PN-ABE-012
64101

FARM SIZE AND EFFICIENCY IN IRRIGATED AGRICULTURE:
The Case of the Zimbabwe Communal Areas

by

Peter C. Bloch

Land Tenure Center
University of Wisconsin-Madison

July 1988

,1'

FARM SIZE AND EFFICIENCY IN IRRIGATED AGRICULTURE:
The Case of the Zimbabwe Communal Areas

Peter C. Bloch

1. The Problem

One of the principal socioeconomic issues in irrigation design is the size of holdings to be allocated to individual farm households. The arguments for larger versus smaller plot sizes are summarized in Bloch et al.[1]; the major ones concern equity -- allocating a given amount of irrigable land to larger or smaller numbers of poor farmers -- and efficiency -- obtaining the maximum economic returns from a given resource outlay.

The equity arguments can be made without reference to the data, because they simply point out that a given area divided among a larger number of farmers is a more even distribution of land. This is not true of the efficiency arguments, however, which are by their very nature empirically based. In the principal published work that has addressed this issue, Berry and Cline[2] summarize a large number of studies in many parts of the world, finding evidence of a negative correlation between

---

1. Peter C. Bloch; Lucie Colvin Phillips; James C. Riddell; Jayne L. Stanning; Thomas K. Park. Land Tenure Issues in River Basin Development in Sub-Saharan Africa. Research Paper No. 90, University of Wisconsin Land Tenure Center, April 1986, esp. pp. 9-12.

2. R. Albert Berry and William R. Cline. Agrarian Structure and Productivity in Developing Countries. Baltimore: Johns Hopkins University Press, 1979.

farm size and productivity. They do not, however, focus on or distinguish irrigation from dryland farming, or on systems with the complicated patterns of interdependence among farmers that irrigation necessitates. Many of the studies about India and Pakistan include irrigated and unirrigated land in the same samples, frequently without distinguishing them from one another.[3] This distinction is important, because: a) irrigated farms have the potential to be much more productive than dryland farms in most ecological zones; b) most of the green revolution technologies have been developed for irrigated farms; c) irrigated farms are typically smaller than dryland farms.

If Berry's and Cline's results about the relation between farm size and efficiency are correct, there is no tradeoff between efficiency and equity, and therefore the appropriate land policy is to distribute the available land among as many people as wish to have access to it. If their generalization is not true, then there is a tradeoff, and policymakers must balance equity and efficiency goals in designing land policies. In any event, there is a risk of serious error if policy is based on preconceptions rather than facts.

Africa has contributed little to the comparative international experience because little economic research has been done on African irrigation. The principal summary of agricultural research

3. see, e.g. Lawrence J. Lau, and Pan A. Yotopoulos. "A Test for Relative Efficiency and Application to Indian Agriculture." American Economic Review, 61(1), 1971, pp. 94-109, and Mahmood H. Khan and Dennis R. Maki. "Effects of Farm Size on Economic Efficiency: the Case of Pakistan." American Journal of Agricultural Economics. 61(1), February 1979, pp. 64-69. Khan and Maki, incidentally, find that larger farms are more economically efficient than small farms.

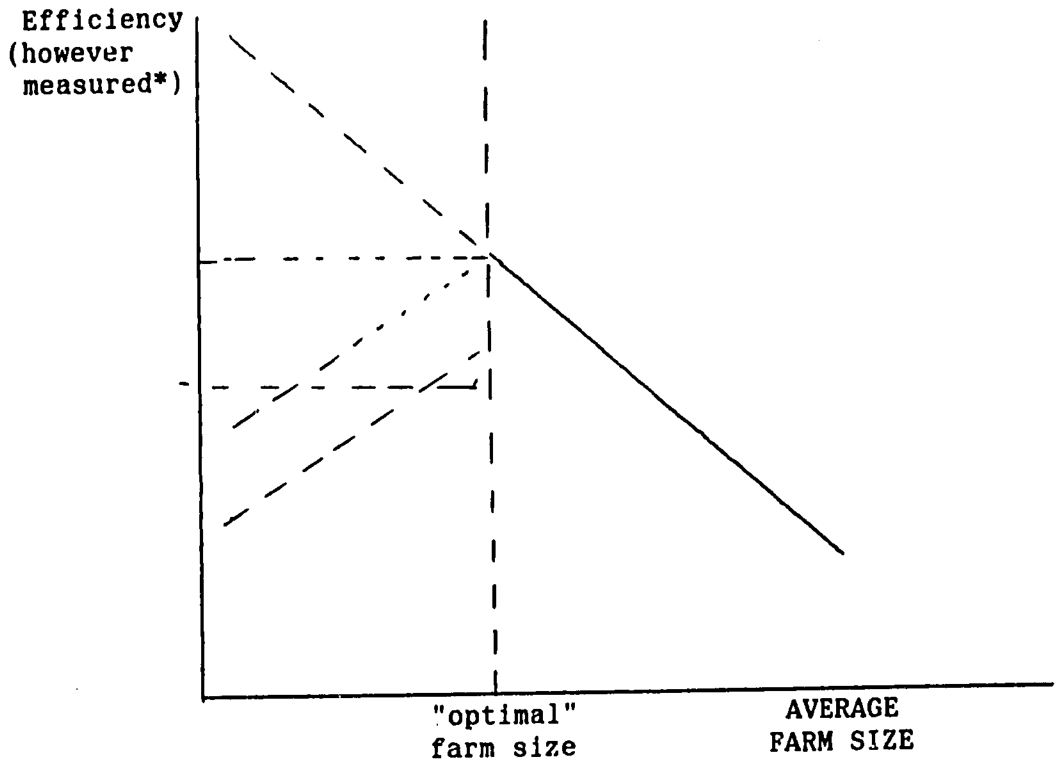
on Africa[4] suggests that small-scale irrigation may be more efficient than large-scale irrigation in Africa. But the size of the scheme is not necessarily correlated with the size of the individual farmer's holdings, and it is the latter which is relevant for equity considerations. It should be noted that on most public irrigation projects in Africa, what is called a large individual holding is quite small; in Zimbabwe's communal area schemes, the largest holdings are under 3 hectares. Thus we do not have the same degree of concern about equity outcomes as if we were studying Brazilian dryland farms, for example.

