A Note on Multi-Stage Sampling with List Frames: The PARC/CIMMYT Survey of Wheat Harvest Technology

MELINDA SMALE
AGRICULTURAL ECONOMIST

PARC/CIMMYT Paper No. 87-9

Pakistan Agricultural Research Council/
International Maize and Wheat Improvement Centre
Collaborative Programme

1987

A Note on Multi-Stage Sampling with List Frames: The PARC/CIMMYT Survey of Wheat Harvest Technology

Melinda Smale
Agricultural Economist

PARC/CIMMYT Paper No. 87-9

Pakistan Agricultural Research Council/
International Maize and Wheat Improvement Centre
Collaborative Programme

1987

Contents

Preface

Social scientists of various institutions in Pakistan conduct numerous farm-level surveys to contribute micro-economic information to policy analysis and agricultural research decision-making. Unfortunately no reliable sampling frame exists at the village level to facilitate random sampling. Hence many social scienties use non-probability sampling methods, which do not allow inferences to be drawn about population parameters.

In the study reported here, the team made a deliberate effort to use probability sampling methods and statistically valid criteria as the basis of the survey design. The discussion of the difficulties encountered and lessons learned should be a valuable guide to other social scientists wishing to improve sample survey methods in rural Pakistan.

Derek Byerlee
Director
Economics Program
CIMMYT

Acknowledgements

T. J. Byram, Eric Waldhaus, and Jamil Rajput of the Agricultural Data Collection (Food Security Management) Project were generous and patient in advising us about the statistical aspects of the survey. Nasir Khan and Ismail Khan, CIMMYT's driver/field assistants, persevered so that the team could meet its design requirements. Without the kindness and cooperation of the farmers and laborers of Gujranwala District, the survey would not have been possible.

A Note on Multi-Stage Sampling with List Frames: The PARC/CIMMYT Survey of Wheat Harvest Technology

Introduction

In collecting farm-level data for the purpose of initiating social science research on selected cropping systems, many researchers in Pakistan use random sampling of villages with non-probability sampling of farmers within villages. Given typical timing and logistical constraints, this relatively low-cost method often provides sufficient data to meet information needs. Studies generated through non-probability sampling provide useful background information, and often serve to identify critical issues for further research.

When the purpose of a survey is to generate estimates of population characteristics for policy-makers, the reliability of the estimates becomes more important. The major shortcoming of non-probability methods is the researcher's inability to calculate sampling errors and confidence intervals because the probability of selecting the elements of the sample is unknown. For this reason, such sampling methods are often termed "nonmeasurable" or "judgment" designs (Hansen, Hurwitz, and Madow, pp. 8-9). In many research settings, however, non-sampling errors constitute a large part of total error of estimates. Non-sampling errors are common to both probability and non-probability surveys, and their magnitude and direction can rarely be measured (Casley and Lury, pp. 86-88).

During March and April of 1987, the PARC/CIMMYT Collaborative Program in Economics sponsored a pilot survey on wheat harvest technology in the rice-wheat production area of the Punjab. In the harvest technology survey, the team experimented with a form of multi-stage probability sampling using list frames. While the sample design described in this note reflects specifically the type of information needs identified for this survey, the lessons learned in the implementation of the method may have general implications for similar sample surveys undertaken by Pakistan's social science researchers.

Objective of the Survey of Wheat Harvest Technology

The objective of the study was to contribute to the existing body of knowledge on custom combining in Pakistan through a pilot survey in the irrigated rice-wheat production area of Punjab. The information needs identified fell into the following general categories:

1) updated comparison of harvest wage and combine rental rates;

2) information on other cost factors and perceptions affecting farmers' choice of harvest technology, including the costs of harvest delay;

3) updated information on the employment characteristics of wheat harvest laborers; and

4) descriptive information on how wages and contracts are negotiated.

Three population subgroups were identified as sources of data: (1) potential combine-users, (2) actual combine-users, and (3) wheat harvest laborers. Operational definitions of these subgroups are described below.

