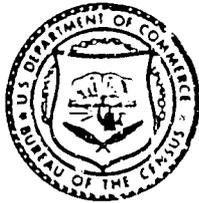


HAITI HOUSEHOLD EXPENDITURE AND CONSUMPTION SURVEY

NOVEMBER 1986 - SEPTEMBER 1987

PROCEDURAL HISTORY



U.S. Department of Commerce
 BUREAU OF THE CENSUS
 International Statistical Programs Center
 Washington, D.C.



REPUBLIQUE D'HAÏTI
 Ministère de l'Économie et des Finances
 INSTITUT HAÏTIEN DE STATISTIQUE ET D'INFORMATIQUE

HAITI HOUSEHOLD EXPENDITURE AND CONSUMPTION SURVEY

November 1986 - September 1987

PROCEDURAL HISTORY

**International Statistical Programs Center
U. S. Bureau of the Census**

PREFACE

The Household Expenditure and Consumption Survey (HECS), carried out November 1986-September 1987, represents a major accomplishment in the history of national sample surveys in Haiti. It is the first national survey of this scope to use a probability sample design. In the development of operational procedures and of survey documents, sound survey and sampling methodologies, consistent with recognized statistical standards were emphasized. It is therefore important to record the survey experiences.

This report presents the procedural history and methodological documentation of the HECS. Its primary purpose is to make the survey methods used available to a wide audience.

A procedural history is useful for the understanding of the survey design and procedures and, of the strengths and weaknesses of the final product. It also provides a good guide for the design of future surveys since it documents both accomplishments and difficulties and gives the rationale underlying the decisions made. In addition, the experiences described can be useful to others who want to conduct similar surveys.

It is important to note that this documentation does not cover all the details of the survey design and implementation; it provides only a summary. Numerous other survey documents have been prepared such as the Interviewer's Manual, the Field Supervisor's Manual, the Office Operations Manual, among others. Appropriate references are made throughout the text, particularly in Chapter III. These documents provide greater detail on the different survey activities.

This report has been prepared for the U. S. Agency for International Development by the U. S. Census Bureau. It represents the combined efforts and inputs of the many people and organizations whose collaboration made the project possible.

ACKNOWLEDGEMENTS

The Household Expenditure and Consumption Survey (HECS) was carried out by the Institut Haitien de Statistique et d'Informatique (IHSI) with the support of the U.S. Agency for International Development (USAID). The U.S. Census Bureau, the U. S. Department of Agriculture (USDA), and the Center for Agricultural and Rural Development (CARD) provided technical assistance. There were many persons who contributed to make the survey possible.

USAID/Port-au-Prince

The USAID/Port-au-Prince Mission was instrumental in the conceptualization and development of the project. James Walker, Chief of the Office of Economic Analysis, and Giovanni Caprio arranged funding and technical assistance, provided technical input at the developmental stage, and gave continued managerial support throughout the survey. This support as well as support for promoting Haiti's overall statistical system were continued through Edward Clarke, who replaced Dr. Walker.

IHSI

The IHSI was responsible for the actual implementation of the survey. Matthieu Duplan and his staff were the IHSI counterparts who, together with USAID, conceptualized and also arranged funding for the project. He was followed by Jacques Vilgrain and Louis Smith. Mr. Smith was with the project the longest and his commitment was one of the moving forces behind the survey's success. Survey implementation was possible thanks to the competent services provided by the HECS team composed of Mesdames Fernande Pierre-Louis and Danilia Altidor and, Messrs. Raymond Gardiner, Burel Decopain, and Odnel Eleazard, who were responsible for the different technical and administrative aspects of the survey including survey design, pilot survey implementation and analysis, design of data collection and data processing procedures, etc. Data processing services were provided by Raoul Nelson, Frantz St. Fleury St. Fleur and staff.

Invaluable clerical support, including the typing of the questionnaire, manuals and forms, was given by Livia Pierre-Lys and others. The contributions of the field and office staff, notably those of Immaculee Noel and Cyrano Morinvil who directed office operations, were of critical importance to the success of the HECS. The participants are too numerous to mention, except collectively: the cartographers, interviewers, office clerks, data entry operators and supervisors all provided reliable data through their diligent efforts in the mapping/listing, data collection and data processing activities.

Census Bureau/ISPC

The International Statistical Programs Center of the U. S. Census Bureau provided technical services in the areas of survey design, sampling, statistical methods, and survey management. Marjorie Dauphin was the resident adviser in Haiti for the length of the survey and directed all phases while providing technical assistance in all aspects of the survey. Sandra Rowland assisted in the design of the survey and the direction of short-term technical assistance. David Megill designed the sample, specifying sample selection and estimation procedures. Miguel Cuevas and Kate Fleischer worked on the design of the questionnaire, the interviewer's manual, the supervisor's manual, the office operations manual as well as several control forms. The Census Bureau team was also substantially involved in the design, implementation and evaluation of the HECS pilot survey.

USDA/NEG

The Nutrition Economics Group (NEG) of the U. S. Department of Agriculture (USDA) carried overall responsibility for the data processing and analysis for AID. Shirley Pryor arranged for funding and provided technical and managerial cooperation throughout the development and implementation of the HECS.

CARD

The Center for Agricultural and Rural Development (CARD) of Iowa State University, under contract with USDA/NEG, was responsible for processing and analyzing the HECS survey results for AID. Under the direction of Stanley Johnson, CARD personnel, notably Helen Jensen, contributed significantly to the design of the questionnaire, the analysis of pilot survey results, the design of data collection procedures and the computer processing of the HECS results, particularly the computer edit.

