Survey costs and rural economics research

John McIntire, *International Crops Research Institute for the Semi-Arid Tropics, Sahel Center, Niamey, Niger*

The recent vogue of farming-systems research in Africa among scientists, donors, and bureaucrats has arisen largely from their frustration at the slow progress of African agriculture. The fashion is sustained by the convictions that profitable technical packages exist, that scientists fail to exploit farmers' knowledge in research, and that existing methods increase research costs by unnecessarily extending the payoff period.

The conviction that profitable packages exist encourages governments and development agencies to search for effective ways to supply the techniques to farmers. The argument is that inputs to farmers, especially information as it is supplied by extension, are a sufficient and necessary condition for adoption of new techniques. The belief that scientists use farmers' knowledge inefficiently, notably by failing to understand farmers' objectives, explains much of the emphasis on village work and especially on doing more than demonstrations in farming-systems research. The argument about profitable packages partly explains the insistence on quick results because it assumes that many of the fundamental (i.e., long-term) problems have been solved. These influences give farming-systems research its principal characteristics: close link to extension; involvement of many disciplines (including social scientists); bias toward quick results; and prejudice against fundamental research.

Wealthy lobbies support farming-systems research strongly and, by implication, the assumptions upon which it is based. These assumptions determine how the lobbies spend their money and how this spending affects farmers. It is important, therefore, to understand the economics of farming-systems research, to relate its economics to its objectives, and to define efficient methods given costs and objectives.

This paper analyzes the costs of the two principal types of methods — intensive (emphasizing quantitative data collection and analysis) and extensive (searching for a qualitative understanding of the farmers' environment and their responses to it). I believe that the differences between the methods are smaller than the similarities and that there is some scope for combining them to exploit the virtues of both.

Intensive surveys

ICRISAT's economics program has used intensive surveys in India (since 1975), Upper Volta (since 1980), and Niger (since 1982). Small

numbers of villages — six originally in India, six in Upper Volta, four in Niger — are studied after a literature review and preliminary visits to identify suitable sites in different agricultural zones (Jodha et al. 1977; McIntire and Matlon 1981).

Field enumerators reside in the villages and visit samples of 25–40 households every 1–3 weeks. After censuses of people, fields, animals, and machinery, the regular interviews (sometimes known as cost-route interviews) are conducted on crop production, crop and livestock transactions, and transactions in inputs and in land, labour, and capital. In crop production, enumerators follow all inputs and outputs by plot. These data are complemented by special studies on, for example, soil fertility, millet marketing, crop by-products, and cowpea storage.

The short-term aim in these studies is to identify and to quantify variables limiting crop production. From village data, for example, we construct input–output tables of crop production. On the input side are flows of materials and primary factors; on the output side are flows of crops and by-products. Using the tables, we estimate productivity to guide technical research. Because the villages represent agroclimatic zones, the results can be extrapolated (whether immediately or by verification surveys) to other areas.

The long-term aim is to ask fundamental questions about the economies of the semi-arid tropics, answers to which can guide research allocation and policy. For the semi-arid tropics, such questions include: What is the magnitude of farmers' aversion to risk? What are the main determinants of mechanization? What role do markets play? What are the common nutritional deficiencies? How is income distributed? How do farmers respond to changes in supply and demand? How economically efficient are various activities?

Extensive surveys

Extensive surveys begin, as do intensive ones, by defining research areas by the principal exogenous variables in the farming system: rain, soil, altitude, and population density. Zones are then evaluated with rapid surveys of local conditions, such as cropping patterns, mechanization, and chemical inputs. More detailed, exploratory surveys are done (ideally in the cropping season) to verify the findings and to determine what the farmers consider to be the constraints within the zones. The results provide the basis for a set of recommendation domains on farming systems.

The approach is to describe, rapidly and qualitatively, the resources in a farming system, their allocation, and the constraints to their fuller use or to an increase in their productivity. The description is "qualitative" only in that the researchers do not attempt to measure precisely the endogenous variables in the system or to quantify the constraints. Rather, the approach provides educated estimates, from careful interchanges with farmers, of the boundaries for the treatments in technical experiments — for example, cycle length in varietal tests and fertilizer rates in agronomy trials. The boundaries for the variables define the domains for the tests.