Budget and Time Constraints

Survey objectives were to be met with a modest budget, and within a time period of two-three months total for data collection and two months total for analysis and report preparation. The survey team was composed of 4-6 enumerators (economists and rural sociologists), two driver/field assistants, and one survey designer (economist).[1]

Another minor logistical consideration was the desirability of completing the interviews before the commencement of Ramazan, the Islamic month of fasting. Wheat harvesting is one of the peak labor seasons of the crop year, and in 1987, harvesting work began only one week before Ramazan. Respondent burden and workload considerations affected the questionnaire design more than the sample design, although the overall laborer sample size was reduced by 50 percent in order to complete work before Ramazan.

Selection of Survey Area

Irrigated rice-wheat production in Pakistan is concentrated in Gujranwala, Sheikhupura, Gujrat, Lahore, and Sialkot Districts of the Punjab. The team purposely selected Gujranwala district as representative of the zone. Secondary sources suggest that, among rural populations of the Punjab, the Gujranwala population has recently enjoyed a comparatively rapid increase in rural wages. Rising rural wages in the District have accompanied the development of local industries and supporting services, and favorable agricultural growth rates.

Informal interviews with combine rental companies also indicated that they presently concentrate and are likely to expand operations in this District. Several company representatives stated that although the farm size characteristics of the wheat-cotton area of Southern Punjab are best suited to wheat combining, the need for combine companies to reduce combine travel time and client transactions costs implies that, in the near future, a large number of companies will continue to combine wheat in areas where they can also combine rice. [2]

Original Sampling Plan

Definition of Population Subgroups

The team decided to attempt a probability sample of each of the three population subgroups cited above. Combine-users were defined as farmers who had rented a combine at least once to harvest either rice or wheat before the 1987 wheat harvest. Potential combine-users were defined as farmers operating at least 25 acres (10 ha.), who had never used a combine to harvest either rice or wheat prior to the 1987 wheat harvest. With this definition, 'potential combine-users' are actually large-sized hand-harvesters or reaper-harvesters. Wheat harvest laborers were defined as laborers cutting wheat for the same hand-harvesters during the 1987 harvest. The definitions are explained in the following paragraphs.

Given the brief time frame of the study, the survey was originally designed to collect data only on wheat harvesting. In the informal survey and pre-test, the team discovered some potential relationships between the demand for rice combining and the demand for wheat combining. Rice is the primary commercial crop in the District. Consequently, the team decided to define a combine-user as a farmer who had rented a combine to harvest either crop. The team designed the questionnaire for more detailed data collection from wheat combine-users, with the idea that more comprehensive data could be collected from rice combine-users in a subsequent study. No farm size limitation was imposed on combine-users, and combine-owners were not included in the population.

Through informal surveys with custom combine companies, the team attempted to formulate a profile of potential combine-users. In the early years of custom combining, many companies were unable to impose lower limits on the crop area of their clients because of the need to promote the new technology and meet short-term operating costs. Each combine rented by the companies has a capacity of approximately 50-60 acres per day, but because of land fragmentation, small field sizes, and variation in road quality, many combine companies stated that they averaged only about 25 acres per day. This minimum requirement can be most

easily satisfied by harvesting one 25-acre block per day. Over the long-term, companies will be obliged to meet at least this quota in order to cover their costs. One square of land, or 25 acres, also represents one of the bounds on farm size classes used in the Pakistan Agricultural Census.

The survey designer decided that 25 acres represented a rough approximation of the minimum farm size of potential clients in the coming years. The team did not attempt to impose a lower limit on fragment size, harvested acreage in either crop, or distance of fields from finished or metalled roads. Although these additional criteria would have improved the definition of the potential-user population, they would have dramatically increased the cost of list frame construction.