CONTENTS

| | <u>Page</u> |
|---|-------------|
| Preface..... | i |
| Acknowledgements..... | ii |
| CHAPTER I. OVERVIEW OF SURVEY | |
| Survey Objectives..... | 1 |
| Overview of Scope and Design | 1 |
| Operational Phases..... | 2 |
| HECS Activities Not Completed..... | 6 |
| CHAPTER II. QUESTIONNAIRE DESIGN | |
| Content..... | 7 |
| Initial Design Considerations..... | 7 |
| Pretest and Pilot Survey..... | 8 |
| Final Organization of Questionnaire..... | 10 |
| Question and Response Type..... | 10 |
| Use of Source Codes..... | 11 |
| CHAPTER III. OPERATIONAL DOCUMENTATION | |
| Reference Manuals..... | 12 |
| Instruction Manual for Mapping and Listing..... | 12 |
| Instruction Manual for the Supervision and Validation of Mapping and Listing Operations and for the Selection of Housing Units..... | 12 |
| Interviewer's Manual..... | 13 |
| Field Supervisor's Manual..... | 14 |
| Office Operations Manual..... | 14 |
| Data Verification and Correction Manual..... | 15 |
| Difficulties with Reference Manuals..... | 15 |
| Survey Forms..... | 15 |
| Data Collection and Sampling Forms..... | 15 |
| Control Forms..... | 16 |
| Administrative Forms..... | 16 |
| Data Processing Control Forms..... | 16 |

CHAPTER IV. FIELD OPERATIONS

| | |
|--|----|
| Mapping and Listing..... | 17 |
| Pilot Survey..... | 17 |
| Full-Scale Operation..... | 17 |
| Difficulties with Mapping/Listing Operation..... | 18 |
| Data Collection..... | 19 |
| Calendar..... | 19 |
| Pilot Survey..... | 19 |
| Field Organization of Full-Scale Survey..... | 21 |
| Training of Field Staff..... | 22 |
| Preparation of Field Materials..... | 23 |
| Interviewer Departure and Arrival..... | 23 |
| General Field Procedures..... | 24 |
| Household Interviews and Market Visits..... | 25 |
| Field Quality Control..... | 26 |

CHAPTER V. DATA PROCESSING OPERATIONS

| | |
|--|----|
| Office Operations..... | 27 |
| Office Calendar and Organization | 27 |
| Pre-Machine Edit, Coding and Conversion..... | 29 |
| Computer Processing..... | 29 |
| Data Entry and Creation of Data Files..... | 29 |
| Machine Edit and Imputation..... | 30 |
| Tabulations..... | 30 |

CHAPTER VI. SAMPLE DESIGN AND ESTIMATION PROCEDURES

| | |
|--------------------------------------|----|
| Sampling Plan..... | 31 |
| Universe and Unit of Analysis..... | 31 |
| Sampling Frames..... | 31 |
| Stratification..... | 32 |
| Sample Size and Allocation..... | 32 |
| Sample Selection Procedures..... | 34 |
| Estimation Procedures..... | 36 |
| Basic Weighting Procedures..... | 39 |
| Weight Adjustment Factors | 41 |
| Estimation Formulas..... | 44 |
| Other Estimation Considerations..... | 45 |
| Calculation of Variances..... | 45 |

CHAPTER VII. EFFECT OF POLITICAL SITUATION ON THE SURVEY

| | |
|---|----|
| Interaction With Survey Activities..... | 47 |
| Effect on Data Obtained..... | 48 |

CHAPTER VIII. RECOMMENDATIONS FOR FUTURE SURVEYS

| | |
|-----------------------------------|-----------|
| Recurrence of Surveys..... | 49 |
| Survey Management..... | 49 |
| Questionnaire Design..... | 50 |
| Data Processing..... | 50 |
| Sample Design..... | 50 |

LIST OF TABLES AND FIGURES

| <u>Tables</u> | <u>Page</u> |
|---|-------------|
| Table 1. Coverage Rate of Target Enumeration Areas..... | 4 |
| Table 2. Household Response Rates by Month..... | 5 |
| Table 3. HECS Sample Allocation | 33 |

| <u>Figures</u> | |
|---|----|
| Figure 1. Calendar of Key Survey Dates..... | 3 |
| Figure 2. Field Operations Calendar (EBCM-A8)..... | 20 |
| Figure 3. Office Operations Calendar (EBCM-A9)..... | 28 |
| Figure 4. Illustration of First Stage Sample Selection..... | 35 |
| Figure 5. Worksheet for the Selection of Housing Units (EBCM-3)..... | 37 |
| Figure 6. Worksheet for the Assignment of the Visiting Week to the Household (EBCM-4)..... | 38 |
| Figure 7. Worksheet for the Calculation and Transmission of Weights (EBCM-5)..... | 40 |

CHAPTER I

OVERVIEW OF SURVEY

To acquaint the reader with the survey, the present chapter gives an overview of the survey objectives, scope, design and operational phases. In the following chapters, the survey design and operations are discussed in detail.

Survey Objectives

The objectives and scope of the HECS were established as a result of discussions among personnel of the IHSI and USAID/Port-au-Prince. The survey had multiple objectives. The main objectives of the IHSI were to update the national accounts and the base weights for the consumer price index and to obtain a data base for national and regional planning. They also needed data to permit the estimation of consumption out of household production. USAID needed data to evaluate the impact of Title III economic assistance programs on different socio-economic groups. In addition, baseline information would be obtained on food consumption, nutritional status and health as well as on a diversity of other socio-economic variables on Haitian households. This would constitute a useful by-product for data users in general as there was little current information on the subject.

Overview of Scope and Design

The HECS questionnaire covered the following subjects:

- housing characteristics and expenditures;
- general and economic characteristics of household members;
- inventory of food on hand (beginning and end of week);
- food expenditures (collected 4 times a week);
- expenditures for non-food items;
- consumption out of the household;
- agricultural production and expenses; and
- health characteristics.

Nine separate estimation domains were formed in addition to the national level: the four planning regions (North, Transversale, Ouest, Sud) divided into urban and rural zones of residence and the Port-au-Prince Metropolitan Area as a separate stratum. At these levels, the precision in terms of the coefficient of variation can be expected to be within 15% for predominant characteristics. Estimates for characteristics for which the number of observations is inadequate can only be produced at the national urban/rural level because, at lower levels, they would suffer from high sampling errors.

The HECS used a probability sample design, the only method which permits unbiased and reliable estimation of the population parameters from the sample. The sample design allowed the sample to be representative in both time and space. Data collection was designed to take place over the course of 52 weeks, thus taking into account seasonal variability, at the rate of 240 households every 4 weeks (for a total sample size of 3,120), scientifically chosen to represent all households in Haiti. The sample design is discussed in greater detail in Chapter VI.

