UNCLASSIFIED
CLASSIFICATION

PD-AAN-088
Report Control Symbol U-447

PROJECT EVALUATION SUMMARY (PES) — PART I

1. PROJECT TITLE	2. PROJECT NUMBER	3. MISSION/AID/W OFFICE
INTEGRATED AREA DEVELOPMENT STUDIES	520-0249	USAID/Guatemala

4. EVALUATION NUMBER (Enter the number maintained by the reporting unit e.g., Country or AID/W Administrative Code, Fiscal Year, Serial No. beginning with No. 1 each FY) 83-02

☐ REGULAR EVALUATION ☒ SPECIAL EVALUATION

5. KEY PROJECT IMPLEMENTATION DATES	6. ESTIMATED PROJECT FUNDING	7. PERIOD COVERED BY EVALUATION
A. First PRO-AG or Equivalent FY 78	A. Total $ 2,081,158	From (month/yr.) January 1982
B. Final Obligation Expected FY 82	B. U.S. $ 1,438,000	To (month/yr.) December 1982
C. Final Input Delivery FY 83		Date of Evaluation Review 06-22-83

8. ACTION DECISIONS APPROVED BY MISSION OR AID/W OFFICE DIRECTOR

A. List decisions and/or unresolved issues; cite those items needing further study. (NOTE: Mission decisions which anticipate AID/W or regional office action should specify type of document, e.g., airgram, SPAR, PIO, which will present detailed request.)	B. NAME OF OFFICER RESPONSIBLE FOR ACTION	C. DATE ACTION TO BE COMPLETED
1. No further evaluations should be scheduled.	C. Carr	06-30-83
2. Planning Council should be encouraged to prepare a user's guide for Studies I, II, III and Activity I.	R. McLaughlin	06-30-83

9. INVENTORY OF DOCUMENTS TO BE REVISED PER ABOVE DECISIONS

- ☐ Project Paper
- ☐ Financial Plan
- ☐ Logical Framework
- ☐ Project Agreement
- ☐ Implementation Plan e.g., CPI Network
- ☐ PIO/T
- ☐ PIO/C
- ☐ PIO/P
- ☐ Other (Specify) _____
- ☐ Other (Specify) _____

10. ALTERNATIVE DECISIONS ON FUTURE OF PROJECT

- A. ☐ Continue Project Without Change
- B. ☐ Change Project Design and/or ☐ Change Implementation Plan
- C. ☐ Discontinue Project

11. PROJECT OFFICER AND HOST COUNTRY OR OTHER RANKING PARTICIPANTS AS APPROPRIATE (Names and Titles)

Cecil McFarland, ORD

12. Mission/AID/W Office Director Approval

Signature

Typed Name: Charles E. Costello

Date: June 22, 1983

AID 1330-15 (3-79)

13. SUMMARY

The project terminated December 31, 1982. Although the studies were carried out as scheduled, recent users questioned the validity and reliability of the date generated. A special evaluation to determine the validity and reliability of the data found the data are of unknown validity and probably reflect selection biases and nonsample errors. Before using the data a comparison of datasets should be made, a user's guide prepared for understanding the data and their limitations, and the raw data of Study I should be properly weighted. Any further USAID assistance to the Planning Council in creating a data bank should focus on systems analysis.

14. EVALUATION METHODOLOGY

The special evaluation was to assess the accuracy and usefulnessof the data generated by the project. The evaluation plan in the project paper called for evaluation at the purpose level at approximately six-month intervals for one year after the project. During the last six months of project activity, users of the data questioned the reliability of the information produced by the project and, therefore, USAID/Guatemala scheduled this special evaluation to ascertain whether the material was sufficiently valid to meet the general project purpose of improving the appropriateness of public investment programs to address the highest priority needs for infrastructure and services in locations where the need is greatest.

The U.S. Bureau of the Census organized an evaluation team consisting of an agricultural statistician, a mathematical statistician, and a data processor. The scope of work and methodology employed are described in Annex I, "Statistical Review and Recommendations for the Guatemala Integrated Area Development Studies (IADS) Project", page 1.

15. EXTERNAL FACTORS

During the last six months of project activity, the National Economic Planning Council, the implementing unit, underwent a radical reorganization with its staff being reduced from 350 to 70 and its role being challenged. Two acting Secretary Generals supplied administration for the first months of the reorganization until the appointment of a permanent Secretary General in 1983.

The five-year plan under preparation by the National Economic Planning Council at the time of the March 1982 coup was shelved and thus far no plan has been published by the current

GOG. Only one agency concerned with the project INFOM continues to use the 5-year plan which used information generated under this project.

16. INPUTS

As stated in PES 82-02, the majority of project funds financed the teams of analysts, technicians, programmers and secretaries who executed the various studies and a Title XII contract with Iowa State University (ISU) to provide the GOG agencies with advisory assistance. The evaluation team found that the definition of the selected variables and their classification breakdowns were not clearly defined early in the study efforts (page 6, Annex I). Availability of qualified technical assistance on a timely basis, in-country data processing capability, uncoordinated GOG inputs, and political turmoil within the country, affected project implementation (page 7, Annex I).

17. OUTPUTS

As reported in PES 82-02, the studies "Definition of Rural/Urban Hierarchy and Inventory of Infrastructure and Services" (Study I) and "Inventory of Natural Resources and Determination of Economic Potential of Each Level of the Rural/Urban Hierarchy" (Study II) have been completed. The evaluation team notes that the original structure of activities was altered, except for Study I variables, the data cannot be considered site-specific, and Study I data validity is limited to sites with 500 or more population. Data for actual land use and individual access and travel are suspect without evidence of sufficient homogeneity, particularly in the upland municipios. To date there appears to be limited interest by GOG agencies in the data except for some use by staffs of the Planning Council and INFOM (see pages 4-6 of Annex I).

18. PURPOSE

As stated in the Project Paper, the purpose of the project is:

"The development and adoption by the GOG and non-Government agencies of a systematic planning methodology for determining priorities and allocating resources to provide the infrastructure and services required to achieve the goal." The evaluation team considered that the basic purpose -- "the development of a statistical base on which integrated rural development plans might be built" -- a worthwhile objective. Of concern was that the project started with a set of potentially useful variables without a clear methodology for decision-mak-

ing use and without sufficient attention given to systems analysis for analytical and planning use. The primary data collected under the project did not use strict probability methods and the selection of the sample and survey techniques violated the canons of accepted scientific practices. The data cannot be definitively assessed nor is it possible to evaluate their reliability or extent of error. Therefore, for the user, the data is of unknown validity, and doubtlessly reflects various selection biases and nonsample errors (Pages 14-15, Annex I).