Accurate identification of labor households and labor market participants often requires a costly preliminary village census. To save time and reduce costs, the survey designer decided to use the potential-user as a cluster of actual wheat harvest labor participants. From a census of the hired harvest laborers cutting wheat on each potential-user's farm, a sample of participants can be drawn during the harvest. Using this definition, the laborer sample represents the population subgroup that will be displaced if each potential-user shifts from hand-harvesting to wheat combining.

Sample Size and Cluster Size

Overall sample size was limited by cost and logistical considerations, or a modest operating budget and staff. A total sample size of 280 farmers and laborers, or 40 farmers per subgroup and 200 total laborers was considered sufficient to conduct simple statistical tests on subgroup means and proportions.[3]

The team planned to draw village-clusters at random from a list of villages located in areas where combine use was concentrated, draw farmers from a list of all potential and actual combine-users in each village-cluster, and draw wheat harvest labor participants from the labor group employed by each potential-user at harvest time. Despite associated difficulties in estimating sampling errors, a cluster design was selected because of its low cost where little is known about the population of interest and list frame construction is necessary.

Within a brief time period, and with two vehicles and 4-6 enumerators, a maximum of only 10 primary sampling units was feasible. With 10 primary sampling units, each unit would provide an average of 4 secondary units per cluster (per subgroup), and yield 40 total laborer clusters with an estimated average of 5 laborers per cluster.

Sampling Frame

Originally, the team planned to interview a random sample of custom combine companies operating in Gujranwala District, and either obtain client lists or determine areas of concentration of clients within the District. Companies were hesitant to provide this information for the following reasons:

1) most had been operating for only one or two cropping seasons, and had not yet developed a fixed clientele or well-defined circumference of operation;

2) with the rapid increase in the number of companies operating in the area, particularly in rice combining, heavy competition meant that the information was considered to be confidential.

The hesitation of the companies to provide this information reflects, in large part, the youth and competiveness of the industry. Those company representatives willing to be interviewed (nine of the 12-15 believed to be currently operating in the rice-wheat area) suggested that, within Gujranwala District, clients were scattered throughout the Hafizabad area, Wazirabad and Gujranwala Tehsils. The team therefore adopted a different sampling frame, which is described in the next section.

Actual Sample Design

First-Stage

Frame and Sampling Procedure: The team used the 1981 District Census Report for Gujranwala District to construct the first-stage sampling frame. The 1981 District Census provides the most recent listing of administrative subunits. The District is divided into Gujranwala, Wazirabad, and Hafizabad Tehsils, with each Tehsil subdivided into urban and rural administrative units. Rural administrative units are classified under QH in alphabetical order. Within each QH, the administrative units are subdivided into Potwari Circles, also listed in alphabetical order. Each Potwari Circle consists of from 1 to 7 villages, and varies in population size and farm acreage.

The team chose the Potwari Circle as the appropriate first-stage sampling unit. [4] From the full 1981 list of 209 Potwari Circles located in Hafizabad QH, and Wazirabad and Gujranwala Tehsils, 10 were selected using a random number table. The first-stage sample consisted of the clusters of combine-users and potential-users registered within the 10 Potwari Circles.

Outdated Frame: In two of the 10 Potwari Circles, administrative changes since the 1981 Census had resulted in a split of one Potwari Circle into two. The team retained the 1981 definition as suitable for representing a cluster of farmers, and listed the farmers whose land was considered to be located in the villages included in the 1981 Potwari Circle. Changing the Potwari circle definition in the course of the survey would have altered the probability of cluster selection. If the parameters under study had been administrative in nature, or related to the Potwari Circle definition, the outdated list would have created analytical problems.

Second Stage

Frame Construction: With the intent of saving time and reducing costs, the survey team combined the second-stage frame construction and the questionnaire pre-test procedure. During the same field trip, some team members compiled the list of all combine-users and potential-users in the Potwari Circles while other team members pre-tested the questionnaire.