Operational Phases

1. Preparatory Phase

In a collaborative effort among USAID/Port-au-Prince, the U. S. Census Bureau, USDA's Nutrition Economics Group, the Center for Agricultural and Rural Development (CARD) and the IHSI, a questionnaire and a data collection methodology were designed in September 1985 for the implementation of the HECS. To assess the feasibility, validity, and efficiency of the methodologies developed, a nationwide pilot survey of 90 households was held in May 1986. Such aspects as field operations, sample design, questionnaire, listing and interviewer's manuals, interviewer training, and others were evaluated. Based on the pilot results, recommendations were made for the final survey design and major revisions went under way. The final preparations for the full-scale survey fieldwork took place in record time, between July and October 1986. A calendar of key dates is furnished in Figure 1.

2. Data Collection Phase

Mapping and listing operations started in July 1986 and lasted through August 1987. These operations preceded the actual data collection by 1 or 2 months and laid the ground for the selection of sample households.

The interviews for the survey began November 4, 1986. Every month 24 interviewers worked in one enumeration area each, visiting 10 households over a 4-week period. Though scheduled to end on November 1, 1987, data collection closed as of September 20, 1987, after the 11th survey month, because of security concerns related to the country's political situation. Hence, the HECS data collection covered 11 consecutive periods of 4 weeks, except for minor discontinuities for short periods of up to 2 weeks total.

The total surveyed sample consists of 260 enumeration areas and 2,593 households. This coverage was equivalent to 83.3% of the original target sample of 312 enumeration areas for 13 months. Of the covered enumeration areas, 11.2% were substitutes. Overall, the household response rate was excellent: 93.8%, with a substitution rate of 10.7%. The data collection results are summarized in Tables 1 and 2.

Figure 1. CALENDAR OF KEY SURVEY DATES

| Activity | Beginning Date | Ending Date |
|--|----------------|-------------|
| INITIAL DESIGN | | |
| Initial Sample Design Recommendations | Apr-85 | Apr-85 |
| Preparation of First-Stage Sampling Frame and Selection of First-Stage Sample | May-85 | Jun-85 |
| Initial Questionnaire and Methodology Design | Sep-85 | Sep-85 |
| PILOT SURVEY PREPARATIONS | | |
| Questionnaire Pretest and Revision | Oct-85 | Nov-85 |
| Design of Data Collection Operations | Nov-85 | Dec-85 |
| Preparation of Reference Manuals for Pilot Survey | Nov-85 | Jan-86 |
| PILOT SURVEY | | |
| Mapping and Listing | Dec-85 | Dec-85 |
| Selection of Sample Housing Units | Dec-85 | Dec-85 |
| Recruitment and Training of Interviewers | Apr-86 | Apr-86 |
| Data Collection | May-86 | May-86 |
| Processing of Results | Jun-86 | Jun-86 |
| Analysis of Results and Recommendations for Full-Scale Survey | Jul-86 | Jul-86 |
| PREPARATIONS FOR FULL-SCALE SURVEY | | |
| Revision of Questionnaire | Jul-86 | Sep-86 |
| Revision of Reference Manuals and Other Survey Forms | Jul-86 | Oct-86 |
| Recruitment and Training of Interviewers | Sep-86 | Oct-86 |
| Training of Coders/Editors | Dec-86 | Dec-86 |
| Data Processing Systems Design | Sep-86 | Dec-86 |
| Specification of Estimation Procedures | Jul-86 | Jul-86 |
| FULL-SCALE SURVEY (repeated monthly) | | |
| Mapping and Listing | Jul-86 | Aug-87 |
| Selection of Sample Housing Units | Nov-86 | Sep-87 |
| Data Collection | Nov-86 | Sep-87 |
| Pre-Machine Edit and Coding | Dec-86 | Nov-87 |
| Data Entry | Jan-87 | Dec-87 |
| Machine Edit | Sep-87 | - |

Table 1. HECS COVERAGE RATE OF TARGET ENUMERATION AREAS (SDEs)

| Month | Number of Target SDEs | Number of Covered SDEs | Number of Substit. in Covered SDEs | Coverage Rate of SDEs | SDE Substitution Rate |
|--------------|-----------------------|------------------------|------------------------------------|-----------------------|-----------------------|
| 01 | 24 | 24 | 3 | 100.0% | 12.5% |
| 02 | 24 | 24 | 2 | 100.0% | 8.3% |
| 03 | 24 | 24 | 4 | 100.0% | 16.7% |
| 04 | 24 | 24 | 4 | 100.0% | 16.7% |
| 05 | 24 | 24 | 3 | 100.0% | 12.5% |
| 06 | 24 | 24 | 1 | 100.0% | 4.2% |
| 07 | 24 | 24 | 1 | 100.0% | 4.2% |
| 08 | 24 | 24 | 1 | 100.0% | 4.2% |
| 09 | 24 | 24 | 1 | 100.0% | 4.2% |
| 10 | 24 | 23 | 3 | 95.8% | 13.0% |
| 11 | 24 | 21 | 6 | 87.5% | 28.6% |
| 12 | 24 | 0 | 0 | 0.0% | - |
| 13 | 24 | 0 | 0 | 0.0% | - |
| 1 THROUGH 11 | 264 | 260 | 29 | 98.5% | 11.2% |
| 1 THROUGH 13 | 312 | 260 | - | 83.3% | |

Table 2. HECS HOUSEHOLD RESPONSE RATES BY MONTH

N. B. The number of target interviews was 240 for each month.

| Month | Number of Eligible HUs in Attempted Sample | | Number of Completed Interviews | Number of Substitutions in Completed Interviews | Response Rate Including Subst. | Substitution Rate |
|--------------|--|---------|--------------------------------|---|--------------------------------|-------------------|
| | Original | Reserve | | | | |
| 01 | 231 | 30 | 238 | 26 | 91.2% | 10.9% |
| 02 | 231 | 21 | 240 | 20 | 95.2% | 8.3% |
| 03 | 232 | 29 | 237 | 28 | 90.8% | 11.8% |
| 04 | 229 | 22 | 240 | 23 | 95.6% | 9.6% |
| 05 | 230 | 14 | 240 | 17 | 98.4% | 7.1% |
| 06 | 229 | 27 | 240 | 29 | 93.8% | 12.1% |
| 07 | 232 | 20 | 240 | 20 | 95.2% | 8.3% |
| 08 | 230 | 20 | 240 | 20 | 96.0% | 8.3% |
| 09 | 229 | 26 | 238 | 25 | 93.3% | 10.5% |
| 10 | 218 | 36 | 230 | 39 | 90.6% | 17.0% |
| 11 | 200 | 29 | 210 | 31 | 91.7% | 14.8% |
| 1 THROUGH 11 | 2,491 | 274 | 2,593 | 278 | 93.8% | 10.7% |

3. Data Processing Phase

The completed questionnaires were brought in to the office each month where they were hand-edited, coded and entered on computer files for subsequent processing and analysis.