The basic hypothesis to test is that there is an optimal holdings size, below which plots are too small either to benefit from appropriate intensification technologies or to maintain the active interest of the farmer, and above which the Berry-Cline scale diseconomies set in. The best way to visualize this hypothesis is to consider that the relationship between holdings size and efficiency is not linear, and may be discontinuous. In other words, the Berry-Cline results apply only above a certain minimum plot size (perhaps one-half to one hectare, but we have no firm evidence yet to support this), but below that size efficiency is lower. One possible way to depict this graphically is shown below.

---

4. Carl K. Eicher and Doyle C. Baker. Research on Agricultural Development in Sub-Saharan Africa: A Critical Survey. MSU International Development Paper No. 1. East Lansing: Michigan State University Department of Agricultural Economics. 1982). One of the few interesting pieces of work on this issue is Hasan Tuluy's article on Senegal River irrigated rice farming in Scott R. Pearson; J. Dirck Stryker; Charles P. Humphreys; et al. Rice in West Africa: Policy and Economics. Stanford, CA: Stanford University Press, 1981.

Figure 1
Hypothesized Relationship Between Farm Size and Efficiency



* Most correctly as total economic resource cost per unit of output. It is likely that it will have to be measured imperfectly given our lack of information on shadow prices of inputs and outputs in Zimbabwe, but efficiency should definitely not be measured simply as yield per hectare.

2. The Context -- the Zimbabwe communal area irrigation schemes

Irrigation has been an important part of the commercial agricultural system in Zimbabwe. Most of the irrigated area is in the half of the national area allocated during the colonial period to white commercial farmers, who produce wheat and cotton for domestic sale and export. Beginning as early as the 1920's, however, the government attempted to introduce irrigation into the African half of the country, notably in the Sabi River valley in the southeast. The origin and early development of the Sabi smallholder irrigation schemes are described in Roder (1985). In more recent years, expecially since the 1970's, other regions have experienced some irrigation development, mostly in the form of so-called "comma-hectare" perimeters, where farmers are allocated micro-plots of 0.1 hectare or so to supplement their dryland farming activities. The evolution of land tenure arrangements on both types of schemes is discussed in Stanning and Bloch (in Bloch 1986).

The Sabi schemes range in size from about 50 to 500 hectares, and are located adjacent to villages whose names they share. They draw most of their water from the Sabi River and its tributaries, though tubewells supplement the surface water on one of them. The farming systems on the schemes are quite similar -- essentially the same cropping pattern is followed and plot sizes are uniform. Average holding per farmer varies by a factor of 2, from 0.7 ha (nearly 2 plots per farmer) to 1.4 ha (nearly 4). Holdings sizes on communal-area irrigation schemes in other regions of Zimbabwe are less uniform, but the principal crops are still the same: maize as the food crop, and wheat, cotton and beans as the cash crops. Over the entire set of schemes, holdings sizes range from under 0.1 to over 1 ha.

Farmers had limited freedom of action on these schemes during the period for which the data exist. There were professional scheme managers who tried to enforce cropping patterns and supervised the purchase and distribution of inputs. Marketing was generally done on a scheme-wide basis. The "within" variance in variables affecting farmer performance is thus not very large. Therefore, the use of aggregate data, i.e. using the scheme as the unit of analysis, as we are constrained to do here, probably does not do as much violence to reality as one might initially suspect.

3. <u>The procedure</u>

The holdings size issue can be addressed by standard econometric techniques. One can estimate production functions or yield equations with an added variable representing farm size by ordinary least squares; the farm size parameter will provide a test of relative technical efficiency of larger versus smaller farms. The problem is that while the parameter estimates from such models will be unbiased, they will not be consistent. Also, they only test for technical, and not for economic, efficiency.