Extensive methods have no long-term aim and are not geared to answering fundamental questions. Their proponents assert that the intensive

approach makes inefficient use of scientists' time, that farmers' needs are pressing, and that extensive methods sacrifice little important precision — "important" in respect to bias in trials designed from the results of extensive surveys.

Similarities

The methods agree about much. In fact, extensive methods are perfectly consistent with intensive ones, and, at ICRISAT, we have used them to identify research topics and sites. They agree on zoning to determine what constitutes a representative sample and to guide research allocation. The methods agree on the importance of farmers' knowledge, considered as a rational appreciation of the system and of changes in it. The methods agree on the necessity of a multidisciplinary approach. They share a systems approach; they view endogenous variables such as fertilizer use and mechanization as determined by exogenous variables.

The methods' agreement on the importance of farmers' knowledge implies, first, that the researcher will have to find out some of what the farmer knows, i.e., be directly involved (intensive methods have been accused of precluding this or minimizing it at any rate). Second, it implies that neither method can be described as upstream or downstream because both view farming-systems research as a circle, not a line, as is necessarily implied by notions like upstream and downstream. Whether one begins at the point on the circle where farmers define the problems or where researchers do depends upon the information available at the beginning of a research program.

The intensive and extensive methods differ mainly in how precisely they estimate endogenous variables and in how much importance they give to long-term aims. Advocates of extensive approaches do not deny that precision and long-term perspectives are important; they assert that the costs of greater precision and of more time spent on a single sample exceed the benefits and, therefore, that extensive methods are more efficient than are intensive.

Casting the debate between the two methods as one of the costs of precision in cross-section data and of quantity in time-series enables one to examine their relative costs.

Survey costs

I tabulated ICRISAT's long term (5 years) survey costs from actual intensive surveys in Mali and Niger for 1982 and from budget requests for 1983 (Table 1). Similar budgets (Table 1) were produced for extensive surveys, although the figures were artificial in that the technical coefficients (e.g., professional staff years/sample unit) were estimated from published accounts. Costs from Niger and Upper Volta were applied to the technical coefficients.

From published accounts of extensive surveys (CIMMYT 1978, 1980), I calculated the numbers of staff years in all categories necessary to survey a given number of households. Each number was multiplied by the number of scientists and then multiplied by its annual cost. The costs of local personnel

Table 1. Intensive and extensive survey costs (US$).[a]

	Intensive					Extensive	
	Upper Volta 1983	Niger 1983	Niger 1982	Mali 1982–83	Mean	CIMMYT 1980	Zambia 1978
Capital	7134	17368	13153	4972	10657	1707	1363
Variable	187333	122363	124022	28314	115508	93961	69003
Total	194467	139731	137175	33286	126165	95668	70366
Households	149	107	100	80		80	60
Area (ha)	866	1328	1328	800		NA[b]	191
Population	1604	1132	1132	800		NA	300
Cost/household	1305	1306	1372	416	1157	1196	1173
Variable cost/household	1257	1144	1240	354	1060	1175	1150
Cost (excluding international professionals)/household	676	430	434	318	494	258	348
Cost/ha	225	105	103	42	117	NA	369
Cost/person	121	123	121	42	108	NA	235
Capital (%)	3.67	12.43	9.59	14.94	8.45	1.78	1.94

[a] The table is printed in integer format and may have rounding errors; francs CFA 350 = US$ 1. The discount rate to amortize capital items was 12%/year. Four-wheel-drive vehicles were amortized over 4 years; motorcycles and bicycles were amortized over 2 years; and all vehicles were given a 20% salvage value at the end of amortization. Houses and furniture for field staff were amortized over 5 years, and field equipment (e.g., scales) over 2 years. Office equipment and microcomputers were amortized over 3 years. Some capital costs were tax free (vehicles especially); others, such as construction materials, included duties. Of the variable costs, the most costly item in the budgets was internationally recruited professional staff — for each one, I assumed $75 000/year. Other variable costs were local professional and support staff salaries, office and field supplies with a service life of at most 1 year, communications, vehicle maintenance, temporary labour, and international travel. All these costs included taxes, except those for gasoline in Upper Volta and Niger.

[b] NA = not available.

were assumed to be roughly equivalent to those in Upper Volta and Niger. That assumption could be changed, but it is reasonable if one wishes to compare two methods in the same country.