Work began on the machine edit for these data files, that is, identification and correction of illegal, out-of-range or inconsistent entries; but, due to the suspension of financial and technical assistance for the HECS, this activity has not been completed. See Chapter VII for a discussion of the political situation.

HECS Activities Not Completed

At the time the survey activities stopped in December 1987, all 11 months of collected data had been entered on computer files - although minor data entry verification problems still had to be corrected. Therefore, to produce the HECS estimates, several operations must take place. These operations can be summarized in the following three phases: the computer edit and imputation phase, the estimation phase and the publication phase.

These phases involve in turn several activities. The computer edit and imputation phase (review and correction of data files), which was only in its beginning stage when the program cuts occurred, involves: finalization of edit computer programs; manual review of edit programs' outputs and of questionnaires to reconcile errors listed; specification of corrections to the data file based on this review; correction of data file and identification of nonresolvable errors; and, imputation of uncorrected errors.

Procedures for the estimation of population parameters and their variances have already been specified. Initial estimation weights have been calculated for all months. Before the actual calculation of estimates can take place, however, several activities must be completed. These include adjusting weights for months 7 through 11 for changes in occupancy status, the finalization of the tabulation plan (variables and levels of estimation, table outlines), and the formulation of computational specifications for the tabulation program. The computer processing includes the preparation and execution of the tabulation program and the variance program. The outputs of the estimation phase are the basis for the data analyses.

The publication phase involves: the preparation of tables of estimates, the preparation of a scope and reliability statement, typing and editing the publication documents, and finally the printing itself.

CHAPTER 11

QUESTIONNAIRE DESIGN

Content

The content of a questionnaire is generally determined by the objectives of the survey. As indicated in Chapter I, the Household Expenditure and Consumption Survey (HECS) had multifold objectives which explains the sizable length of the questionnaire - 220 pages with a total of 160 questions grouped into 14 separate sections. It may be recalled that the Institut Haitien de Statistique et d'Informatique (IHSI) had among its main objectives to update the national accounts and the base weights for the consumer price index while USAID/Port-au-Prince's primary interest in the survey was to obtain baseline data on the food consumption, nutrition status and health of Haitian households. The data were to be used in a series of background and policy studies to be conducted by the Center for Agricultural and Rural Development (CARD) at Iowa State University. These studies were, in turn, to be used by the Government of Haiti as a basis for evaluating the changes in policy contemplated or proposed in the 1987 Title III Agreement.

Taking into account the survey sponsors' various information needs, the final questionnaire was designed to contain the following 14 sections:

- I. Housing Characteristics and Related Expenses
- II. General Characteristics of Household Members
- III. Economic Characteristics of Household Members
- IV. Inventory of Food Items and Beverages
- V. Food and Other Daily Expenditures
- VI. Food Consumed Outside the Household
- VII. Expenditures for Services and Non-Food Items
- VIII. Payments for Goods and Services Purchased on Credit
- IX. Income
- X. Health
- XI. Agricultural Production
- XII. Cost of Agricultural Inputs
- XIII. Livestock Production
- XIV. Home Production and Sale of Selected Products

Initial Design Considerations

The essence of a successful survey is having a questionnaire that has been planned, carefully designed, and thoroughly tested. The HECS questionnaire had all of the ingredients of an effective questionnaire, since it adhered to the above principles. The first draft of the HECS questionnaire was developed in group sessions at the Census Bureau in Washington, in September 1985, during a 3-week work-visit by IHSI. Other institutions participating in the sessions were CARD, USDA and USAID/Port-au-Prince.

Prior to their Washington visit, IHSI staff developed a preliminary tabulation plan which was used, along with the information provided by the analysts from CARD, as the basis for the design of the questionnaire. This first questionnaire had 17 sections.

One of the most important considerations in the questionnaire design was the accuracy of the data to be collected. Given the food-buying habits of low-income groups where food is purchased and consumed the same day and respondents do not recall their purchases, it was determined that a minimum of four visits during the week were required to minimize recall problems and hence obtain more accurate estimates of weekly consumption. Consequently, an extensive list of precoded food and food-related item prompts was included for each of the four visits. Although these lists made the questionnaire longer, they served to remind respondents of most of their purchases.

Another important consideration was the languages in which to produce the questionnaire. Although most of the Haitian population speaks Creole, especially in the rural areas, few know how to read or write Creole since Creole literacy is only a recent phenomenon. Moreover, the language lacks complexity and offers vocabulary limitations for the formulation of technical concepts. From a technical point of view, it was thus more practical to write the original questionnaire in French. What needed to be considered, however, was whether a second version of the questionnaire, translated into Creole, was required. Because of the additional time and cost this would require and because it was easier to find interviewers who could read French and orally translate the questions into Creole, the decision was made to maintain only the French version and have the interviewers translate into Creole in non-French speaking households.

Pretest and Pilot Survey

To test the questionnaire designed, a "mini-test" of four interviews was conducted in October, 1985. Due to time constraints, the four visits planned were combined into two. The result was a 2 to 3 hours of interviewing per visit which was not indicative of the time it would have taken if the original schedule had been followed. The mini-test, however, revealed the need to conduct a separate inquiry into weight measurements and their relationship to price since consumers, in low-income areas, were not conscious of what they purchased in terms of weight. The mini-test also led to a shortening of the questionnaire. The number of sections was reduced from 17 to 14, and the number of questions from 192 to 158.

The HECS questionnaire was again tested in May 1986 in a nationwide pilot survey. It was conducted over the course of 4 weeks in nine sections d'énumération (SDEs), or enumeration areas, by nine interviewers. A total of 90 interviews were conducted. The pilot survey was a success since it pinpointed some of the shortcomings of the questionnaire and provided a basis for revising it.

Two reports were prepared detailing the findings of the pilot survey. The first, "Field Observation Report", summarized the major fieldwork problems while "Questionnaire Evaluation Report" focused on particular problems found in the completed questionnaires. The following aspects were evaluated for each of the questionnaire sections: the influence of the questionnaire format on the data recorded, the validity of certain answers, the type and frequency of errors, and the consistency among different variables. To assist in the evaluation, computer summaries of household data were prepared on the software package LOTUS 1-2-3.