One could overcome these difficulties by using models such as those developed by Lau and Yotopoulos (1971). They estimate a profit function to assess the relative economic efficiency of large versus small farms. The estimates of production relationships which come out of this model have the desirable econometric properties of consistency and statistical efficiency, but the data requirements are somewhat severe.

Another technique which could be used is that of the multicrop production function.[5] This very recently developed technique recognizes that in complex farming systems, many of the needed data, especially on the input side, will be available only as aggregates rather than attributed to individual crops. On the plus side, the technique is more forgiving of data gaps. On the minus side, it is more econometrically ambitious, therefore more likely to be perceived by non-econometricians as the use of a sledgehammer to crack a very thin-shelled nut. Both the Lau-Yotopoulos and the Just methods merit consideration in spite of their weaknesses, because they overcome some of the greatest problems of simple OLS estimation of production or yield functions.

However, the benefits of rigor may be outweighed by the costs of over-stringent data requirements. The econometric arguments about using profit functions instead of production functions to make direct estimates of relative efficiency have to do with consistency and (statistical) efficiency. The only evident bias in estimates is that due to grouped data; see Maddala (1977, pp. 69, 273). Given the weakness of the data, the asymptotic properties are a second-order problem. Production and yield function estimates will be sufficient. If so, estimate the Cobb-Douglas form of the production function:*

(2) $Q = A\ K^a\ L^b\ T^c\ M^d$

where Q is the volume of output;
M is the volume of purchased inputs;
K is the stock of capital applied to the productive system;
L is the labor input applied to the productive system; and
T is the land area under cultivation

Any difference in efficiency between two groups of farms should show up as a difference in the constant A, so that a dummy variable would be as satisfactory here as in the profit function. The specification of the function for estimation purposes should be:

---

5. The standard reference is R. E. Just; D. Zilberman; and E. Hochman. "Estimation of Multicrop Production Functions." American Journal of Agricultural Economics. 65 (1983), pp. 770-780. Useful extensions are contained in C. Richard Shumway; Rulon D. Pope; and Elizabeth K. Nash. "Allocatable Fixed Inputs and Jointness in Agricultural Production: Implications for Economic Modeling." American Journal of Agricultural Economics. 66 (1984), pp. 72-78.

(3) $\ln Q = \ln A + a_1 D + b_1 \ln K + b_2 \ln L + b_3 \ln T + b_4 \ln M$,

where D is the dummy designed to capture differences in efficiency and the other variables are as defined above.

4. The data

Various characteristics of the irrigation schemes in communal areas have been reported since the early 1970's by the Department of Rural Development (DERUDE) of the Ministry of Agriculture, in a series of annual reports on Agricultural Production in Communal Land Irrigation Schemes and ARDA Estates. Production, marketing and input data for fifty or more smallholder irrigation schemes are available for a period of up to eight years. The schemes range in size from about 5 to over 200 hectares, with average plot sizes varying from less than 0.05 ha. to over 1.0 ha. per plotholder. Some of the schemes are over 50 years old (the Sabi River schemes -- see Bloch 1986), and some were established in the last few years. A variety of crops is produced; most of the farmers produce several cash crops as well as basic foodgrains. Most of the schemes suffer periodically from water shortages, especially during the drought of the early 1980's, but many are able to double-crop anyway. Most schemes use fertilizer, draft oxen, and contract plowing. The labor information is the least satisfactory, because it refers to "persons on the plot" by age and sex, rather than giving their activities. The "other persons" listed may or may not be hired laborers. But in order to use "persons on the plot" as a proxy for labor input we only have to assume that there is a roughly similar percentage of effort devoted to irrigation by households on the different schemes, or alternatively that labor input per "person on the plot" varies systematically with other information (such as livestock ownership) that we possess.

Data are available for all of the Sabi valley schemes for the years 1974/75 to 1982/83, and on a more scattered basis for the schemes in other parts of the country. The usable data included in the report are defined as follows:

Output: production (in tons) of each crop. Only six(?): wheat, maize, cotton, sugar beans, groundnuts, and tomatoes, were analyzed, since the others represented less than __ per cent of output.

Land: area (in hectares) planted to each crop.

Labor: The number of labor-days is not available, either in the aggregate or by crop. The number that is available is the number of people living on the scheme, broken down by sex and by age (over 16, children 7-16 at home, children 7-16 not in residence, children under 16). One can select any subset of the number of people as the labor input variable. Another indicator of labor is simply the number of plotholders. Using this would imply that each scheme's ratio of family labor per farmer is equal. Either possible definition of this variable is obviously problematic: to the extent that irrigation farming is not a full-time activity, assuming that all workers are employed full-time (or an equal proportion of full time) on the scheme will overestimate labor input to a greater extent in schemes where irrigation is minor than where it is a major use of family labor. After much experimentation it was determined that the results were not sensitive to the choice of labor variable.