Capital costs for the extensive surveys were the field vehicle, scientific equipment, and the microcomputer. (Reports of extensive surveys make no mention of the last item, but it is fair to include one given the current cost advantage of micros in Africa.) The costs for these three items were assumed to be the same as they were for intensive surveys (the mean of four surveys). The unit capital costs were multiplied by rates of use — for example, the four-wheel-drive vehicle was assumed to be used for 2 months, a use rate of 0.167.

All operational costs except vehicle maintenance were assumed to be equal to the mean of the intensive surveys. Vehicle maintenance was held at 60% of the intensive mean because enumerators' motorcycles were left out of the extensive surveys. I assumed that office supplies, communications, international travel, and gasoline for the vehicle would not differ between the surveys. In the extensive surveys, I assumed two internationally recruited scientists because farming-systems research teams described in the CIMMYT documents included at least that number.

In terms of costs, the questions are:

- What is the annual cost of each method?
- What is the total cost of each method over the research period?
- Is one cost structure less flexible than the other so that it would lose more if the original research direction were wrong?

- Does one method produce results faster?
- Are there common costs so that advantages of both methods can be exploited?

The mean cost of intensive methods is roughly $1157/household. The range is from $1372 (Niger in 1982) to $416 (Mali). The mean of intensive surveys without the costs of international scientists was $494, ranging from $318 to $676. Expressed in $/member of the survey population, the mean is $117 and the range from $42 to $123.

The estimates for extensive surveys were $1194/household, as estimated from a methodological paper (CIMMYT 1980), and $1169, as estimated from a demonstration of the method in Zambia (CIMMYT 1978). These estimates do not differ significantly from those for intensive surveys. The estimates per hectare and per person in Zambia are much greater than any of the individual estimates for intensive surveys; although this result is clearly a reflection of small family and farm sizes in Zambia, it shows that one cannot always assume extensive surveys are cheaper. Excluding international staff from extensive surveys reduces their costs greatly and makes them less expensive than intensive. The costliest intensive survey was $676, whereas the cheapest extensive was $258. The average intensive ($494) was about 66% more expensive than the average extensive ($297).

For calculations of research expenditures over 5 years at a discount rate of 12%, I took the Niamey 1983 data as typical of an intensive study and those for Zambia to be typical of an extensive one (Table 2). The cost/household is about 24% greater in intensive surveys, although the costs per person and per hectare are greater in the extensive survey done in Zambia. At a 24% discount rate, the relative comparisons do not change, but intensive surveys have a higher cost/household partly because of the capital costs incurred early in the research. Even when only variable costs are considered, intensive surveys are about 15% more expensive than extensive surveys. Flexibility in costs depends on the share of fixed capital and on the care with which research problems are first defined. Extensive methods are

Table 2. Present values of survey costs (US$) at 12% and 24% discounts.[a]

	12%		24%	
	Niger Intensive	Zambia Extensive	Niger Intensive	Zambia Extensive
Capital	53695	30123	44000	24475
Variable	447071	697704	340487	531369
Total	500766	727827	384487	555844
Households	500	900	500	900
Area (ha)	6638	2862	6638	2862
Population	5660	4500	5660	4500
Cost/household	1002	809	769	618
Variable cost/household	894	775	681	590
Cost — international professionals/household	326	208	254	160
Cost/ha	75	254	58	194
Cost/person	88	162	68	124
Capital (%)	10.72	4.14	11.44	4.10

[a] Table may have rounding errors, francs CFA 350 = US$ 1.

more flexible than intensive ones because they have lower relative capital costs — but the average share in the intensive surveys is only 8.5% anyway, most of which is spent on enumerators' houses. Other capital — vehicles, computers, furniture — is movable at low cost and is flexible with both methods.

The costs wasted because of poorly designed research, necessitating abandoning a site or a topic, are equal to the annual survey cost multiplied by the time lost. Because annual costs are similar in the two surveys, neither type has a higher expected cost unless one assumes that one type is more likely than the other to begin wrong.

Advocates of the extensive approach argue that their method works faster and with a bigger sample. Collinson (CIMMYT 1980:11) asserts, for example, ". . . the benefits from wide coverage of small farmer populations dramatically outweigh those from a more intensive, numerate approach among fewer populations."