Among the major findings were:

- the format of the questionnaire was confusing to the interviewers as the sections were not arranged by visits thus necessitating their flipping through the questionnaire in search of the questions corresponding to a particular visit;
- some of the food items listed in the questionnaire were considered a luxury and were rarely or never consumed in poorer households; mention of those items made the respondent uncomfortable;
- the duplication of food items reported in the sections concerned with food received as gifts and/or produced at home with that of food purchased;
- the need to include a section to account for expenses and the number of meals that household members consume outside the household, as well as meals consumed inside the household by non-household members; and,
- the Creole translation of the questions from the French during the interview was more inconsistent than could be tolerated. There was great variation in the translations not only among interviewers but also among interviews conducted by the same interviewer.

Of the questionnaire's 14 sections, those that presented the most inconsistencies were the ones on agricultural production and cost of agricultural inputs. There was great reluctance on the part of the respondents to answer detailed questions about their farm holdings. Respondents also had great difficulty in estimating the quantities produced of the different products, as well as remembering when they had last harvested a given product. In addition, with the HECS methodology being oriented to the collection of socio-economic household data rather than agricultural (farm) data, there were doubts as to the reliability and validity that could be expected from the agricultural data in the HECS.

To address these and other difficulties identified during the pilot survey, a series of meetings was held at IHSI from June 30 to July 4, 1986. Present at these meetings were the representatives from IHSI, CARD, USDA, USAID/Port-au-Prince and the Census Bureau. The review meetings produced a report titled, "Recommendations for the Design of the Full-Scale Survey" which, together with the "Questionnaire Evaluation Report" and the "Field Observation Report" constituted the guiding documents for

the improvement of the questionnaire. Virtually all of the recommendations outlined were incorporated into the revised version of the questionnaire.

Final Organization of the Questionnaire

Since the organization of the pilot survey questionnaire by topic made it difficult for interviewers to use, the questionnaire was rearranged to facilitate its handling. Instead of having one questionnaire encompassing all four visits, the questionnaire was divided into four separate modules. Each module was complete with its corresponding sections and questions according to the visit number and was physically separated from the others. To prevent the intermixing of modules from one questionnaire to another, the household identification data were marked on each module before the interviewers' departure to the field.

The sections assigned to each visit were:

First visit: Sections I - VI

Second visit: Sections V - VIII

Third visit: Sections V, VI, and IX

Fourth visit: Sections IV - VI and X - XIV

All questions about housing characteristics and identification information of household members were collected during the first visit since they are the easiest to answer and the least likely to cause anxiety for the respondent. On the other hand, the sections relating to agricultural production and costs were left for the last visit.

Question and Response Type

One of the principal features of the HECS' questionnaire was that all inquiries were written as complete questions. The verbatim questions took more space to print on the questionnaire but helped interviewers grasp the intent of the questions and served to lessen inconsistencies from interviewer to another. In French-speaking households, verbatim questions allowed interviewers to proceed smoothly with the interview and relieved them of the burden of having to form questions in the course of the interview. For non-French-speaking households, a uniform Creole translation of the questions, prepared by the technical staff, was provided on a separate sheet. This Creole translation helped standardize the way the interviewers asked the questions without having a second questionnaire totally in Creole.

Most of the questions in the HECS' questionnaire elicited numerical responses since they dealt with values and quantities. Wherever feasible, questions were also designed to channel responses to a relatively small

number of answer choices precoded on the form. This type of question which produces a brief structured response is usually referred to as a "closed question". Closed questions allow the interviewer to save time recording answers and, since they are usually precoded, they permit entering answers directly into the computer without a separate coding operation. Finally, a few questions were of the type known as "open question" since they produced an unstructured type of response that is recorded on the questionnaire in the respondent's own words. These were subject to office coding prior to data entry.

Use of Source Codes

The HECS questionnaire design was planned for efficient data processing by using source codes which eliminated keypunching cells of the questionnaire that were not filled. Source coding is a data entry scheme in which each cell of information in the questionnaire is identified by its unique code. With this system only the cells with actual responses (including the source codes) have to be keyed. It was chosen for use in this questionnaire, because the number of expected nonzero responses in any given questionnaire was relatively small in comparison to the total number cells.

The 20,303 source codes that were allocated to the questionnaire were not assigned to questions continuously. For example, Section I of the questionnaire consisted of source codes 00100 through 00124, while Section II began with source code 00200 - a gap of 76 source codes. Ideally, the total number of source codes should not be so high and the number of digits should be no more than three. However, given the large number of questions on different topics in the HECS questionnaire, it was difficult to proceed differently.

CHAPTER III

OPERATIONAL DOCUMENTATION

A. REFERENCE MANUALS

In order to carry out the survey procedures, six operational manuals were written. They are:

1. the Instruction Manual for Mapping and Listing
2. the Instruction Manual for the Supervision of Mapping and Listing and for the Selection of Housing Units
3. the Interviewer's Manual
4. the Field Supervisor's Manual
5. the Office Operations Manual
6. the Data Verification and Correction Manual

These manuals served both as training guides and as reference documents by survey personnel throughout the survey operations. They were invaluable in ensuring clarity and uniformity in the interpretation of procedures and they serve as primary references for survey documentation. A description of each manual follows.

Instruction Manual for Mapping and Listing

This manual explains the objectives of the mapping and listing operation and their relation with the sample design and data collection. It covers basic canvassing procedures and provides detailed instructions on how to complete the listing record. It was used to train the listing enumerators and the listing supervisor. The manual is divided into seven chapters which cover the following: the objectives of the operation, the role of the listing enumerator, operational concepts and definitions, materials and documents needed by the enumerator, how to identify and canvass the enumeration area, how to update the maps, detailed instructions for filling out the listing record, instructions for reviewing and transmitting documents, and finally an example of the listing record itself.

Instruction Manual for the Supervision and Validation of Mapping and Listing Operations and for the Selection of Housing Units

These instructions concern the second-stage sampling frame directly: (1) they offer guidelines for the validation of this frame, which is the result of the mapping and listing operations, and (2) they provide step-by-step instructions for the selection of housing units (HUs). Even though these instructions were addressed to a small number of people, they were needed because of the several steps involved and the importance of adhering to the survey design. Without written references, it would have

been easy for the listing supervisor or the sampling statistician to make procedural errors or to skip vital validation steps.

