.1887	1.1167*	-0.0063					.886
			(0.1230)	(0.0078)					(491)
3	G	0.0456			1.6150*		0.6516*		.962
					(0.0419)		(0.0299)		(1604)
4	G	0.2588			1.3149*	0.0244*	0.7038*	0.0003	.964
					(0.1219)	(0.0089)	(0.1071)	(0.0080)	(850)
5	ln G	-0.0095	0.2142*						0.7045
			(0.0123)						(298)
6	ln G	-0.0423			0.28198*		0.1705*		0.7184
					(0.0123)		(0.0191)		(161)
			ln T		ln I		ln U		
7	ln G	0.2298	0.9174*						.972
			(0.0139)						(4383)
8	ln G	0.9245			0.7405*		0.2586*		.972
					(0.0260)		(0.0161)		(2177)
					ln (I/U)				
9	ln (G/U)	0.9244			0.7416*				.948
					(0.0156)				(2253)

Note: Total number of observations is 127.

Equations 5 and 6 provide estimates of semi-log equations and 7 and 8 provide estimates of double-log equations. Using the RSS of linear equations 1 and 3 and semi-log equations 5 and 6, we tested the null hypothesis that irrigated and unirrigated lands have the same effect on G. The computed F Statistic rejected the null hypothesis for both linear and semi-log equations at better than 99 percent confidence levels.

Comparison of double log equations 7 and 8 with semi-log equations 5 and 6 clearly shows that double log specification is far superior in terms of explanatory power ($\bar{R}^2$). Once again, comparison between logarithmic equations 7 and 8 shows that though the conventional representation of farm size in terms of T has the same explanatory power as our representation in terms of I and U, the conventional equation is misspecified and as such the farm size elasticity parameter is not reliable. According to the conventional equation 7, there exists diseconomies of scale since the value of the farm size elasticity parameter (0.9174) is statistically significantly less than unity at 99% confidence level. In contrast, our equation 8 in terms of I and U has a homogeneity of (0.9991), which is not statistically different from unity in terms of the F test at 99% confidence level. Also, analysis of linear and log-linear equations in terms of standardized values of G (dividing G by its geometric mean) further revealed that log-linear is superior to linear specification. Once again, in the "intensive use" of land also, defining the size of farm in terms of two attributes of land namely land with and without irrigation facilities leads to inferences concerning the relationship between farm size and land use contrary to what is reported in the empirical literature.

Our results show no diseconomies of scale either in the "extensive" or "intensive" use of land per se when the size of farm is defined in terms of its two major attributes, namely, land with and without irrigation facilities.

VI. Conclusions

In this paper, we showed that the conventional way of defining farm size in terms of physical size of operational holding would lead to misleading inferences since mere physical size does not distinguish between land that has irrigation facilities and those that do not. Since irrigation makes a substantial difference to the level of both extensive and intensive use of land, the size of farm should be defined at least in terms of the two attributes of land, namely, lands with and without irrigation facilities. We showed that using the conventional definition of farm size in the regression equation leads to biased estimates of elasticities of land use with respect to size resulting in the misleading inference that there are diseconomies of scale in the use of land. In contrast, with the same data set, when we used our broader definition of size, it resulted in the inference that there are no diseconomies of scale in land use.

REFERENCES

Berry, A.R. and Cline, William R., *Agrarian Structure and Productivity in Developed Countries*, Johns Hopkins University Press, Baltimore, 1979.

Bhalla, Surjit S., and Prannoy Roy, *Misspecification in Farm Productivity Analysis: The Role of Land Quality.* Oxford Economic papers, March 1988, 40:(1)53-73.

Bharadwaj, K., *Production Conditions in Indian Agriculture: A Study Based on the Farm Management Surveys*, Cambridge:Cambridge University Press, 1974.

Cline, W.R. *Economic Consequences of a Land Reform in Brazil*, North Holland, 1970.

Dhawan, B. V., *Irrigation in India's Agricultural Development*, Sage Publications, New Delhi, 1988.

Government of India (GOI), *Report of the Irrigation Commission*, Volume I, Ministry of Irrigation and Power, 1972.

Government of India (GOI), *Report of the National Commission on Agriculture*, Part V, Resource Development Ministry of Agriculture and Irrigation, 1976.

Pal, S. P., *Contribution of Irrigation to Agricultural Production and Productivity*, National council of Applied Economic Research, New Delhi, 1985.

Rao, P.M. and R.L. Miller, *Applied Econometrics*, published by Prentice-Hall of India, Private Limited, New Delhi, 1972.

Sampath, R.K. and Gopinath, *Land Distribution in India: Its Nature and Economic Implications*, Margin, Vol. II No. 2 April pp 17-34, 1978.

Seckler, D. W., and R. K. Sampath, *Production and Poverty in Indian Agriculture*, ISARD Working Paper No. 1 (December), 42 pp., Colorado State Univ., Fort Collins, 1985.

Sen, A.K., *Size of Holdings and Productivity*, Economic Weekly, Annual Number, 1964.