19. GOAL/SUBGOAL

The broad objective to which this project was seen as contributing at the time of project design was:

"Increase small farmer incomes and improve well-being in rural areas." As stated on Page 59 of the Project Paper, "the desired effort at the goal level will come about indirectly after the outputs of this project are utilized to plan new infrastructure and services projects and then subsequent implementation."

The only identified user of Study I data is INFOM but it cannot use the information without first validating the data. Therefore, this project can be considered as contributing minimally to goal achievement.

20. BENEFICIARIES

Not pertinent.

21. UNPLANNED EFFECTS

Not pertinent.

22. LESSONS LEARNED

Definition of variables and their classification breakdowns need to be clearly defined early in a study effort. Development of table formats, map layouts, and specific analytical needs prior to questionnaire development will facilitate a clear understanding of what variables are being considered, what accuracy is needed, and what effort is required to obtain the data.

Limited participation of Guatemalan agencies in the development of the Project Paper and Project Agreement resulted in the failure to establish an integration of relevant Guatemalan statistical work already underway at the Ministry of Agriculture and the National Institute of Municipal Development.

Potential users should be brought into the planning stage so that their needs and concerns are explicitly taken into account in the database system design.

XD-AAN-088-A

STATISTICAL REVIEW AND RECOMMENDATIONS FOR THE GUATEMALA INTEGRATED AREA DEVELOPMENT STUDIES (IADS) PROJECT

by

H.A. Green and Enrique Gomez
International Statistical Programs Center

Anthony G. Turner
Statistical Methods Division

U.S. Bureau of the Census

February 1983

STATISTICAL REVIEW AND RECOMMENDATIONS FOR THE
GUATEMALA INTEGRATED AREA DEVELOPMENT STUDIES (IADS) PROJECT

Table of Contents

		Page
I.	Introduction	1
II.	Statistical Evaluation Concerns	2
III.	Informational Requirement and Resource Issues	4
IV.	Survey Design and Operational Issues	7
V.	Sampling and Related Measurement Issues	9
VI.	Data Processing Issues and Concerns	11
VII.	Conclusions and Recommendations	14
ANNEX I.	List of Persons Contacted	17
ANNEX II.	List of Documents Reviewed	18
ANNEX III.	List of Additional References	21

STATISTICAL REVIEW AND RECOMMENDATIONS FOR THE
GUATEMALA INTEGRATED AREA DEVELOPMENT STUDIES (IADS) PROJECT

I. Introduction

At the request of USAID/Guatemala the U.S. Bureau of the Census has agreed to provide statistical services to evaluate the accuracy and usefulness of data collected under the Guatemala Integrated Area Development Studies (IADS) project. The scope of work for this evaluation consists of:

A. Review of the sample and survey designs used for data collection and evaluation of their efficiency with reference to stated goals of the survey and the analysis desired.

B. Evaluation of the data collection procedures used and all relevant documentation, including questionnaire design, manuals, auxiliary field forms, and field editing, followup and other quality control procedures.

C. Evaluation of the data processing after field collection, including coding, clerical editing, data entry, and computer processing specifications for editing and tabulation.

D. Based on findings in Items A-C above, evaluation of the validity and reliability of the data.

The Census Bureau Evaluation team consisted of an agricultural statistician, a mathematical statistician, and a data processor. Following discussions with a former Mission RDO, S&T/DS/UD officials and contractor personnel, and a review of documentation available from S&T/DS/UD files and from the contractor, the agricultural statistician and the data processor undertook a two-week TDY to Guatemala to collect and review additional documentation available there and to interview personnel of the involved Guatemalan agencies. Copies of relevant documentation were brought back to Washington for further review and analysis by the entire team. (See Annex I for the list of persons contacted and interviewed, and Annex II for the list of documents reviewed.)

The IADS project undertook five separate but interrelated statistical activities which correspond to the five basic components of the project. These components were:

A. Study I, which was based largely on a survey of 1987 community leader groups as well as secondary data to obtain an inventory of available infrastructure and services and to delineate a rural/urban hierarchy of services.

B. Study II, which was based largely on a survey of 398 farmers as well as secondary data to obtain an inventory of the natural resource base and to determine the agricultural potential in each of 26 ecosystem zones.

C. Study III, which was based largely on a survey of 314 households representing a three subregion subsample of the 1987 Study I communities as well as secondary data to determine patterns of access, travel, and movement of goods.

BEST AVAILABLE COPY

D. Activity I, a study of local participation in development projects, which was based on questionnaire addendums to the above three surveys and an in-depth examination of three projects selected on the basis of data obtained from an auxiliary survey of 25 rural development agencies.

E. Activity II, which provided for the establishment of an information center focused on a data bank consisting of data compiled during the project. (Development of a planning method to determine investment priorities was also part of this activity).

II. Statistical Evaluation Concerns

One can generally assume that there exists no statistical effort without data error; there is no method of collecting and processing data that will be entirely error-free. The objective of data collection and processing is to provide information sufficiently accurate for the user's purpose. Errors can be tolerated if they will not seriously affect the character of decisions based on analysis of the data. Hence the objective of statistical evaluation is:

A. To determine the extent of error (accuracy) -- i.e., what is the statistical variance and/or possible bias in the data.

B. To determine the nature of error -- i.e., how prevalent is the error, where is it concentrated, what is its source. Data cannot be evaluated without detailed knowledge of how it was collected and processed.

C. To resolve error impact -- i.e., by correcting for error through imputation or other statistical adjustment; by cautioning on use and detailing validity restrictions; and by correcting future survey design/implementation to preclude such error and/or to obtain better information for imputation purposes.

Most data have errors which remain undetected; moreover, even when presence of error is detected there often are no obvious corrective actions that can be taken because the nature and/or extent of error remains unknown.

The determination of error involves both objective and subjective elements and to a considerable extent it is an art. In essence, it is a process which consists of comparing one value and how it was collected/processed with another and how it was collected/processed. It involves evaluating the likely validity of the various collection/processing procedures, including if possible an assessment of the quantitative impact of erroneous procedures or actions on the values. For the present assignment it is not feasible to examine the respective values of individual variables included in the IADS statistical effort. We can, however, examine the statistical operational design and implementation experience to the extent that documentation and personal recollection are available.