The manual is divided into two parts addressed, respectively, to the listing supervisor and the sampling statistician. The first part covers the following topics: how to use the listing supervisor's work assignment and control form; an example of this form; the field duties of the listing supervisor; how to verify the completed maps and listing sheets; and how to validate the listing results against census counts. The second part covers: the review of listing results before beginning the sample selection process; the renumbering of eligible units; the HU drawing process; the assignment of the visiting week and of the serial number; the identification and control of selected units; examples of completed listing records after review by the statistician; examples of the sample selection form, the week assignment form and the map showing selected HUs and visiting weeks; and finally, an example of the sample summary sheet showing sample selection results.

Interviewer's Manual

This manual, in addition to its use as a comprehensive training and instruction guide for the interviewers, serves as the main reference for the survey data collection methodology. It consists of 11 chapters outlined below:

1. Introduction: survey objectives, overview of sampling and data collection methodologies;
2. Interviewer's Duties and Performance, including rules regarding interviewer's conduct, timeliness and supervision;
3. Interviewer's Check-Out Procedures;
4. General Field Procedures: step-by-step approach starting with the interviewer's arrival, contacting local authorities and making other public relation visits, answers to common questions asked about the survey, familiarization with local foods and lifestyles, locating and identifying sample HUs, recognizing and handling changes in occupancy status and, general substitution procedures;
5. General Interviewing Procedures: including getting ready for the interview, obtaining cooperation, establishing rapport with the respondent, and handling refusal and not-at-home cases;
6. The Visits: visiting schedule, how to modify schedule in cases where the pre-assigned schedule cannot be followed, preparations for the next visit;
7. General Application of the Questionnaire: how to ask questions, interviewer instructions, probing and, interview control techniques;

8. How to Record Information in the Questionnaire, including how to make corrections;
9. Detailed Instructions for Each Questionnaire Item, including concept definitions and probing requirements;
10. Field Edit Instructions: discussion of why, when, and how to conduct field edit; specific error checklist included;
11. Instructions for Visits to Local Market, for the purpose of obtaining price and weight data for reported food items.

The appendix to the Interviewer's Manual includes instructions for the interviewer's use of control forms and for the procedure of measuring weight and height. Reference materials such as tables to permit the estimation of age and a 3-year calendar are also included.

Field Supervisor's Manual

The supervisor's manual served as training guide and as reference guide for the field supervisors. It is divided in three chapters: Responsibilities and Tasks of the Supervisor, Instructions for Field Supervision and, Instructions for the Use of Control Forms. The following topics are discussed: field supervisor's role, duties, and required qualifications; materials and documents needed by the supervisor; supervisory area definitions and work organization; instructions for the control and flow of materials in the field; how to obtain cooperation in the field; criteria for timeliness and quality control (validation of selected unit's status, interview observation and validation, field editing of questionnaires); evaluation of interviewer's conduct and performance; and means of communication among interviewers, supervisors and the central office.

Office Operations Manual

The office manual served mainly as an operational guide for the pre-machine processing operations and for the preparation of sample materials each month. Its eight chapters cover the following main topics: calendar and organization of office activities, duties of office clerk, instructions for questionnaire check-in and transmission, general edit instructions, question-by-question edit instructions, illustrations of how corrections should be made, coding instructions, and instructions for converting measurements to standard units. The office manual also includes examples of control forms associated with the office activities and reference tables for the coding of questionnaire items.

Data Verification and Correction Manual

This manual refers specifically to the machine edit operation. It was used to train the team of data reviewers which also used it as a reference guide during their work. The manual begins by describing the machine edit activities and the duties of the data reviewers. This introduction is followed by general instructions and illustrations for checking and resolving errors. Detailed instructions are then provided for reviewing the output from each type of edit program: range checks, validity (screening for illegal entries) checks and consistency checks. In addition, the manual shows the main control forms used and lists the program specifications for each of these three types of edit.

Difficulties with Reference Manuals

Preliminary versions of the Interviewer's Manual and the Listing Manual were completed before the May pretest; these versions were substantially revised following the pretest experiences. Unfortunately, because of time constraints, the pre-machine edit and coding operations could not be pretested. A price had to be paid for this: after coding had begun, classification errors were found in the coding tables; as a result the first 2 months of data had to be recoded and the Office Manual had to be revised. This experience thus reinforces the emphasis usually placed on pretesting essentially all survey operations before their implementation.

A problem common to most of the manuals was that the final printed versions were not ready in time for training. This problem was due to untimely budget disbursements paired with severe time and personnel constraints; in the case of the Office Manual it was also due to the fact that it was not pretested and some procedures changed after the manual was considered complete. Even though fairly advanced draft versions were used, additional training was necessary to improve the users' mastery of the final versions.

B. SURVEY FORMS

A list of the forms used throughout the survey is provided below.

Data Collection and Sampling Forms

| | |
|--|---------|
| Listing Record | EBCM-1 |
| Data Collection Questionnaire | EBCM-2 |
| Creole Reference for Questionnaire | EBCM-2a |
| Worksheet for the Selection of HUs | EBCM-3 |
| Worksheet for Assignment of the Visiting Week to the HU | EBCM-4 |
| Worksheet for the Calculation and the Transmission of Weights | EBCM-5 |
| Record of Local Market Data | EBCM-6 |

Worksheet for the Calculation of
Variances (SUPER-CARP)

EBCM-SC1 to EBCM-SC8

Control Forms

| | |
|---|----------|
| Master Control Form of Sample EAs..... | EBCM-C1 |
| Mapping/Listing Supervisor's Work Assignment Form..... | EBCM-C2 |
| Sample Summary Sheet..... | EBCM-C3 |
| Questionnaire Lot Transmittal Form..... | EBCM-C4 |
| Communication and Problem Documentation Form..... | EBCM-C5 |
| Interview Observation and Verification Form..... | EBCM-C6 |
| Monthly Work Assignment Form for Interviewers and Supervisors..... | EBCM-C8 |
| Monthly Report of Interviewers' Performance..... | EBCM-C9 |
| Cumulative Index of Supervisors' Progress..... | EBCM-C10 |
| Pre-Machine Edit and Coding Work Assignment Sheet..... | EBCM-C11 |
| Record of Corrections Made (Errors Found) in the Questionnaire during Manual Edit..... | EBCM-C12 |
| Mapping/Listing Observation and Verification Form..... | EBCM-C13 |
| Questionnaire Log-In Control Form..... | EBCM-C14 |
| Record of Pages to be Keyed..... | EBCM-C15 |
| Identification Form for Transmittal of Questionnaires to Data Entry..... | EBCM-C16 |