According to my calculations, extensive methods could cover 180 240 households/year. The population covered depends on household size, and the area depends on household size and on farming techniques. In ICRISAT's surveys, intensive methods cover 80 150 households/year. Extensive methods, therefore, work about twice as fast as intensive ones. If each extensive sample is drawn from a different population, then extensive methods permit inferences about larger populations than do intensive surveys.

The speed of extensive methods is an advantage only if three surveys are conducted annually. This is possible but requires quick work and means increased costs if new field assistants have to be recruited at each survey site. It would be particularly difficult in areas of language fragmentation.

The major common costs — international staff, four-wheel-drive vehicles, field staff needed in a more or less fixed proportion to international staff, data processing, and office supplies — and the low share of fixed capital in both methods imply that farming-systems research teams can easily exploit both methods, in particular by joining the immediacy of the extensive method to the analytic power of the intensive one.

Simulating benefits

The relevance of any method is its effect on output. Because no one can accurately quantify how research has affected food production in Africa, it is impossible to put a value on the effects. Still, simulations of how intensive and extensive methods benefit farming-systems research are possible. The simulations sketch answers to questions important for research design: Should research be concentrated in areas with high or low potential? Does the urgency of results affect research methods? Do lags in adoption affect the choice of methods? What sizes of target populations are necessary to repay various research investments?

I have constructed a model that simulates research costs and benefits. It assumes that there is a 5-year project, in which the donor can choose the intensive or the extensive method. Either method increases agricultural growth within 10 years, and the changes in the per-person income that

Table 3. 10-year present values (US$'000s) of benefits (1% increment) from intensive surveys, with and without 6-year benefit lag, at an original income of $150/person and an original growth rate of 1%/year.

	Discount (%)	
	12	24
Income at 1% growth	888	574
With 1% benefits, no lag	930	597
With 1% benefits, 6-year lag	894	577
Cost	501	384
Breakeven target population ('000s)		
No lag	11.72	16.62
6-year lag	82.94	156.66

[a] The values for income are the 10-year present values of per-person income under the conditions assumed for original income level and growth

exceed the expected annual growth are attributable to farming-systems research. The new level of income is the basis for calculations of the growth the next year. I assumed initial income level to be $150/year, corresponding to rural income in many African countries. The annual expected rate of growth (trend rate) is 1.0%/year. The first increase in growth brought about by farming-systems research is 1.0% — that is, the trend rate is doubled, so that the new rate is 2.0%/year.

I ran the model to see what sizes of target populations were necessary to repay research costs. The sizes of the target populations were tested for sensitivity to the rate of discount; the lag in technology adoption; the original income level; and the trend rate of growth.

Assuming no lag in the entire population's adoption of beneficial techniques, I found that a target population of almost 12 000 is necessary to repay intensive survey costs at a 12% discount rate (Table 3). Another way of looking at the result is that an intensive project providing immediate benefits to 12 000 people has an internal rate of return of 12%. At a 24% rate of discount, a target population of 17 000 is necessary. Extensive research needs a target population of 17 000 at 12% discount and 24 000 at 24% (Table 4).

The benefit cost calculations for both methods are sensitive to the rate of discount: varying the rate by 100% (from 12% to 24%) causes about a 41% increase in the necessary target population, implying an elasticity of 0.41. The extensive method is no more or less sensitive than is the intensive one. In other words, the urgency of results, used to justify the use of rapid, extensive methods, does not affect the choice of methods.

Advocates of extensive methods argue that their methods produce benefits quicker. If this were so, then such methods would have smaller target populations to repay research costs. I have evaluated this argument by assuming that intensive methods have a lag of 6 years before they produce benefits but that extensive methods have only a 4-year lag.

A 6-year lag in benefits from intensive research at a discount rate of 12% increases sevenfold the target population necessary to repay the costs (Table 3). At a 24% discount, a 6-year lag increases the target population from 16 620 to 156 660. An 8-year lag increases the target population to 612 000. Similarly, in extensive research, time lags increase the target populations

Table 4. 10-year present values (US$'000s) for benefits (1% increment) from extensive surveys, with and without 4-year benefit lag, at an original income of $150/person and an original growth rate of 1%/year.