Statistical operational design and implementation involve many aspects any of which may introduce or compound error. Major aspects of design and implementation are:

A. Specification of information requirements and resource inputs

- definition of variables
- accuracy needed
- tabular format and other presentation characteristics
- availability of technical and other resource input (survey, sampling, data processing skills, facilities, equipment and supplies)
- overall statistical design
- timing of total operation and its various phases

B. Survey design and operations

- consideration of field conditions
  - unusual population and environmental characteristics
  - maps, secondary data, etc.
  - logistical/data collection support
  - institutional arrangements
- questionnaire design (in accord with principles which experience has shown to be effective in use, response to definitions of variables, likely characteristics)
- field operational aspects
  - enumerator selection and training
  - supervisor selection and training
  - listing and other sample frame operations
  - quality control measures (including validation, recheck and reinterview procedures, document flow records, etc.)
  - logistics (equipment, payroll, information flow)
- pretest of individual survey operations/procedures and pilot survey

C. Sample design and operations

- specification of target population(s)
- identification of sample frame(s) and evaluation of their completeness
- methods of achieving sampling efficiency, including stratification, clustering and optimization
- fidelity of the actual methods of selection at the various stages compared to the intended design
- specification of valid measurement techniques to assess non-sampling error and sampling error

D. Data processing design and operations

- field edit specifications and operations
- office edit specifications and operations
- coding and entry specifications and operations
- machine edit specifications and operations (including computer access and availability of services)
- data storage requirements
- retrieval access and relevance for users (including weighting, tabulation and similar operations)
- analytical operations

In the following four sections we shall examine the operational design and implementation of IADS project statistical activities in terms of their consideration of these aspects.

III. Informational Requirement and Resource Issues

Project Purposes

The detailed and unambiguous specification of information requirements was a task which should have but apparently did not occur at the beginning of the project. To understand this ambiguity concerning the statistical purposes of the project, some detail as to what the project initially entailed and how it evolved is of interest.

As initially conceived in the Project Paper (pp. 15-25) and modified as indicated in a Memorandum of Agreement between USAID/Guatemala and S&T/DS/UD (pp. 8-12), Study I was to provide: a spatially defined inventory of rural infrastructure and services existing in 157 "submarginal" and 49 "marginal" (later 52 for a total of 209) municipios; a series of maps and overlays visually presenting this inventory data; a computerized system for data storage and further processing and analysis, including programs developed for retrieval of the data; a manual explaining the preparation of maps, overlays, and computer classification system; and a final report describing the results of the spatial analysis, including analysis of the hierarchy of the various levels of market places and surrounding areas of influence.

Study II, as initially conceived and modified, was to provide: a spatially defined inventory of the natural resource base of the Study I municipios (including present and potential crop production alternatives); a computerized system for data storage, processing and analysis, and retrieval; a series of maps and overlays presenting the natural resource data; a manual explaining the methods used in collecting and integrating the data as well as preparation of the computer classification system and retrieval programs; a report on analysis of potential on-farm resource improvement projects as a result of an initial screening of the natural resource base, present production patterns, and production potential; and a final report describing the data collected, the results of analyses, and possible applications in establishing investment needs and priorities.

Information forthcoming from the above two studies was to be classified and mapped in such a way that analysts could determine the suitability of an investment according to: (i) the economic potential of the natural resource base at each investment site; (ii) the need for specific types of public infrastructure and services in each site; and (iii) the suitability of locating an investment at each site given the spatial relationship of that site to the surrounding localities. The project was to provide an initial screening of investment possibilities, applying appropriate economic and social criteria, to establish priority investment programs and determine the most indicated sites for such investments. This methodology was to be applied by one or more Guatemalan agencies to the practical problem of choosing infrastructure or service interventions and their locations.

The cited MOA indicated the need for data on additional variables: (i) rate of utilization of urban-based services such as schools, health services, water and sewage systems, public market structures, local transportation systems, electrical energy, agricultural extension and credit services, and similar off-farm agricultural support; (ii) local prices of a broad spectrum of agricultural inputs and some measure of local previous year prices for agricultural products; and (iii) more detailed breakdowns of commercial functions than were initially considered.

In addition, the MOA specified an intensive and more thorough analysis of two areas comprising four or five contiguous municipios and designed to determine the tendency to travel varying distances (i) to acquire or obtain access to each of the full range of urban-based infrastructure and services, and (ii) to obtain inputs or sell products. Data collection for this purpose was to involve up to 200 in-depth interviews with farmers over at least ten different randomly selected subareas of each area. The project would also provide for creation of an extensive and varied set of experiments with local participation in planning, involving a stratified sample of 20 to 25 communities differing in size, access, ethnic composition, and economic base. There was also to be an expanded analysis to relate utilization and perceived needs as functions of household socioeconomic characteristics, presence and attributes of services and infrastructure, and accessibility to such, based on the urban infrastructure and services data (Study I), utilization and pricing data gathered as an augment to Study I, the access and travel information from the access study (termed "Study III in the MOA), and the needs expressed in the local participation experiments.

Due to concerns mentioned in the MOA as well as implementation concerns perceived after promulgation of the Project Paper, the structure of activities was substantially altered. Specifically, project statistical activities were restructured as indicated in the Introduction above due to these MOA concerns as well as such factors as: the limited participation of Guatemalan agencies in the development of the Project Paper and Pro/Ag (which in consequence failed to establish an integration of relevant Guatemalan statistical work already underway at the Ministry of Agriculture and the National Institute of Municipal Development (INFOM) as well as a valid reflection of the institutional structure of Guatemalan agencies); the multiple contribution of successive RDOs in the Mission; the substantive and organizational preferences of the technical assistance contractor; and the political, economic and social disruption in the country during much of the project period. A survey of farmers was explicitly incorporated in Study II. A household survey was included to obtain information on access, travel, and the movement of goods (Study III). To provide the information desired for Activity I local participation questions were appended to the survey of community leaders (Study I), to the agricultural production survey (Study II), and to the survey of households or individuals (Study III). Finally, the development of a planning method to determine investment priorities as well as a computerized data bank was formally instituted as Activity II.

Data from the three studies including the local participation addendum were to provide the basis of a planning methodology which then would be utilized: (i) to identify gaps in infrastructure and services within the municipalities relative to their natural resources and productive capacity; (ii) to establish priorities for investment in additional infrastructure and services in these rural areas; and (iii) to relate priority investments to locally-perceived needs. The project results should permit governmental and non-governmental agencies (including international donors) to utilize the data and the planning methodology to: (i) define with precision the location, nature, and extent of the requirements for rural infrastructure and services; (ii) using agreed upon criteria, determine the relative priorities between locations and alternative investments in infrastructure and services; and (iii) determine the roles, responsibilities and coordination required among governmental institutions operating in the rural areas to assure complementarity between their activities and to avoid duplication of efforts.