Administrative Forms

| | |
|---|----------|
| Field/Office Telephone Communication Form..... | EBCM-A1 |
| Central Office Field Supervision Schedule..... | EBCM-A2 |
| Routing Slip for HECS Documents..... | EBCM-A3 |
| Identification Codes for HECS Documents..... | EBCM-A4 |
| Checklist for Interviewer's Materials | EBCM-A5 |
| Checklist for Supervisor's Materials..... | EBCM-A6 |
| Leave Request and Authorization Form for Office Clerks..... | EBCM-A7 |
| Field Activities Calendar..... | EBCM-A8 |
| Office Activities Calendar | EBCM-A9 |
| List of Selected EAs by Stratum and Sub-Stratum..... | EBCM-A10 |
| List of Selected EAs by Month..... | EBCM-A11 |
| Reference List of HECS Documents..... | EBCM-A12 |

Data Processing Control and Report Forms

| | |
|---|---------|
| Data Entry and Verification Work Distribution Form..... | INFO-1 |
| Data Entry and Verification Request Form..... | INFO-2 |
| Transmission of Corrections for the Data File..... | INFO-3 |
| List of Occupancy Status Changes in Reporting HUs..... | INFO-4 |
| Control File/Data File Match Report..... | INFO-5 |
| Cumulative Report of Number of Questionnaires by Processing Stage..... | EBCM-10 |

CHAPTER IV

FIELD OPERATIONS

The fieldwork for the HECS consisted of (1) listing housing units (HUs) in selected sections d'enumeration (SDEs) or enumeration areas, which were the primary sampling units; and, (2) interviewing the selected households. These operations were carried out every month with a workload of 24 SDEs per month. The listing operation preceded the interviewing by 1 or 2 months. This interval permitted time for the selection of sample HUs and the reproduction of maps.

A. MAPPING AND LISTING

The purpose of the mapping/listing operation was to construct an adequate frame of eligible HUs, the secondary sampling units, responding to the sample design criteria which are later discussed in Chapter VI. In addition to an exhaustive list from which to select the second stage sample with known probabilities of selection for each unit, the listing provided information for the identification and retrieval of the selected units. The maps permitted to establish the physical location of the SDEs and HUs on the ground.

Pilot Survey

During the pilot survey, the operation was tested on the nine SDEs. As mentioned in Chapter II, two reports were produced to evaluate the findings of the pilot survey and recommendations were made for the full-scale survey. It was found that the number of cartographers needed to be increased from five to seven and the number of vehicles to at least two. The listing record (EBCM-1) needed to be revised to provide more recording space for each HU and to include a space for recording such general SDE information as: boundary description, access instructions and, totals for listed HUs and eligible HUs. The maps needed to be more legible and provide better landmarks. Furthermore, the need was identified for extensive preliminary public relation and motivation work to be conducted before the actual operation could start because of extremely hostile conditions found in the field.

Full-Scale Operation

The listing manual and form were revised to incorporate the pilot survey recommendations and a team of seven listing enumerators and one supervisor was trained. The mapping/listing operation for the full-scale survey began in late July 1986 and ended in August 1987. Each month the listing enumerators worked in the 24 sample SDEs updating the SDE census map and completing a listing record for each SDE. All HUs within SDE boundaries

were identified and pertinent information was recorded on the listing forms.

Housing structures were spotted and numbered on the maps; appropriate symbols were used to identify vacant, under construction or occupied HUs. Physical characteristics such as water courses, bridges, landmark buildings, etc., were also noted. The listing was conducted by systematically visiting each dwelling unit within boundaries of the SDE to find how many HUs it had. The enumerator listed the HUs in systematic order and recorded the following information:

- for the SDE as a whole: identification, directions for locating, boundary description, and special remarks.
- for each HU: complete address of the housing structure, village or local area, enumerator-assigned serial number for the structure, number of HUs in the structure, enumerator-assigned serial number for HU within structure, occupancy status, and name of head of household. He also made notations in the remarks column on any unusual information about the household and provided a description when a complete street address could not be obtained.

Difficulties with the Mapping/Listing Operation

Throughout the survey, the operation had a tendency to fall behind schedule; this created a hectic pace for the processing of listing results to determine the sample and for the reproduction of maps for the following month. On one occasion, the data collection had to be postponed by 1 week because sample materials were not ready due to mapping/listing delays. The delays were due mainly to hostile conditions in the field which interfered with the supervisor's coordination efforts. A second problem was that, given the difficulties involved in accessing mountainous areas, one supervisor was insufficient and there were not enough resources to hire more.

One deficiency of the maps was that they were not always drawn to scale. This was due to the fact that the IHSI lacked the necessary equipment for the cartographers. Another factor was that with the mountainous Haitian topography, determining scales on maps was considered a difficult task without the appropriate resources. Nevertheless, this imperfection was inconsequential for the location of units because interviewers were usually accompanied by local guides.

As far as the listing results per se, it was found that the number of HUs differed from the 1980 Census counts often by more than 10%. The 10% figure was pre-determined as a tolerance limit based on experiences in other countries as well as the fact that there was a Census undercount estimated at 10%. The sampling statistician and the listing supervisor thus had to investigate all differences over 10%. However, it turned out that most of these differences were due, not to listing errors and not

necessarily to overall population growth or decline, but to actual migration across SDE boundary lines, that is, population growth and decline within an SDE. In more than one case, it was impossible to recognize the imaginary boundary established for the Census because new HU settlements covered the area.

B. DATA COLLECTION

Calendar

As discussed in Chapter I, the HECS sample was distributed over the 52 weeks of the year to ensure representativeness over time. However, because of political turmoil, it was necessary to terminate the fieldwork after 44 weeks. Interviews began November 4, 1986 and ended September 20, 1987. The complete calendar of the data collection schedule is shown in Figure 2 (Form EBCM-A8). The full-scale survey was preceded by the pilot test, previously discussed in Chapter I, which took place during May 1986.