To construct the frame, the team had originally planned to contact the Potwari in each sampled Circle. In the first few Circles, the team was unable to locate the Potwaris, many of whom resided outside the Circles. Subsequently, the team decided to obtain the information by consulting Lumberdars, Village Union Council Chairmen, and larger farmers.

Several representatives from these groups were usually available to furnish the information for each Potwari Circle. In Circles with strong leadership, the informants sometimes stated that they were able to provide the list for the entire Circle. In other cases, the team visited each village located in the Circle. The quality of the list frame would undoubtedly have been improved by consulting the Lumberdar in each village of the Circle, but the procedure would have required more time.

Combining the listing procedure with the pre-test activity complicated the logistics of the field work. Team members had conflicting objectives, and accomplished neither activity smoothly. Some team members were obliged to return to several of the Potwari Circles to verify list frames. The team concluded that no cost savings resulted from combining the list frame and pre-test activities. If time permits, separating the construction of the list frame from the pre-test activity would be preferable.

The latest Pakistan Agricultural Census indicates that about 11 percent of the farms in Gujranwala District are greater than or equal to 25 acres in size. The fact that large farmers are well known in the Potwari circles greatly facilitated the listing process. The frame constructed from 10 Potwari Circles revealed 235 potential-users and 99 combine-users. Listing of larger

population subgroups, such as a census of all farmers in a Potwari Circle, might not be feasible in a brief time period.

Sampling Procedure: To select the sampling procedure for the second-stage, the survey designer consulted sampling statisticians working for the U.S. Department of Agriculture Statistical Reporting Service and stationed in Islamabad. The lists, or census of the two population subgroups within the Potwari Circles, indicated wide variation in cluster size. The number of potential users in a cluster ranged from 1 to 49, with an average of 23.5. The number of combine-users in a cluster ranged from 0 to 38, with an average of 9.6. [5]

The sampling statisticians recommended setting a lower limit on the number of units drawn from each cluster, with a constant sampling proportion from remaining clusters. The total sample of 42 potential-users and 41 actual-users was distributed according to this recommendation among the clusters. Where the population in a subgroup cluster was 2 or less, a census was taken. Where the population of potential users was greater than 2, the cluster was sampled at the rate of 1 in 6 (.17). Where the population of combine-users was greater than 2, the cluster was sampled at a rate of 2 in 5 (.4).

Listing Errors: The team returned to the Potwari Circles one week after frame construction and pre-testing to conduct the formal survey. Out of 42 potential users, 2 did not meet the definition. One farmer operated only 8 acres, and the other farmer had used a combine on some of his rice acreage in 1986.[6]

One farmer was unwilling to participate in the interview. His refusal apparently resulted from a fear that the team represented the ADBP and had arrived to request loan repayment. The team believes that a second attempt to contact him, after he had been fully informed about the questionnaire content by the other farmers interviewed, might have resulted in his cooperation.

The final sample of potential users contained 39 out of 42 originally sampled, which the SRS statisticians considered as acceptable. The team concluded that the listing procedure was fairly accurate for potential-users, and that with additional time devoted to listing, the frame would have been more reliable.

Out of the original sample of 41 combine-users, 4 had never used combines. The team believes that Potwari Circle informants were less capable of accurately listing combine-users because custom combining was often a new phenomenon in their area. The sample survey data confirms that most combine-users had rented combines for the first time in 1986. Although the SRS statisticians did not consider 4 listing errors to be a severe problem, the listing procedure does not appear to have been as successful with this population subgroup.

Other Difficulties Associated with Random Selection of Farmers: Other difficulties associated with the sampling procedure can be more easily solved in future surveys. In Potwari Circles with strong local leadership, some farmers were irritated by the random sample procedure. Large, influential farmers were insulted that they had not been selected. The team decided that, in these cases, any farmer who desires to be interviewed should be interviewed. Data from farmers not included in the random sample can be retained for background information but not used in any statistical analysis. The SRS statisticians expressed familiarity with the problem, and recommended the same solution.