Project Outputs

It is not our task to evaluate the usefulness of the selected variables and their specificity. Nevertheless, in our judgment the study designs do not reflect these purposes well. The selection of variables appears to have slighted many important economic factors such as prices and employment (especially non-wage). Data are not sufficient to derive even elementary price and cost structures - a primary goal of the project - and without which development potential cannot be adequately evaluated. Variable definitions do not appear sufficiently detailed to provide planners with the commodity, spatial, ethnic and other market structure information desirable for economic interventions. For example, one would expect wide fluctuation of commodity prices over space and time - particularly during political, economic, and social unrest - which simply has not been measured. Microclimatic conditions, practices, and significant detail concerning viable small farm agricultural systems, which might indicate a need for agricultural input and product market interventions, similarly appear to have been overlooked.

Except for Study I variables the data cannot be considered site-specific, in large part but not entirely due to sampling design (discussed further below). Moreover, Study I data validity is limited to sites with 500 or more population. One suspects a fairly high level of heterogeneity in data from the aldeas and caserios (especially as a function of size), and the somewhat arbitrary selection of these smaller communities gives us no reason to expect a selected village to reflect those throughout the municipio with sufficient accuracy for information purposes. Data for actual land use and individual access and travel are similarly suspect without evidence of sufficient homogeneity, particularly in the upland municipios.

The data have found some use by staffs of both the National Economic Planning Council (CNPE) and INFOM. INFOM appears to more readily accept the existence of errors and the need for validation as a matter of course in data use. Hence the possible existence of errors in the data is of less consequence. CNPE appears to require turn-key validated data and hence the dataset is of less usefulness. There appears to be limited interest to date by other agencies. These reactions to the IADS project outputs illuminate the need to bring potential users into the planning stage so that their needs and concerns are explicitly taken into account in the database system design. For example, at least two factors in addition to general validity concerns appear to inhibit interest in the IADS database by the Ministry of Agriculture: (i) the lack of "user-friendly" access; and (ii) the failure to incorporate agricultural sector investment priority decision linkages (in terms of variables, detail, accuracy). In short, it is not clear why some variables were included and why others were excluded.

Definition of variables and their classification breakdowns need to be clearly defined early in a study effort. This facilitates clear understanding of who is and who is not in the survey, what variables are or are not included, and if included, how they are disaggregated. As concerns respondent definitions, this is very important in developing an appropriate sample design that will not be violated in the field. Days and weeks spent specifically laying out the exact data requirements will generally save unnecessary collection of data and hence survey resources. Development of table formats, map layouts, and specific analytical needs prior to even questionnaire development will facilitate a clear understanding of what variables are being considered, what accuracy is needed, and what effort is required to obtain the data.

Project Inputs

With respect to the responsibilities of general line agencies such as INFOM and the Ministry of Agriculture, planning and implementation in Guatemala appears to be in a rather non-integrated institutional context. The IADS project encountered this lack of integration and was unable to establish a meaningful locus of integrated direction for the statistical effort. This potential problem was compounded by the project direction shifting from INFOM to CNPE when the latter remained ill-equipped to provide the integrated focus in light of the underlying institutional interrelationships.

It appears that INFOM and contractor personnel had mutually inconsistent ideas about many aspects of the data collection operation which were not resolved. This was in part due to the late arrival of the contractor on the scene, poor timing for technical assistance trips to Guatemala by the contract advisors, and perhaps simply a lack of communication. The contractor endeavored to incorporate several changes but was not always available to follow through for implementation, and INFOM was forced to proceed in their own way in an effort to get on with the work.

The need for experienced technical input early on seems obvious. For the IADS project experienced technical input was too little too late. For example, even before the contractor was selected INFOM was well on its way toward an IADS statistical effort in light of its own interpretation of data needs; while the nonexistent Steering Committee was replaced by CNPE direction, experienced technical input remained limited. Sufficient (on-site when needed) experienced technical input is essential throughout a statistical effort if quality data are to be obtained. Unfortunately for the project, such assistance was limited not only in initial design considerations but in the design/implementation of the entire survey operation, of the sampling approach, and of the data processing system. (Although K. L. Carey and S. E. Carey, who were in Guatmala for 18 months, apparently provided some input, this was not the primary goal of the local participation activity of the project.)

While overall study duration and funding certainly cannot be considered a project constraint, scheduling and/or logistic bottlenecks and concurrent occurrrence of external events certainly took their toll. The timing of field activities as well as data processing during a period of violence and unrest in Guatemala City and the study region was indeed unfortunate. Additionally, there were reported delay problems with respect to counterpart availability, technical assistance contractor involvement, contractor personnel being on-site, access to relevant technical materials, equipment and supplies availability, timeliness of staff payrolls, and in-country data processing, at least some of which could have been avoided with proper design and scheduling.

IV. Survey design and Operational Issues

Field Conditions

As mentioned above with respect to variable selection and timing bottlenecks, field conditions were apparently not adequately foreseen. While some problems resulted from failure to adequately take into account the innate socioeconomic and physical milieu of the study area, there is little question that the social unrest and violence in the country during the project was a major factor in

creating problems. Whether the unresolved difficulties were primarily a result of this unrest or of problems in basic planning for survey operations is not known; the available evidence suggests both factors. For example, the Study I survey was intended to be administered to community leader groups. Available documentation does not indicate whether selection of individuals for these groups was in fact carried out in an unbiased manner (see comments below on listing operations and on sampling design and implementation). Indeed, there were reportedly several instances where key informants could not be located and other data gathering procedures were substituted. Response bias may also have resulted from selected informants simply not being knowledgeable. Group interviews are generally an unsatisfactory procedure in that very experienced survey personnel are required if response bias is not to be introduced. INFOM documentation reflects awareness of the potential bias here but they apparently felt such bias could be successfully circumvented; we are not so sure.

Questionnaire Design

Survey questionnaires need to be designed carefully and thoroughly in order to obtain the required data. The questionnaire needs to reflect and facilitate respondent recall and accurate response. For the three surveys skip patterns were generally easy to follow and well-placed so that unnecessary questions need not have been asked of respondents; however, some were missed and interviewers had to learn how to avoid difficulties from field experience. Question phrasing and redundancy likely also caused some ambiguity in response. In addition, the questionnaire for Study I was quite long and cumbersome -- taking over three hours to administer -- and so respondent fatique may have lessened response quaility. (Differences in concept between the Guatemalan and contractor technicians as well as problems in appropriate initial and follow-up technical assistance appear to have been factors here).