	Discount (%)	
	12	24
Income at 1% growth[a]	888	574
With 1% benefits, no lag	930	597
With 1% benefits, 4-year lag	901	580
Cost	728	556
Breakeven target population ('000s)		
No lag	17.03	24.02
4-year lag	52.90	91.03

[a] The values for income are the 10-year present values of per-person income under the conditions assumed for original income level and growth.

necessary to repay the costs: at 12% and 24% discounts, a 4-year lag more than triples the target populations (Table 4). If extensive methods actually do produce benefits more quickly than do intensive methods, they have a considerable advantage. For example, with a 4-year lag, extensive methods would require target populations only 58–64% of those for intensive methods with a 6-year lag.

Another question I was able to address using the model was whether farming-systems research should concentrate on areas with low or high potential — a question that is widely debated. One school argues for focusing on areas where the potential return is highest — usually in high-rainfall areas. Another school argues for concentrating on areas where help is needed most — among the poorest farmers in the driest areas. If the location does not affect the productivity of research, then one can concentrate on the areas where the help is needed most. To evaluate these arguments, I varied the original level of per-person income and growth rate to model "favourable" (e.g., high-income, high-growth) areas and "unfavourable" (e.g., low income, low growth) areas. If the size of the target populations did not vary when the trend rate of growth or the original income level was changed, then the productivity would not be affected by location.

Table 5. 10-year present values (US$'000s) for benefits (1% increment) from intensive surveys, with and without 6-year benefit lag, at an original income of $300/person and an original growth rate of 3%/year.

	Discount (%)	
	12	24
Income at 3% growth[a]	1952	1244
With 1% increment, no lag	2049	1295
With 1% increment, 6-year lag	1966	1249
Cost	501	384
Breakeven target population ('000s)		
No lag	5.15	7.41
6-year lag	34.70	65.69

[a] The values for income are the 10-year present values of per-person income under the conditions assumed for original income level and growth.

I assumed an original income level of $300/year and a trend growth rate of 3.0%/year (Tables 5 and 6) and compared the results with those for the lower income ($150/year) and lower trend growth (1%). I found that increases in income enabled more "profitable" intensive and extensive research because the target populations to repay costs were much smaller. The results also showed that the effects of the lags were much reduced by the higher-income assumptions.

The implication is that research should be concentrated in high-rainfall areas. This conclusion is strengthened if one includes the probability of achieving a given level of growth in the calculations. Because the probability of a 1% increase in the growth rate increases with rainfall, expected benefits (a specified increase multiplied by its probability) are greater in high-rainfall areas. If, as is likely, adoption lags are shorter in high-rainfall areas, including a probabilistic lag also favours placing research in high-rainfall areas.

Survey costs and farmers' participation

Farmers' participation has distinct effects on the costs and benefits at each stage in village-based research: design, execution, and analysis. At the design stage, the farmer provides information about constraints and about investments to relieve them. This role differs little between intensive and extensive methods. Errors occur because farmers, with whom the researchers are not well-acquainted, can make systematically misleading statements. Farmers make errors of magnitude — for example, in exaggerating the prevalence of a disease by reporting only extreme cases. These errors arise from confusion, a desire to please, to hide facts, or to mislead in the hope of receiving aid. They can be reduced by checks and by discussion with informed observers, but there are many examples of unexpected discoveries after long periods in what the researchers thought were well-known areas.

The costs of such errors are increases in the time it takes research to pay off. If one can reduce such errors to roughly zero in 1 year, then at most they would add a year to the payoff period. Because the response of survey benefits to lags is nonlinear — for example, 1 year's lag reduces benefits more if it comes after 7 years than after 3 — then the costs of farmers' errors

Table 6. 10-year present values (US$'000s) for benefits (1% increment) from extensive surveys, with and without 4-year benefit lag, at an original income of $300/person and an original growth rate of 3%/year.[a]

	Discount (%)	
	12	24
Income at 3% growth	1952	1244
With 1% increment, no lag	2049	1295
With 1% increment, 4-year lag	1984	1258
Cost	728	556
Breakeven target population ('000s)		
No lag	7.49	10.71
4-year lag	22.48	38.88

[a] The values for income are the 10-year present values of per-person income under the conditions assumed for original income level and growth.

at the design stage are smaller than at later stages. Because researchers using extensive methods spend comparatively little time with the same farmers, they probably suffer higher costs in terms of farmers' errors than do researchers using intensive methods.