Fertilizer: amount of fertilizer (in tons) applied to each crop.

Capital: the only indicator of capital input (other than the irrigation system itself) is the number of draught oxen on the scheme. There is no evidence that these are used only for plowing the scheme land, just as there is no evidence that the labor is applied only to the scheme.

Marketing : sales of crops, in tons except for tomatoes and other vegetables, given in Zimbabwe dollars.

There is a potential of approximately 350 observations: 7 years on each of approximately 50 schemes. The disappearance of some schemes and creation of others reduces this to a total of 273 observations. This is certainly adequate for serious statistical work. It is unclear that more observations would add very much, especially given the problem of measurement associated with the labor variable and the fact that the data are aggregate rather than individual.

5. Preliminary view of the data

Figure 2 is a scatterplot of the value of income per hectare against average holdings size. At first glance, a negative-sloped relationship appears to exist at average holdings sizes between 0.1 and 1 hectare; above 1 ha. or so the slope is close to zero. The negative slope is emphasized by about twenty observations which lie substantially higher than the bulk of others. These outliers have their opposites in the form of a few observations which are below the core clump in the scatterplot.

Figure 3 is a scatterplot of maize yields against average holdings size. Maize is planted on virtually all of the schemes, and is the principal food crop in addition to being marketed. Here the negative relationship between yield and acreage is even more strongly dependent on a small group of observations in the lower size range. The most noteworthy characteristic of this scatterplot is not the slope, but the decreasing variance of yield as area increases. The smallest maize plantings have both the largest and the smallest yields in the sample with a range of 15:1 between them, and there is a wide and fairly even distribution of yields between the extremes. Larger acreages are associated with relatively uniform yields (if a 2:1 range can be called uniform).

The scatterplot in Figure 2 may be interpreted in the same way: a negative relationship between farm size and the variance of yields. There are several competing explanations for this relationship:

1) Institutional factors which determine yield, such as management, condition of the irrigation system, reliability of water delivery, etc., may be more variable on irrigation schemes with small than with large average holdings sizes.

2) Measurement and reporting errors may affect the quality of reported data on schemes with small average holdings size more than that of schemes with large average holdings size.

FIGURE 2: Scatterplot of Average Income per Hectare against Average Holdings Size