In general questionnaires should be structured to permit data entry directly from the forms with minimal coding and transcription effort. For the IADS surveys full-blown coding and transcription procedures were incorporated which likely introduced error. Quality printing of questionnaire forms is also desirable to provide legible and durable pages that are easy to handle and maintain under field conditions. Quality questionnaires also convey to both staff and respondents that the survey is an important and meaningful undertaking. The questionnaires for the IADS surveys were mimeographed on low quality paper which did little to enhance their use.

Survey Operations

Interviewers were selected and trained for each survey, and some field supervision was also provided. However, there is no solid information as to how experienced and well-trained the interviewers and/or supervisors were. Hence the quality of the field staff remains largely unknown. Better training in probing techniques, for example, might well have improved response quality.

A legitimate listing operation was apparently bypassed in all three field surveys, and consequently respondent selection may well have introduced substantial biases (see Sampling Issues below for further comment on this). Listing and other appropriate sample frame operations should be considered as important as the actual survey data collection and steps should be taken to secure the requisite effort.

While INFOM personnel and documentation indicated field check and validation (such as rechecking and comparison with secondary sources) of collected data, this apparently was not continued throughout the survey operation. Expenditures apparently outran available funds and such validation procedures as were incorporated in the survey design were bypassed; completed questionnaires were accepted as they came from the field without further verification. In addition, the secondary data used for much of the information on schools, roads, and electric power do not appear to have been validated for the basic database. INFOM is apparently validating selected statistics as they use the study data for current planning and analysis purposes. However, these corrections are not carried back to the basic database nor does there appear to be any systematic documentation of the extent and nature of the corrections required.

Complete validation should have been (and should be) continued as long as the data for variables of concern contain suspected error of unacceptable extent. Apparently the need for more complete validation was not realized until much of the field work had been completed (a design and/or scheduling problem). When error is detected, not only should the value be corrected but the cause should be determined and documented. This procedure not only permits correction of questionnaire, interviewer and/or supervisor procedures, and other aspects of field survey design/implementation but also allows concentration of validation resources where errors are most likely to be prevalent. It also facilitates the development of better imputation methodologies for subsequent cleaning of the data.

Pretest and Pilot Survey

Although it cannot solve all problems, a proper pretesting and pilot survey effort will greatly improve data quality. In spite of the IADS pretesting and pilot survey efforts, major nonsample errors did develop. Perhaps not enough attention was given to the pretest and pilot survey results, or perhaps the pretest and pilot surveys were poorly designed. For example, using experienced well-qualified interviewers in non-typical survey sites will indicate little of what may happen when less-experienced interviewers hit actual field conditions. Pretest/pilot activities should be structured to evaluate as much of the survey design and implementation procedures as possible -- questionnaire design, listing form layout, means to locate respondents (such as map clarity and accuracy, institutional arrangements, etc.), concept ambiguity among both interviewers and respondents, means to provide field checks and reinterviews, and other field and office operations.

V. Sampling and Related Measurement Issues

The methods of sampling used for the three principal studies in the Integrated Area Development Study project suffer on several accounts. As a result, there is no scientific justification for claiming that the data produced from these samples are valid representations of the populations they reputedly represent.

Of the three studies, Study I comes closest to achieving a valid design. It is inherently appealing and mathematically optimum to oversample large villages as was done for this study. It is not clear, however, why all villages over 500 in population were sampled with certainty, though it is not incorrect to do so. It may have been more efficient to stratify the villages with more refinement, applying a scheme such as the following:

villages larger than 1500 population - take all
villages with 1000-1499 population - sample 2 in 3
villages with 750- 999 population - sample 1 in 2
villages with 500- 749 population - sample 1 in 3

Unfortunately, the available documentation is not sufficiently detailed to judge whether the above scheme, or some variation, would have been more efficient. Nevertheless, the scheme actually used is valid, though perhaps inefficient, and hence the utility of data is not compromised by this minor flaw. Another deficiency with the design of Study I was a failure to match the primary sampling units with the stated objectives. The main unit of analysis was to be municipios. This means that sampling of settlements/villages should have been done for each municipio. The method that was used yielded too few settlements for some municipios, so that some of them (municipios) had to be collapsed at the analysis stage. This is again an inefficiency in the design, but a minor point with regard to whether the data are valid or useful.

There are two major faults with Study I, however, that seriously compromise its usefulness. First is the methodology for gathering the required information at the final stage of sampling. The technique used is essentially "key informant." Though there is a stated hierarchy of individuals who were to be contacted to supply the required data, we are extremely skeptical that this procedure would not quickly degenerate into a convenience sample of whoever happens to be available. When this happens, response biases of unknown direction and magnitude come into play, and we suspect that field quality control efforts were woefully inadequate to minimize this effect.

The second major problem concerns the way in which the estimates were made following data collection. The estimation procedure that was used ignored the sample design. The sample design consisted of two strata - one of certainty villages larger than 500 in population and the other of noncertainty villages smaller than 500. When samples are taken from multiple strata with different rates of selection, the resulting estimation procedure must take account of this through weighting, that is, by multiplying each stratum's raw sample counts by the inverse of its appropriate probability of selection, before aggregating the results across strata. There is no evidence in the documentation that this was done, and so the aggregated data are flatly wrong. This problem could be rectified, however, by going back to the raw data file and properly weighting them, that is, if it can be determined what the probabilities of selection were in the small village stratum. It is stated that the sample of small villages was a 15 percent sample, implying the "weight" would be about 7.5 for this stratum (the weight is 1.0 of course for the large village stratum). However, there were 338 villages sampled from a universe of over 7000 small villages, a sampling fraction of less than 1 in 20 (the corresponding weight would be about 20). Clearly there is a problem in determining the correct weight by relying on the documentation. Perhaps the small villages were not sampled with equal probability but with probability proportionate to size, or some other way. In any event, we would not recommend using the aggregated estimates as currently derived.

Related to the above point about weighted data is use of the "Statistical Package for the Social Sciences" (SPSS) software package for statistical analysis. The statistical functions that SPSS performs are based on the assumption that the sample observations were selected through simple random sampling - a condition violated by the design of Study I with its stratified sample utilizing different

rates of selection. The standard deviations calculated for Study I are therefore wrong and should not be used. This is unfortunate because correctly calculated standard errors are essential for evaluating the reliability of the estimates. Standard errors are only applicable to that portion of the data coming from small villages, since large villages were sampled with certainty and hence not subject to sampling error.