Costs of farmers' participation during the execution stage are generally in the form of unwanted variation in test results. A common error is spreading fertilizer on unfertilized treatments. If this error is known — e.g., if fertilizer is observed in an unfertilized treatment — the researchers can offset it, for example, by using regression analysis, which does not require equal numbers of observations per treatment. This kind of error is damaging in analyses that require equal numbers, such as paired comparisons.

Execution errors, like design errors, prolong the research period and delay the benefits to the target populations. Their distribution depends more on how much input the farmers have (more participation, more error) than on the survey method. I doubt that any village-based research is free of such errors. Although the errors cannot be eliminated, they are likely to be fewer (or at least more likely to be recognized and allowed for in an analysis) in a long-term than in a short-term project because the researchers and farmers have time to identify and eliminate problems in implementation.

Farmers' errors at the analysis stage are similar to those at the design stage. Farmers give biased answers to questions about technologies, probably because they think the researchers want to be told their technology is an improvement. These errors are harmless if there are objective checks on farmers' answers. No one should draw conclusions about yields or about adoption solely from farmers' declarations.

Farmers' errors that introduce random variation into test results increase the sample size necessary to make inferences about a given population. Increased sample size means increased costs and a reduced number of agricultural populations that can be covered with given resources. The bias in farmers' responses at the design and analysis stages increases costs by necessitating expensive objective checks. In the case of crop yields, for example, I have found that farmers understate yields at the design stage and exaggerate them, at least for "improved" packages, at the analysis stage. Uncorrected, these biases increase research costs by making unpromising approaches look better than they are.

Conclusions

My principal conclusions are simple:

- Intensive and extensive methods of research differ little in annual cost/sample unit. Further, they share an approach to farmer-based agricultural research, and they share many cost elements.
- The greatest cost in both methods is for internationally recruited staff. This element far surpasses the costs of local personnel, equipment, or materials and is much more important than assumptions made about discount rates used to value future costs. If this cost can be reduced, then cost comparisons are in favour of extensive methods of research.

- Research should be located in the most favourable areas, if costs and benefits are the criteria: the expected return to research is likely to be greater there and the variance of returns is probably smaller there as well. In West Africa, the distribution of rural income between regions is fairly egalitarian so that regional differences in income distribution should not be too important in the choice of research location.
- The urgency of results from research has little effect on the choice of methods. Although extensive methods have about a 24% advantage in total costs over a 5-year research project, this advantage is not much affected by the rate of discount used to value future costs. Therefore, if the rate of discount reflects donors' impatience for results, one cannot say that even high rates of impatience will make one set of methods better than the other.
- Although lags in benefits from farming-systems research have a large effect on the sizes of the target populations necessary to repay research investments, they do not much affect the choice of research technique. This conclusion, like the previous one, depends on the similarity of costs between extensive and intensive methods.
- Farmers' errors in farming-systems research increase random variation in tests at the execution stage and introduce bias at the design and analysis stages. These errors postpone research benefits and, therefore, increase the target populations necessary to repay research costs. Because the size of the targets is sensitive to benefit lags, reducing farmers' errors is important in controlling costs. There are two ways to reduce errors: use objective methods of analysis to verify farmers' evaluations of technologies, especially about such critical variables as crop yields; and have ample test replicates so that execution errors do not drastically diminish the usefulness of statistical analyses.

Implications

The principal implication is the need to spread the high costs of internationally recruited staff over larger target populations. This is the fastest way of reducing the high cost of research and of extending its benefits. This need is more or less independent of the choice between extensive and intensive methods. It means that much more effort should be made to create standard questionnaires and minimum data sets for extensive surveys (along the lines developed by CIMMYT) to define research zones, whether the extensive surveys are ends in themselves or preliminaries to intensive surveys.

Second, standard questionnaires should be entered into standard data bases accessible to researchers from different zones so that comparisons across zones and years can be easily done. Such comparisons are crucial to an understanding of the fundamental economics of rural areas, without which the research is location specific and anecdotal.

Third, the comparative advantage of research returns in favourable areas argues, analytically, for a concentration of expensive research there and for a concentration of cheap research in the unfavourable areas. Unfortunately, this conclusion is politically unacceptable because fundamen-

tal research is expensive and is needed in the unfavourable areas. One possible approach is for international research investment to be concentrated in unfavourable areas and national efforts in more favourable areas.