The recommended solution obviously raises the cost of the survey effort. An intermediate solution, which the team adopted in several cases, was to conduct informal interviews with a group of farmers who were not included in the sample while enumerators formally interviewed those included in the sample. Major difficulties were encountered in only 3 of the 10 Potwari Circles. In the other 7 cases, either an influential farmer was already included in the random sample or local leaders were not offended.

On several occasions, enumerators encountered resistance from villagers when they sought an individual farmer by name, especially in smaller villages with poorer populations and less centralized leadership. The formal letter of authorization from PARC was useful only in some instances. The presence of the expatriate, the villagers' perception that the vehicle belonged to ADBP, and the enumerators' requests for a specific farmer, caused some suspicion of the team's intent. To protect themselves, farmers and villagers sometimes told the enumerator that the farmer he or she sought lived elsewhere, or was not available. Enumerators often found that after lengthy discussion, or on a subsequent visit, the farmer appeared and was willing to participate in the interview. This type of difficulty is not insurmountable--a more experienced team and more discussion time would have allayed most of the villagers' fears.

Other problems were not specific to the sampling method. Some farmers were annoyed by the time required to complete the interview (interview time averaged about 35 minutes), or were displeased that previous surveys had not resulted in an improvement in their well-being. These difficulties occurred in a minority of cases, and can be expected to occur under any circumstances.

Third Stage

Sampling Frame and Procedure: The total laborer sample size of 200 was reduced by 50 percent because of the desirability of completing the survey before Ramazan with only 4 enumerators. The frame for sampling wheat harvest laborers was provided by the potential-user subgroup survey. At the commencement of the wheat

harvest, the team returned to each hand-harvester to interview wheat cutters. Of the 39 potential-users, 3 farmers used reapers to cut all their acreage or employed family labor in 1986. The number of cutters hired by the remaining 36 farmers in 1986 was used as an estimate of the total number of laborers in the farmer clusters in 1987. Based on the 1986 figures, a sampling proportion of 1 in 3 yielded a laborer sample size of 100.

Farmers were usually unable to list the laborers employed in their fields on a given day by name, and the team felt that requesting names might have created some uneasiness among the laborers. The team therefore chose to sample laborers as they cut wheat in the farmer's field. When the team arrived at the farm, they contacted the farmer, explained the purpose of interviewing his cutters, and requested the number of cutters employed in his fields on that day. The team drew a sample of 1 in 3 from a random number table, and identified the laborers corresponding to those numbers as they counted from left to right in the field. If hired laborers numbered fewer than 3, the team took a census within that cluster.

While some team members interviewed laborers, other team members requested updated harvest wage information from the farmer and measured the moisture content of the wheat cut on that day. Average interview time was 15-20 minutes for laborers, and only 5 minutes for farmers.

Changes in Harvest Technology: With the sampling technique used in the third-stage, actual sample size was unknown until the team reached the fields. Of the 36 farmers who hired casual laborers to cut their wheat in 1986, 6 decided to combine wheat in 1987. This change reduced the total number of laborer clusters to only 30, and decreased the actual sample size.

Harvest Delays and Unusual Rainfall: In most years, wheat harvesting begins during the last ten days of April. When the team began the laborer survey, many farmers had not yet begun to harvest, few interviews could be completed, and the team was obliged to return to the farms at a later date. As the temperature increased, the wheat dried rapidly, and more farmers began to harvest. However, many of the farmers had difficulty locating laborers, and the team was again obliged to circulate among farmers as they waited for the commencement of delayed harvesting. In some cases, farmers reported that a labor group had begun to harvest their fields, but left to work for other farmers. These logistical difficulties for the team were an expression of the labor contracting problems experienced by the farmers, and provided more evidence of existing incentives for adoption of combines in Gujranwala District.

Labor contract delays resulted in the need to continue work during Ramazan fasting. Two days after the commencement of Ramazan, unusual windstorms and abnormally high rainfall struck the wheat crop in the District. The team was unable to reach the last two farmers by dirt road. The harvesting delays that

resulted from the storms were lengthy, and the team could not complete the last two clusters.