Turning to Studies II and III, the sampling technique employed in both instances is a quota sample. The overriding defect in a quota sample is that there is no way to derive the probabilities of selection of units finally included. Hence true estimates cannot be made. The chief bias likely to ensue is that some units may actually have a zero chance of selection, since there is no positive control over which units are taken. With probability methods, on the other hand, every unit in the population has a calculable and non-zero chance of inclusion in the sample. A related defect in quota sampling is that without known probabilities of selection it is impossible to calculate the sampling errors (standard errors) on the data collected, and hence there is no way to assess their reliability. In sum the data are subject to unknown biases and are of unknown reliability, and so using them is extremely risky.

VI. Data Processing Issues and Concerns

The procedures used for the computer processing or machine editing of the data collected under Studies I, II, III, and Activity I of the IADS project were very similar. This section does not deal with them individually; instead it reviews and evaluates the overall data processing procedures used. Our review of documentation available under the IADS project and our interviewing as many as possible of the people involved raised the following issues and concerns:

A. Questionnaire Processability

The design and complexity of the questionnaires used in the IADS project made it necessary to transcribe the data already recorded by the interviewer onto separate transcription forms. These transcription forms were then used by the data entry operators as the source document instead of the actual questionnaires. Transcription should be avoided as much as possible, not only because it is an additional step that slows down the operation, but much more importantly, because it inevitably introduces new errors in the data.

When designing a questionnaire the most important consideration should be to capture the data as accurately as possible, but the ease with which the questionnaire can be processed should also be adequately considered. Therefore, a close collaboration should be established between questionnaire designers and data processing personnel to ensure the feasibility of accurately processing the questionnaires and avoid costly errors in subsequent stages of the operation.

B. Data Entry

The data entry procedures used by the IADS project consisted of keying the coded data from the transcription forms onto diskettes, with an initial verification factor of 100 percent. Because of delays in the schedule and budgetary considerations this procedure was later changed from 100 percent to a sample of the transcription forms.

Usually 100 percent verification should be used in the initial phase of the data entry operation to correct errors and to identify operators with below average performance. After this initial phase, verification on a sample basis should be sufficient. However, operators must qualify for their work to be verified on a sample basis by maintaining a certain level of performance.

The data entry procedures documented by the IADS project were appropriate to ensure a quality data entry operation. However, we suspect that these procedures were not followed as documented because of time and budget restraints in that stage of the processing.

C. Machine Edit: Cleaning the Data

In our view, the machine editing process to which the data collected under the IADS project were submitted was insufficient to guarantee an adequate cleaning of the data.

After the data entry phase, which recorded the data onto diskettes, the computer processing of the data was accomplished using the software package SPSS. The procedure was to print the variables considered most important and manually examine them for completeness and internal consistency. Corrections were applied to the variables in error and corrected SPSS files were generated for subsequent analysis. We found inadequate documentation on how these editing procedures were carried out and on the procedures used for correcting the data. We suspect that this process occurred under time and budgetary constraints with few quality control procedures and did not allow for an adequate machine editing or "cleaning" of the data, greatly affecting the accuracy of the final data.

It is very important for the interpretation of the data that errors and inconsistencies be correctd before the analysis begins. The main objectives for cleaning the data are: (1) to enhance its quality and (2) to facilitate the data processing task of analysis that follows. In our opinion the machine editing to which the IADS data was submitted was insufficient to meet these objectives.

The machine editing of the data should have included, as a minimum, the following steps:

a. Check-in procedure and questionnaire record formatting. This step is necessary to ensure that all questionnaires that are expected are accounted for.

b. Valid values (range) check. This step looks at individual responses to determine whether or not they are plausible values.

c. Consistency check. This step is performed to detect inconsistencies between two or more responses.

d. Correction of the data after machine editing.

We believe that the selection of the software package SPSS to perform the editing of the IADS data was not a good choice. SPSS is very useful for data analysis and an important tool in most statistical and analytical offices.

However, SPSS was not designed as a software package to correct raw data and its use as such is burdensome and inefficient. The process of cleaning the data should be thoroughly controlled in order to correct errors and not to introduce new errors in the data. During the machine editing stage the data should be cycled through the editing program after corrections are made until no more errors are detected. Shortcuts reflecting directly on the quality of the IADS data appear to have been taken in this stage because of time or other limitations.

D. Present Locations of Data

With respect to the present storage locations of data collected under the IADS project we were able to ascertain the following:

a. The Computer Center at San Carlos University has an IBM 3340 disk labeled EIARØ1 on which the computerized data files for Studies I, II, III, and Activity I are stored. A listing of its directory (VTOC DISPLAY UTILITY) showed the following names of principal datasets:

1) For Study I
   - P001A999 (We assume that this dataset contains P001A102, P103A135, and P136A184. See BANCO de Datos, Informe Final Anexo No. 5)
   - P185A200
   - JERAR1
   - JERAR2
   - JERAR3

2) For Study III
   - EIII1
   - EIII2
   - EIII3

The display of its directory also showed that the programs comprising the SPSS software package are stored on this disk along with other datasets some of which could not be identified but that did not seem to be part of the IADS project data. System and/or operational documentation on how the data were created or archived and how they could be accessed was unavailable in San Carlos. Under the terms of the IADS project this disk belongs to CNPE and can only be used with their approval.

b. CNPE has several magnetic tapes with IADS data given to them by the technical assistance contractor. Unfortunately we found no documentation that would indicate the contents of these tapes or their physical characteristics.

c. The leader of the technical assistance contract team indicated that the contractor has retained a copy of the data for Studies I and II and for Activity I. The data for Study III remained in Guatemala.

E. Documentation of the Data Processing System

One of the essential factors in a project such as this one is an effective and clear system documentation. We looked for, but were unable to find, adequate system and operational documentation.

Documentation of data processing should include as a minimum:

a. A systems flowchart showing how individual programs or procedures interact to make up the system with all inputs and outputs clearly labeled.

b. Complete and precise specifications for each program, describing the inputs and outputs and procedures to be followed.

c. All test runs, indicating the test being performed in each case.

d. Instructions for running each program.