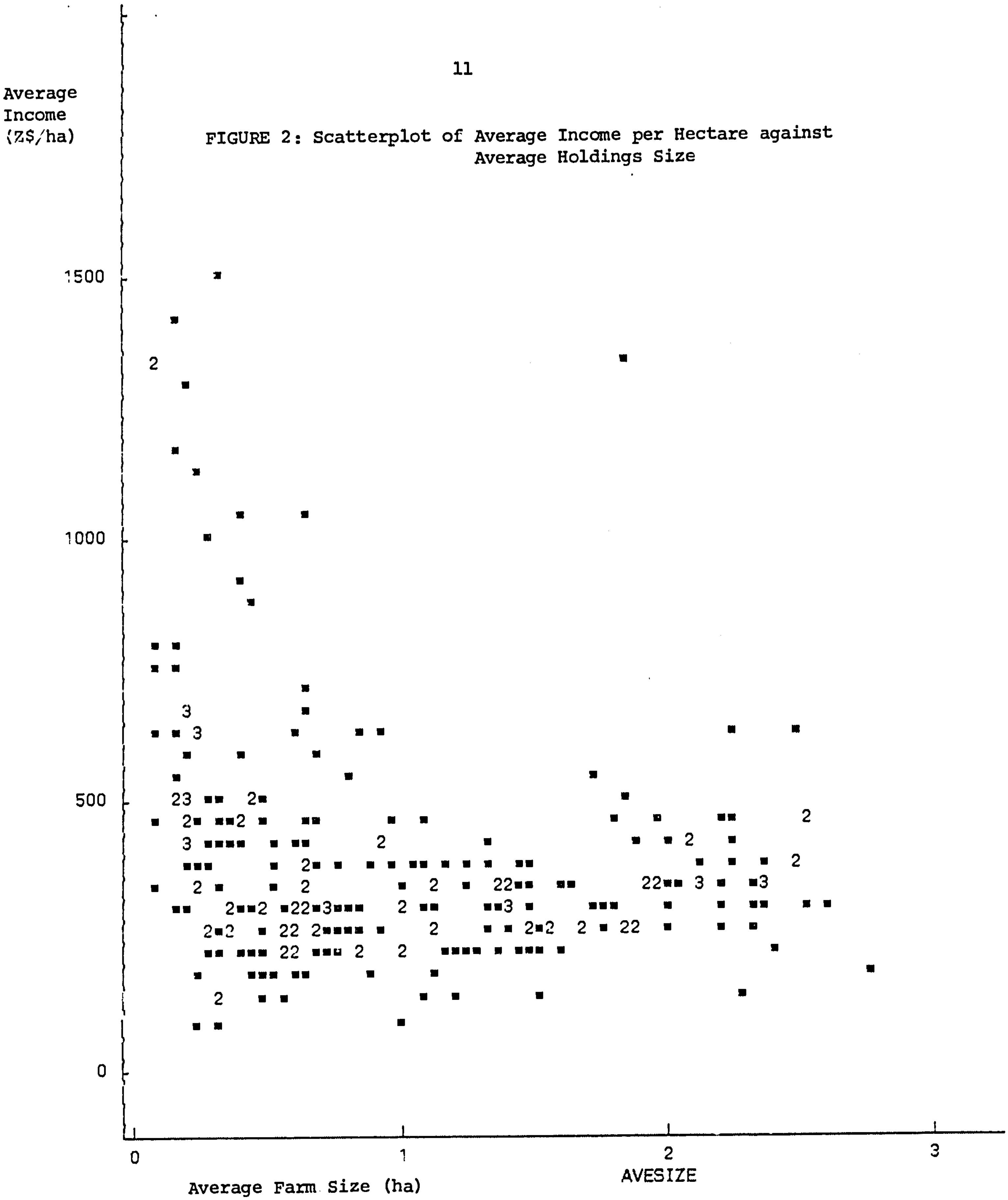
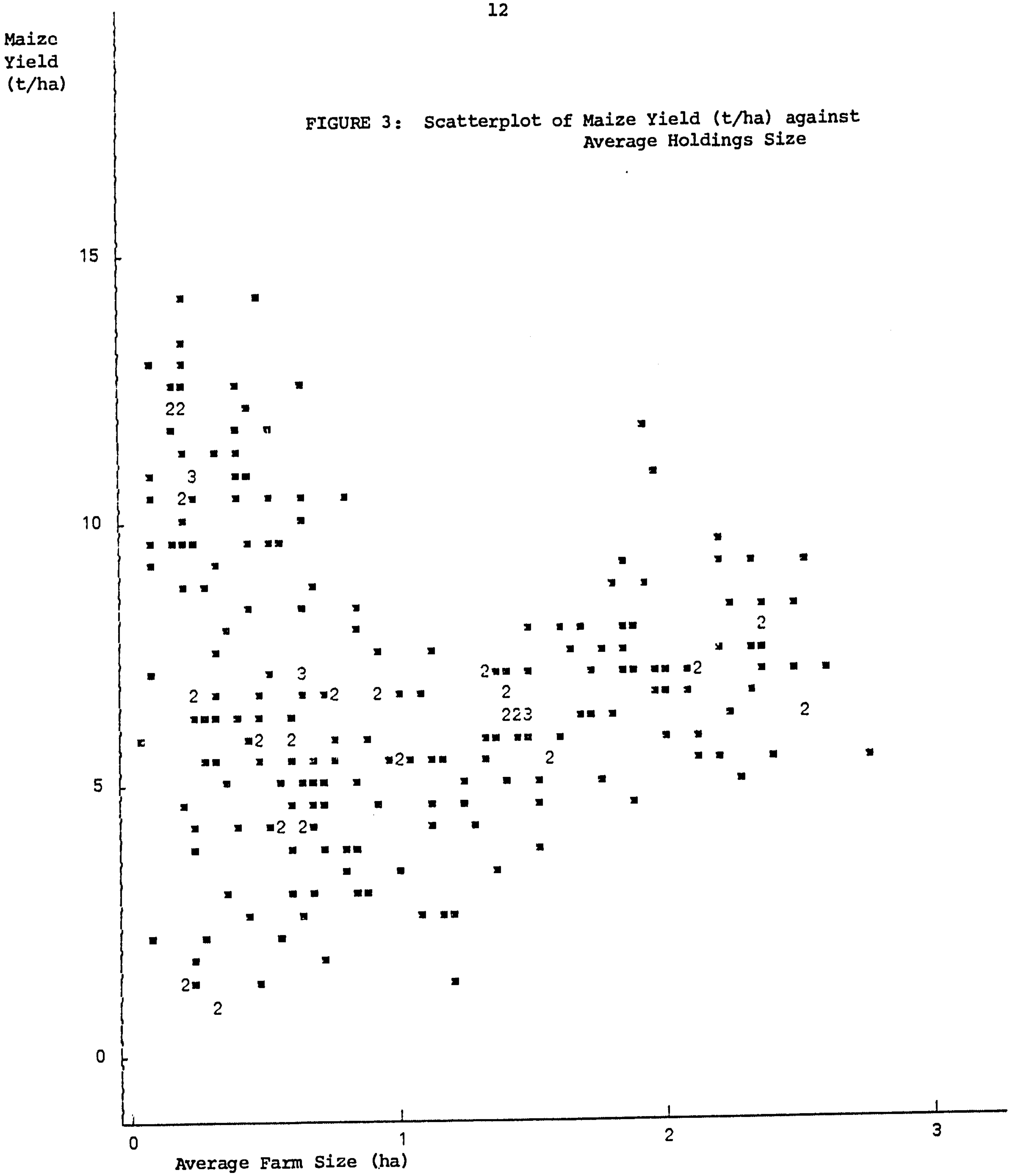


FIGURE 3: Scatterplot of Maize Yield (t/ha) against Average Holdings Size



3) There may be two distinct groups of people with small plots, those with alternatives to irrigation and those without; their strategies for dealing with their irrigated plots are likely to be very different, and so are the outcomes. Holders of large irrigated parcels have less need for alternative sources of income, and their strategy will focus on proper use of their irrigated plot.