Conclusions

In the survey of wheat harvest technology, first-stage sampling units were Potwari Circles. The Potwari Circle, which contains several villages, is a useful cluster for a study of combine-users because of certain characteristics of the custom combine market. For most studies of farmer characteristics, villages provide a natural and suitable cluster of elementary units. Listing farmers within villages requires less time and fewer resources than listing farmers within Potwari Circles, and for this reason, second-stage frame construction is less cumbersome with a village cluster.

The team found that, as a result of administrative changes, their 1981 list of Potwari Circles was outdated. Outdated or inaccurate list frames create sampling bias through altering the probability of selection of elements. In the wheat harvest survey, all villages represented in the previous designation could be identified, and the probability of selection of first-stage clusters was not affected by the change. Fortunately, administrative designations were unrelated to the characteristics under study.

In second-stage frame construction, the team did encounter list frame errors that could not be remedied. The difference in numbers of listing errors for combine-users and hand-harvesters indicates that a list frame is less efficient for some types of surveys. The list of medium and large farmers was fairly reliable because these farmers are generally well-known by leaders of the area and represent a small proportion of the total farmers in the Potwari Circles. The list of combine-users was less reliable because custom combining is a new phenomenon and a farmer's choice of harvest technology is not as well-known to community members as the size of his landholding. If a research team seeks to study all farmers in the village, or small farmers in the village, list frame construction may require substantial time and resources.

Other forms of random-sampling with known probability of selection exist, including types of area sampling, and these merit serious attention in many research settings. Area frames are particularly suitable to analysis of field-specific data, such as yield response. Area frame construction generally requires heavier initial expenditures than list frame construction, but the same frame can be maintained at lower cost and used in successive surveys of various field characteristics.

List frame construction and random sampling created uneasiness among farmers and laborers in some of the Potwari Circles. The uneasiness can be alleviated by interviewing both the farmers included in the random sample and those desiring to be interviewed. Although this solution may raise the total cost of the survey, a larger proportion of total cost is represented by travel to the cluster than by travel and interviewing within the cluser. Once the team has arrived in the village, relatively little extra cost is incurred by a few unscheduled interviews.

The fact that village leadership insisted the team interview certain farmers suggests that without random sampling, some selection biases may occur. The selection bias that accompanies non-probability sampling was more evident during the laborer survey. Data gathered during the questionnaire pre-test, when the team interviewed "available laborers" two weeks before the wheat harvest, appear very different from those obtained with the probability sample.

Among the sampling stages, the third stage was the most awkward. Most of the problems in the third stage were caused by logistical factors, such as unusual storms, and labor contract delays. These factors are unique to the wheat harvest study, and are not likely to be encountered in most survey designs. On the other hand, the sampling method was a fairly accurate way of identifying actual wheat harvest labor participants. The cost and errors involved in the alternative--a village census of laborer households followed by a sample of wheat harvest laborer households--might have been greater.

The sample design adopted for the survey of wheat harvest technology resulted in a total data collection cost of approximately Rs. 75,000. Including all questionnaires, the average cost per questionnaire was about Rs. 385. An average of 3 questionnaires were completed per person, per day. Three-fourths of the total cost was incurred in per diem, lodging and travel expenses to and from clusters. The rate of questionnaire completion is low relative to non-probability samples. Both the high displacement costs and the low rate of questionnaire completion reflect the need for repeated trips to Potwari Circles to arrange interviews with the specific farmers identified by the random sample procedure. Total costs and the rate of questionnaire completion could be modified dramatically by arranging for accomodations in or at closer proximity to the Potwari Circles.