VII. Conclusions and Recommendations

The IADS statistical effort has illuminated some of the advantages as well as problem areas in the development of a systematic integrated planning information system. The basic project purpose -- the development of a statistical base on which integrated rural development plans might be built -- constituted a worthwhile objective; the DS/UD increment complemented and in our opinion made for a more rounded and appropriate approach. However, a distinction should be made between a specific user-oriented information system and a general-use databank. In the former, statistical design needs to start with detailed specification of the end product and work back to the required data collection requirements. In the latter, design generally starts with an on-going data collection/storage/retrieval system and efforts are made to improve the linkage both to the user and to the data source. The IADS project appears to have started with a set of potentially useful variables without a clear methodology for decision-making use and without sufficient attention given to systems analysis (assuring efficient data collection and processing, storage, and ready access) for analytical and planning use. One would expect a major project of this size and complexity to require: (i) strong clear on-site direction and coordination of the project as a whole; (ii) strong on-site statistical support in terms of sample design, survey methodology, and data processing (including computer software and hardware); (iii) strong on-site experienced area development analysis support. Unfortunately these requirements were not met.

INFOM has indicated that they are using the Study I data but not without validation after being derived from the data base. This implies that from experience they have found the data to be of questionable reliability and that validation is required before use. Erroneous data exist in part because in parts of the study area there has been substantial guerrilla and/or army activity, often resulting in destruction of key infrastructure and service elements, such that even data valid at the time of the survey may no longer reflect present conditions. However, the available information on the design and implementation experience also indicates a strong element of potential unreliability at the municipio level; at the specific intervention project level the data leave even more to be desired.

The small sample size used for the Study II survey raises severe questions about the reliability of data on actual land use patterns. The ecosystem zone data are perhaps more valid but their relevance in describing the range of micro-climatic conditions to be found upon spatial disaggregation remains questionable.

The level of fear and violence in much of the study area during the survey operations raise grave questions particularly about the Study III data. Indeed, one cannot expect unbiased data when police escorts are necessary to carry out the interviews.

It is conceivable the data that have been accumulated by means of the primary collection studies are suggestive or indicative of socioeconomic conditions and attitudes in the area of study. Unfortunately the sampling and survey techniques employed violate the canons of accepted practice, and therefore the data cannot be definitively assessed. Nor is it possible to even evaluate their reliability or extent of error. As a result, it is really quite impossible to defend the data on scientific grounds. What this means from a user's standpoint is that the data are of unknown validity, doubtlessly reflecting various selection biases and nonsample errors, though we have no way of judging how serious these biases and errors may be.

Studies involving primary data collection should be designed, supervised, and carried out by survey methodologists and practitioners. Sample design, in particular, should be seen as a great deal more than a hip-pocket exercise that can be executed by a novice in his spare time. Proper sample design is the very foundation upon which inferential studies are based. Casual selection of sample cases that are thought to be "typical" or "representative," but which are not vigorously controlled by probabilistic methods to ensure representativeness, simply does not cut the mustard and cannot be defended on any scientific basis. In short, one cannot describe a population, or otherwise conclude or infer anything about it based on a sample, unless the sample is selected properly, and this means through strict probability methods. This was not the case with the primary data collected under the IADS project.

If the mission chooses nevertheless to use the IADS data it should do so with considerable caution. To facilitate such use we recommend the following:

a. A comparison of the respective datasets stored at San Carlos University CNPE, and Iowa State University is essential to determine which are the most useful for data bank purposes. Information available to us does not indicate whether or not the data at San Carlos, CNPE, or Iowa State were submitted to the same editing processes and transformations. In addition, the results of updating and validating activities by INFOM and other agencies should be incorporated in terms of corrections and revisions in the basic datasets.

b. A well-documented and straightforward users guide should be written to provide the users of the IADS data (Studies I, II, and III, and Activity I) with the information necessary to understand the data and their limitations.

c. The data from Study I could conceivably be used beneficially if the raw data can be weighted properly (see page 10). This would require going back to the initial data file and assigning the appropriate weight, if it can be calculated from available documentation, to each basic record.

d. CNPE is formulating a proposal to upgrade the IBM 4341 computer in the Ministry of Finance to use it as the principal Guatemalan repository of the data collected under the IADS project. The users of this data bank, such as CNPE, INFOM, and the Ministry of Agriculture, would have unrestricted access via remote stations (Cathode Ray Tube display units). Such a system would have a function in the institutionalization of systematic integrated planning information in Guatemala particularly as the existing IADS data are validated and new data are added. In order for this approach to be sound, it would require (1) a systems analyst be hired to organize, administer and develop this data bank and provide technical assistance and training to its users, and (2) a qualified professional, with specialized knowledge of computer hardware, be consulted in the upgrading of the computer and the acquisition of new equipment. Initial tasks of the systems analyst might include implementation of the above recommendations to facilitate IADS data use by the Mission.

ANNEX I

List of Persons Contacted

Ministry of Agriculture

1. Direccion de Recursos Naturales Renovables (DIRENARE)

   - Ing. Ricardo Masaya Andrade
   - Ing. Juan Humberto Mancur
   - Ing. Marco A. Aguilar C.
   - Ing. Alejandro A. Hernandez F.
   - Ing. Carlos de Leon Prera

2. Unidad Sectorial de Planificacion Agricola (USPA)

   - Lic. Ricardo Avila
   - Lic. Manuel Eliseo Herrarte M.

National Planning Council: Consejo Nacional de Planificacion (CNPE)

- Lic. Miriam W. Sieckavizza A.
- Lic. Ruben Flores
- Lic. Axel Rayo
- Lic. Anselmo Alvarez

Municipal Development Institute: Instituto de Fomento Municipal (INFOM)

- Lic. Juan Francisco Leal
- Lic. Eric A. Rosales

USAID/Guatemala

- Peter Kolar, Acting Director
- Cecil McFarland, Rural Development Officer
- George Like, Assistant Rural Development Officer
- Richard D. McLaughlin, Program Officer

USAID/Washington

- Eric Chetwynd, S&T/MS
- Steve Miller, formerly S&T/MS
- Clem Weber, LAC/DR
- Phil Warren, LAC/DR

Iowa State University, Dept. of Community & Regional Planning

- Jerry Knox
- Roberto Morales
- Paul Anderson

Clark University Dept. of Economic Geography

- Jerry Karaska

List of Documents Reviewed

Study I --

Inventario de Datos de Equipamiento, Infraestructura y Servicios: Informe Final, Guatemala: INFOM, Octubre 1981