We cannot use the present data set to choose from among these explanations in any formal statistical way. We can, however, make some comments about which factors are likely to dominate. First, explanations 1) and 2) are really more relevant to the size of irrigated perimeters rather than to the size of individual plots. The larger Zimbabwe schemes have resident professional managers, whereas the smaller ones must rely on itinerant managers; this may affect both delivery of irrigation services and data quality. On the other hand, it is hard to distinguish between size of scheme and average size of holdings, because there is quite a strong association between the two (Pearson correlation coefficient of 0.59).

Explanation 3) has interesting implications. Farmers with small irrigated holdings cannot rely on them for their entire means of subsistence. This is explicitly recognized in Zimbabwe, where the "comma-hectare" irrigation schemes are intended to supplement rather than to replace dryland farming activities. Plotholders with viable dryland agriculture may choose to put little effort into their irrigated plots, whereas those who are not so lucky may be constrained to apply as much labor as possible to the small plots, thus achieving very high yields per hectare (even though returns to labor are low). Mediocre yields may occur in cases such as one situation documented in Senegal (Bloch and Sella, 1987): if the rainy season begins well, farmers neglect their irrigated holdings; if the rains falter they return to the irrigated plots, but too late to achieve good yields.

6. The "productive but inefficient" hypothesis

Schultz (1964) describes the behavior of subsistence farmers in a way that has been labeled the "efficient but poor" hypothesis. According to Schultz, farmers adapt rationally to the constraints imposed by nature and institutions, and that poverty is not a sign of inefficient use of available resources, but rather a reflection of the inadequacy of resources. On the Zimbabwe communal-area irrigation schemes, we may have a converse situation: the observed high yields are not a sign of efficient use of available resources, but rather of excessive input use. Farmers are constrained by their poverty and their lack of alternative sources of income to allocate too much labor to small irrigated plots. This situation would be reminiscent of the one depicted by Geertz (1963) on Javanese rice farms, where increasing population pressure on a limited land area led to ever-increasing labor intensity and higher yields per hectare: "agricultural involution" with constant returns to labor at best.

A formal statement of the "productive but inefficient" hypothesis is:

> **High yields on small farms are the result of higher than optimal quantities of inputs being applied to plots of land of suboptimal size by farmers with limited alternatives.**

The data upon which this paper focuses permit the testing of this hypothesis only to a limited extent. The principal problems have to do with the measurement of inputs, notably labor. In the published reports from Zimbabwe the information on labor is limited to the number of persons living on the schemes (broken down by gender and age) rather than the amount of labor actually applied to the irrigated plot. As we have seen, this fact leads to low estimates of the contribution of labor to output, because variability of labor input is caused in part by existence of unmeasured alternative uses of labor. The following section presents some preliminary evidence about the hypothesis, in full recognition of the inadequacy of the data.

6. Data Analysis

Variables used are defined in Table 1, and estimates of the production function are shown in Table 2. Consider first the "classic" Cobb-Douglas function with four inputs, shown in the center of the table. Land, labor and fertilizer all enter as significant determinants of output, and the sum of the input elasticity coefficients is extremely close to 1.0. The impact of the oxen variable is negligible. Nearly 90 per cent of the variation in output is explained by variation in inputs, a very tight fit. The impact of the size dummy alone is shown at the left side of the table. Total output is a positive function of average holdings size, but the coefficient is less than 1.0.

The combination of the production function and the size dummy is shown at the right side of the table. The input elasticities remain close to the same, with land's value falling and labor's rising but the sum remaining very close to 1.0. The size dummy is positive and significant. This suggests that schemes with larger holdings sizes are more productive than those with smaller holdings sizes once inputs are taken into account.

In Table 3, income per hectare (call it aggregate yield) is the dependent variable, and inputs per hectare the explanatory variables. This equation is related to the production function, where output and inputs are divided by area. The simple relationship between aggregate yield and the size dummy is negative: aggregate yields are higher on schemes with small than with large holdings sizes. This is the story shown in Figure 2. When inputs are taken into account, however, as in the right-hand equation, the relationship becomes positive. In other words, higher input use per hectare over-explains the higher yields on smaller holdings. In still other words, small farms are productive in that their aggregate yields are high, but they are inefficient in that they use inputs to excess.