The calculations do imply that for the same total budget, a non-probability design would yield a larger sample size. However, comparing sample sizes between non-probability and probability designs has little meaning. Since sampling errors are affected by sample size, and sampling errors are not measurable with non-probability samples, the effect of increases in sample size with non-probability samples is indeterminate. For probability samples, large gains in precision can result from

increasing the subgroup sample size (proportion of sample with characteristic under study) to 200. With cluster designs, larger sample sizes are needed to obtain the same level of precision as with a simple random sample. For the wheat harvest study, this would have implied a minimum of 600 interviews, which was not feasible with the available resources.

The final consideration, then, is whether a probability sample is better than a non-probability sample. The answer to this question lies with the type of information needed. If a researcher intends to present statistical inferences about a population to policy-makers, a probability sample is necessary because it provides a measure of the reliability of the estimates. Even with a small probability sample, some inferences can be stated with confidence. If, on the other hand, the researcher's goal is to identify key issues for further research, a combination of extensive informal work and a non-probability survey of a well-defined population subgroup is likely to be a more efficient use of resources.

Notes

[1] Survey costs totalled approximately Rs. 75,000, excluding the salary of the survey designer and computer analysis charges. Other salary costs represented about 25 percent of the total. Per diem, diesel, other field expenditures, and miscellaneous supplies composed the remaining 75 percent of total costs. The informal survey, pre-testing of the farmer questionnaire and list frame construction, the formal farmer survey, and the laborer survey each required about a week of field work. Each field activity required a week of office preparation. Estimated cost per questionnaire (including combine-user, potential-user, laborer, and wage update/moisture content schedules) was about Rs. 385.

[2] A total of 29 custom combine projects had been financed by the ADBP as of January 1987, while additional loan projects awaited approval. The estimated number of self-propelled combines in operating condition was approximately 250-300. Discussions with ADBP representatives indicated a total of 12-15 companies currently operating in Gujranwala District, with expansion expected in the 1987 crop year.

[3] One rough guideline in small sample surveys is a minimum of 20 to 30 observations per characteristic, with some additional observations to allow for substandard questionnaires or problem cases. Although statistical research has shown that sampling errors decrease rapidly as sample size for one characteristic rises to 200 (with modest reductions thereafter), the team did not have the resources to collect data from 200 elements per subgroup. With fixed resources, increasing sample size can also contribute to greater non-sampling error by reducing resources devoted to data quality (Fowler, pp. 42-43).

[4] The informal survey that preceded the sample survey suggested that, while combine use is fairly widespread within the Tehsils, custom combine companies prefer clients in some proximity to metalled or finished roads and tend to seek clients with large acreages or villages where farmers with contiguous fields can organize large acreage blocks. The survey team concluded that the village represented too small a cluster of farmers to use as a primary sampling unit, because some villages may contain only one or two farmers with suitable acreage requirements or fields that are accessible by combine.

[5] The structure of the list frame itself suggests what was found in the informal surveys about modes of contract negotiation. In their first years of operation, companies often attempt to book large tracts of land through associating with farmer groups. Potwari Circles with a

proportionally small actual user population tended to be more isolated, and may also have contained a higher proportion of smaller, poorer farmers. Potwari Circles with a high proportion of users may have booked acreage blocks with combine companies.

[6] Some farmers owned or operated a square, but their rice and wheat acreage in 1986 was under a square because they kept land in fallow, planted a higher proportion in fodder and other crops, and in one case, shifted from crop enterprises to cattle production. These farmers were considered as potential-users. A more complete definition in terms of acreage harvested in wheat or rice would have resulted in a more accurate representation of a potential-user, but a more detailed definition would have required considerably more time during the listing procedure,and may not have been feasible.

References

Casley, D.J., and D.A. Lury, Data Collection in Developing Countries, Clarendon Press, Oxford, 1981.

Fowler, Floyd J., Survey Research Methods, Sage Publications, Beverly Hills, 1984.

Hansen, Morris H., William N. Hurwitz, and William G. Madow, Sample Survey Methods and Theory, Volume I, John Wiley and Sons, Toronto, 1953.