Anexo 1, Listado de Lugares Investigados por Departamento
Anexo 2, Boletas Originales
Anexo 3, Boleta Final
Anexo 4.1 Documentos de Encuesta: Guía del Encuestador (1a y 2a partes)
Anexo 4.2 Documentos Auxiliares de Encuesta
Anexo 5, Documento de Adiestramiento-
Cursillo de Adiestramiento para Encuestadores:
- Programa
- Programación...que Se Realizara para Evaluar y Seleccionar Personal de Campo para el Estudio I
- Instructivo de Programación y Descripción de las Actividades de Reclutamiento, Selección y Contratacción del Personal de Campo Permanente
- Instrucciónes Generales para las Actividades sobre Llenado, Revisión y Corrección de Boletas
- Técnica de la Entrevista
- Instructivo General para Recolección, Verificación y Procesamiento de Datos
- Uso y Manejo de Material Cartográfico
- Relaciones Humanas en el Trabajo de Campo
- Examen Parcial de Evaluación
- Examen Final de Evaluación

Anexo 6, Manual Crítica y Codificación
Anexo 7, Guía para Visitas a Instituciónes
Anexo 8, Listado de Códigos
Anexo 9, Información Contenida en Inventario, segun Numero de Variables

Study II --

Sub-Programma II, Estudios Integrados de Áreas Rurales, 8 volumes, Guatemala: Ministerio de Agricultura

Clima y Vegetación, Septiembre 1981
Clasificación Agrológica (Recurso Suelo), Septiembre 1981
Unidades de Suelos y Potencial de Suelos, undated
Uso Actual de la Tierra, Agosto 1981
Producción Agrícola, Septiembre 1981
Producción Agrícola: Anexo, Septiembre 1981
Recursos Hidráulicos, Septiembre 1981
Resumen, Octubre 1981

Sistemas de Recolección de Datos en Agricultura, (Manuel E. Herarte M.), Guatemala: Ministerio de Agricultura, 1981

Primera Encuesta Agrícola de la Región I, Guatemala: Ministerio de Agricultura, Julio 1981

Segunda Encuesta Agrícola de la Región I: Año 1980, Guatemala: Ministerio de Agricultura, Abril 1982

Study III--

Patrones de Acceso (Estudio III): Informe Final, Guatemala: INFOM, Diciembre 1981

Anexo 1, Listado de Lugares Poblados Investigados en la Muestra del Estudio III por Región
Anexo 2, Mapas de Frecuencia de Viajes
Anexo 3, Distribución de los Viajes por Propósito y Medio Segun Departamento y Nivel de Análisis de Acuerdo al Lugar de Origen
Anexo 4, Contiene Boleta de Investigación, Códigos Utilizados y Documentos de Apoyo en Crítica y Codificación

Activity I --

Participación Local en la Aldea "Sonadora", Guatemala: INFOM, Octubre 1981

An Opportunity for Development: A Study of Local Participation in Highland Guatemala, (K.L. Carey and S.E. Carey), Ames: Iowa State University, 1981

Activity II --

Banco de Datos: Informe Final, Guatemala: INFOM, Octubre 1981

Organización del Espacio en el Franja Central de la República de Guatemala, Guatemala: INFOM, Octubre 1981

Propuesta Técnica para la Implementación del "Sistema Nacional de Información para el Desarrollo (SINADES)", Guatemala: CNPE, Octubre and Diciembre 1981

Fase de Integración, Estudios Integrados de las Áreas Rurales, Guatemala: CNPE, Abril 1982

Informe Final
Resumen
Anexo A, Enlaces
Anexo B, Uso Actual de la Tierra, Cultivos por Municipio y Departamento
Anexo C, Uso Potencial: Grupos de Cultivos por Subsistema
Anexo D, Producción Agrícola Sistema Occidental
Anexo E, untitled (selected data from the 1980 Encuesta Anual de la Industria Manufacturera)
Anexo F, Uso Potencial por Grupos de Cultivos, Uso Actual Segun Variables y Uso Actual Segun Cultivos por Subsistema
Anexo G, untitled (selected data from Estudio I, Inventario de Infraestructura y Servicios (1980))
Apendice, Códigos del Lugar Poblado

A Rural Development Planning Program for Guatemala, (Jerry Knox et al), Ames: Iowa State University, unpublished draft dated November 17, 1982

A Rural Development Planning Program for Guatemala, Tecnical Report: Data Processing Procedures, (Jerry Knox et al), Ames: Iowa State University, unpublished draft dated November 13, 1982

24

Modelo de Planeación para la Identificación y Localización de Proyectos de Inversion, Guatemala: INFOM, Octubre 1982

Propuesta para Creación e Implementación del "Centro de Información y Análisis para el Desarrollo Municipal del Instituto de Fomento Municipal (CIADEM)", Guatemala: INFOM, undated draft.

Related AID Memoranda and Reports --

Land and Labor in Guatemala: An Assessment, (Richard Hough et al), Guatemala: USAID/Guatemala, September 1982

Rural Demand for Urban Service Systems (IADS-Guatemala) Project, Eric Chetwynd/Files memo dated February 20, 1981

Report to USAID/Guatemala: Integrated Area Development Study, (G. Karaska and J. Lombardo), Worcester (Mass.): Clark University COAD, July 9, 1982

Evaluation of Agricultural Sector Planning Activities in Latin America and the Caribbean, (J.T. Riordan), Cambridge (Mass.): Abt Assoc., June 1982

Memorandum of Agreement between AID/Guatemala and DS/UD/AID/Washington on Integrated Area Development Studies (IADS) and Rural Demand for Urban Service Systems (RDUSS) projects, undated draft.

Participatory Aspects of the Current Methodology Development Project, F.O'Regan/D. Peacock memo dated January 30, 1981

Report on Trip to Guatemala; June 20-25, 1978, M. Conroy/J. Dickey memo with cover letter dated July 17, 1978

Trip to Iowa State University to Review Integrated Area Development/Urban Functions in Rural Development Project in Guatemala, Steve Miller/Files memo dated December 14, 1982

Integrated Area Development Studies (Guatemala), Project Paper, Project No. 520-0249, Washington: AID, May 1978.

(Untitled) Evaluation Report to USAID/Guatemala: Integrated Area Development Studies Project, (T. McKee and J.T. Riordan), Guatemala: USAID/Guatemala, undated draft.

List of Additional References

Cochran, William G., Sampling Techniques, Second Edition, New Delhi: Wiley Eastern Limited, 1963

Hansen, Morris H., William N. Hurwitz, and William B. Madow, Sampling Survey Methods and Theory, Volume I, Methods and Applications, New York: John Wiley and Sons, 1953

Jessen, Raymond J., Statistical Survey Techniques, New York: John Wiley and Sons, 1978

Kish, Leslie, Survey Sampling, New York: John Wiley and Sons, Inc., 1965

Williams, Bill, A Sampler on Sampling, New York: John Wiley and Sons, Inc., 1978