TABLE 1: Variable Definitions

Labor: number of adults (over 16) living on the scheme in the families of plotholders

Fertilizer: tons of fertilizer applied to all crops

Oxen: number of draught oxen owned by plotholders

Area: number of hectares planted per year in both cropping seasons (thus a double-cropped, fully-planted scheme would have Area=twice the scheme's area)

Total Income: the value of all crops produced over the year, at official or market prices deflated by a price index with 19xx=100 (see appendix X)

Size: a dummy variable	equalling 1	if average holdings size is greater than 1.0 ha., and
	equalling 0	if average holdings size is less than or equal to 1.0 ha.

Size cat. 1 - 4: dummy variables differentiating holdings size categories:
1: value of 1 if size is between 0.51 and 0.75 ha., and 0 otherwise
2: value of 1 if size is between 0.76 and 1.00 ha., and 0 otherwise
3: value of 1 if size is between 1.01 and 1.50 ha., and 0 otherwise
4: value of 1 if size is greater than 1.50 ha.
(omitted category -- all dummy variables equal zero -- is size 0 - 0.50 ha.)

Region: dummy variables differentiating provinces:
1: Manicaland
2: Mashonaland
3: Matabeleland North
4: Matabeleland South
5: Midlands
(omitted province -- all dummy variables equal zero -- is Masvingo)

Year: dummy variables differentiating years
(omitted year -- 1975)

TABLE 2: Production Function Estimates

Dependent Variable: Log(total income) size dummy =1 for 1.0 ha.+

VARIABLE	Coeff.	t-ratio	Coeff.	t-ratio	Coeff.	t-ratio
Constant	8.761	87.93	5.756	37.55	5.697	37.06
Size dummy	1.732	11.06			0.232	2.46
Log(labor)			0.137	3.70	0.181	4.43
Log(fert.)			0.351	9.15	0.361	9.44
Log(oxen)			0.013	1.03	0.009	0.68
Log(area)			0.517	13.45	0.439	8.87
Adjusted R^2	.312		.887		.889	

Note: Coefficients with t-ratios greater than 1.97 are statistically significant at the 5% level on a two-tailed test.

TABLE 3: Aggregate Yield Function Estimates

Dependent variable: Log(total income per hectare); size dummy =1 for 1.0 ha.+

VARIABLE	Coeff.	t-ratio	Coeff.	t-ratio	Coeff.	t-ratio
Constant	5.926	118.30	6.085	47.27	5.673	35.04
Size dummy	-.274	-3.48			0.374	4.00
Log(labor/ha.)			0.020	0.33	0.156	2.31
Log(fert./ha.)			0.479	7.93	0.487	8.31
Log(oxen/ha.)			0.070	1.87	0.069	1.91
Adjusted R^2	.040		.394		.430	

Unfortunately there are adjustments that must be made to this picture: omitted variables may play an important part in these results. There is some reason to believe that regional differences are important, and correlated with such included variables as average holdings size. It is also likely, given Zimbabwe's turbulent history in the late 'seventies, that annual variation may greatly influence the data. We therefore should correct for these influences by including them in the model. There are solid econometric reasons for doing this, too: the fact that the sample is a pooled cross-section and time-series means that the individual observations are not independent, as the equations estimated above assume (footnote on seemingly unrelated regressions).

In addition, the holdings size dummy variable is quite restrictive, because it only separates the sample into two categories of schemes -- large and small -- and does so quite arbitrarily (see Appendix I for results of redefining the size dummy to differentiate average holdings sizes above and below 0.5 ha.). A better procedure is to determine a vector of dummy variables which differentiate several size categories. There is no magic number of such categories; in the work reported upon here we use four dummy variables to differentiate among five categories.

Results of the estimates of the production function are shown in Table 4, and of the aggregate yield function in Table 5. The left-hand columns repeat the production function equations shown in Tables 2 and 3. The center column adds the four size dummy variables, the coefficients on which represent deviations from the omitted size category: holdings size below 0.5 ha. The right-hand columns add region and year dummy variables. In Table 4's center column, the largest size dummy has a significant and positive coefficient, suggesting that farms in this category are more productive than the smallest farms once one corrects for

TABLE 4: Production Function Estimates

Dependent Variable: Log(total income)

VARIABLE	Coeff.	t-ratio	Coeff.	t-ratio	Coeff.	t-ratio
Constant	5.756	37.55	5.815	35.40	5.910	33.47
Log(labor)	0.137	3.70	0.170	4.01	0.158	3.96
Log(fertilizer)	0.351	9.15	0.345	9.06	0.366	9.20
Log(oxen)	0.013	1.03	0.005	0.36	-.004	-.294
Log(area)	0.517	13.45	0.468	8.77	0.384	6.66
Size cat. 1			-.233	-2.42	-.099	-1.01
Size cat. 2			0.160	1.17	0.170	1.32
Size cat. 3			-.015	-.11	-.126	- .95
Size cat. 4			0.310	3.16	0.079	0.77
Region 1					0.605	3.45
Region 2					0.648	3.15
Region 3					-.027	-.22
Region 4					0.033	0.27
Region 5					0.309	2.09
D1976					0.178	1.88
D1977					-.058	-.59
D1978					0.001	0.01
D1979					0.310	2.67
D1981					0.040	0.37
D1982					0.051	0.49
Adjusted R^2	0.887		0.894		0.909	

TABLE 5: Aggregate Yield Function Estimates

Dependent Variable: Log(total income per hectare)

VARIABLE	Coeff.	t-ratio	Coeff.	t-ratio	Coeff.	t-ratio
Constant	6.085	47.27	5.650	29.68	5.474	24.96
Log(labor/ha.)	0.020	0.33	0.166	2.35	0.193	2.83
Log(fert./ha.)	0.479	7.93	0.495	8.28	0.557	8.96
Log(oxen/ha.)	0.070	1.87	0.073	2.03	0.076	2.14
Size cat. 1			-.080	-.74	0.065	0.61
Size cat. 2			0.250	1.85	0.267	2.14
Size cat. 3			0.019	.15	-.126	-.99
Size cat. 4			0.314	3.30	0.090	0.92
Region1					0.489	3.61
Region2					0.557	3.28
Region3					0.009	0.08
Region4					-.064	-.57
Region5					0.267	1.98
D1976					0.138	1.48
D1977					-.119	-1.22
D1978					0.011	0.11
D1979					0.225	1.95
D1981					0.020	0.20
D1982					-.024	-.22
Adjusted R^2	0.394		0.445		0.538	

NOTE: Coefficients with t-ratios greater than 1.97 are significant at the 5% level on a two-tailed test.

input use. Conversely, the second-smallest size (Size cat. 1) has a significantly negative coefficient. When one accounts for effects of regional and annual differences, however, both of these effects disappear, although the signs stay the same.

Aggregate yield equations are shown in Table 5. Here, as in the production functions, the largest size dummy is significantly positive, but the smallest is negative but not statistically significant. In the full model including region and year, the largest size dummy becomes nonsignificant, and Size cat. 2 (farms between 0.75 and 1.0 ha.) becomes significantly positive.

7. Extensions

In experiments not reported here, econometrically more rigorous estimates were made, using both weighted least squares estimates of the Cobb-Douglas production and aggregate yield functions and also estimates of translog production functions. There were no important additional insights gained from such endeavors.

Similar estimates can be conducted for individual crops, especially those which are grown on most schemes, such as maize, wheat and cotton. The results are even less conclusive there: labor input has no significant contribution to output in even the simple production function and the size dummy variables follow no predictable pattern. Further investigation of the farm size/productivity relationship will require better data sets.

In summary, the preliminary evidence supporting the "productive but inefficient" hypothesis disappears when more complete versions of the production and aggregate yield functions are estimated. This may be due to data weaknesses or to the incorrectness of the hypothesis.

8. Conclusion

This paper set out to explore the relationship between farm size and efficiency of production. The goal was to attempt to identify something like an "optimal farm size" that was not the smallest conceivable farm as the conventional generalization of Berry-Cline results appears to suggest. This goal was not achieved, but in the process of exploring the data an interesting finding emerged: the large variance in results achieved by small farms. Some farmers produce enormous amounts on small plots, and others produce very little. It was then hypothesized that the productive farmers are using excessive amounts of inputs, especially labor.

To test the "productive but inefficient" hypothesis satisfactorily, the following information should suffice:

a. family labor allocations (person-days or half-days will suffice) among dryland and irrigated farming and nonfarm activities.

b. purchased inputs, including wage labor.

c. output by crop on dryland and irrigated plots.

d. nonfarm income, both earned and unearned (rents, remittances).

e. institutional information about management, irrigation system maintenance, water availability, etc.

We expect to obtain these data in our Senegal research program, from irrigated perimeters with a wide range of plot sizes.

Appendix I: Production and Aggregate Yield Functions with Different Size Dummy

TABLE A: Production Function Estimates

Dependent Variable: Log(total income) size dummy = 1 for 0.5 ha.+

VARIABLE	Coeff.	t-ratio	Coeff.	t-ratio	Coeff.	t-ratio
Constant	8.702	82.86	5.756	37.55	5.656	36.35
Size dummy	0.650	10.74			0.104	2.78
Log(labor)			0.137	3.70	0.194	4.62
Log(fert.)			0.351	9.15	0.365	9.56
Log(oxen)			0.013	1.03	0.008	0.62
Log(area)			0.517	13.45	0.421	8.17
Adjusted R^2	.299		.887		.890	

TABLE B: Aggregate Yield Function Estimates

Dependent variable: Log(total income per hectare); size dummy = 1 for 0.5 ha.+

VARIABLE	Coeff.	t-ratio	Coeff.	t-ratio	Coeff.	t-ratio
Constant	5.948	114.11	6.085	47.27	5.569	33.81
Size dummy	-.114	-3.78			0.181	4.72
Log(labor/ha.)			0.020	0.33	0.183	2.72
Log(fert./ha.)			0.479	7.93	0.507	8.72
Log(oxen/ha.)			0.070	1.87	0.077	2.16
Adjusted R^2	.047		.394		.444	