Pilot Survey

Given that the HECS was the first survey of this scope in Haiti, the pilot survey was an indispensable reference in determining the final field organization and procedures. The two reports issued from the pilot survey (the Field Observation Report and the Questionnaire Evaluation Report) summarized the difficulties as well as the successes of the proposed methodology and provided a basis for improvements.

The main areas requiring solutions were: 1) better quality control for the interviews: better explanations of the procedures in the manual, more comprehensive training of interviewers, specific instructions for field edit, and more field supervisors; and, (2) arrangements to ensure the week-to-week continuousness of the data collection while taking into consideration the following logistical problems: insufficient commuting time between the end of interviews in one SDE and the start of interviews in a different SDE, not enough vehicles to distribute and pick up survey materials throughout the country on a timely basis each month, how to arrange time for refresher training and personal leave for the interviewers and, how to make payments of per diem and salaries in the field in a safe and efficient manner.

Figure 2

EBCM-88

CALENDRIER DES ACTIVITES DE TERRAIN

| Periode Mensuelle | BUREAU CENTRAL | TERRAIN | | | | | BUREAU CENTRAL |
|----------------------|---|--|--|------------------------------|-------------------------------|--|---|
| | Organisation du Depart des Enqueteurs (mercredi) | Arrivee Enqueteur sur le Terrain (dimanche) | 1 ^o . Reunion avec Superviseur (lundi) | Debut Collecte (mardi) | Fin Collecte (dimanche) | Reunion Depart avec Superviseur (lundi) | Retour Questionnaires et Enqueteurs (mercredi) |
| 1 | 29-Oct-86 | 02-Nov-86 | 03-Nov-86 | 04-Nov-86 | 30-Nov-86 | 01-Dec-86 | 03-Dec-86 |
| 2 | 26-Nov-86 | 30-Nov-86 | 01-Dec-86 | 02-Dec-86 | 28-Dec-86 | 29-Dec-86 | 31-Dec-86 |
| 3 | 24-Dec-86 | 28-Dec-86 | 29-Dec-86 | 30-Dec-86 | 25-Jan-87 | 26-Jan-87 | 28-Jan-87 |
| 4 | 21-Jan-87 | 25-Jan-87 | 26-Jan-87 | 27-Jan-87 | 22-Feb-87 | 23-Feb-87 | 25-Feb-87 |
| 5 | 18-Feb-87 | 22-Feb-87 | 23-Feb-87 | 24-Feb-87 | 22-Mar-87 | 23-Mar-87 | 25-Mar-87 |
| 6 | 18-Mar-87 | 22-Mar-87 | 23-Mar-87 | 24-Mar-87 | 19-Apr-87 | 20-Apr-87 | 22-Apr-87 |
| 7 | 15-Apr-87 | 19-Apr-87 | 20-Apr-87 | 21-Apr-87 | 17-May-87 | 18-May-87 | 20-May-87 |
| 8 | 13-May-87 | 17-May-87 | 18-May-87 | 19-May-87 | 14-Jun-87 | 15-Jun-87 | 17-Jun-87 |
| 9 | 10-Jun-87 | 14-Jun-87 | 15-Jun-87 | 16-Jun-87 | 12-Jul-87 | 13-Jul-87 | 15-Jul-87 |
| 10 | 08-Jul-87 | 12-Jul-87 | 13-Jul-87 | 14-Jul-87 | 09-Aug 87 | 10-Aug-87 | 12-Aug-87 |
| 11 | 05-Aug-87 | 09-Aug-87 | 10-Aug-87 | 11-Aug-87 | 06-Sep-87 | 07-Sep-87 | 09-Sep-87 |
| 12 | 02-Sep-87 | 06-Sep-87 | 07-Sep-87 | 08-Sep-87 | 04-Oct-87 | 05-Oct-87 | 07-Oct-87 |
| 13 | 30-Sep-87 | 04-Oct-87 | 05-Oct-87 | 06-Oct-87 | 01-Nov-87 | 02-Nov-87 | 04-Nov-87 |

Field Organization of Full-Scale Survey

In consideration of the survey design, and taking into account the experiences from the pilot survey, the data collection was organized in the following manner:

- 13 successive data collection periods of 4 weeks each
(referred to as the "survey months")
- 24 work areas per survey month
(a work area was defined to be an SDE)
- 1 interviewer per SDE
- 10 sample households per SDE interviewed at the rate of
2 or 3 households per week
- 4 visits per household to complete an interview (Tuesday,
Thursday, Saturday and Sunday)
- 5 supervisory areas
(i.e., four regions plus Port-au-Prince)
- 2 supervisors per region and 1 for Port-au-Prince

The data collection staff consisted of a total of 48 interviewers and 10 supervisors (including 1 substitute). The 48 interviewers, which included both men and women, were divided into two groups of 24 which alternated in and out of the field each month. Even though only 24 interviewers were required at a time to cover the 24 SDEs, the rotation arrangement was necessary for the following reasons. The required time for two-way travel, for public relations, reconnaissance and location of HUs, and for covering 4 weeks of interviewing was 5 weeks. Given that, (1) there was to be no interruption in the 52-consecutive-week schedule, and that (2) all SDEs changed each month, one set of interviewers was not sufficient. The arrangement also served to circumvent the logistical problems identified in the pilot. By returning to the central office each month, the interviewers were able to receive their pay, return the completed questionnaires and forms, and pick up the material for the next SDE. It would have been very difficult for IHSI to otherwise accomplish this because of the lack of vehicles.

To ensure their availability for the entire survey year, it was necessary to keep the interviewers on the payroll even when they were not in the field -approximately 3 weeks every other month. To derive an advantage from this constraint, their time away from the field was used in four ways: (1) to perform the HECS office operations discussed in Chapter V, i. e., editing, coding and control/preparation of materials, (2) to refresh their field operations training, (3) to replace or supplement interviewers in the field when additional help was required, and 4) to grant them personal or sick leave.

The supervision was organized roughly in the following manner. In addition to Metropolitan Port-au-Prince, each of the four planning regions (Nord, Transversale, Ouest and South) was considered a "supervisory work area". The monthly sample consisted of three urban SDEs and two rural SDEs in each region and, four SDEs in Metropolitan Port-au-Prince (the sample allocation is discussed in detail in Chapter VI). There were thus

two supervisors assigned to each region, with a workload varying between two and four SDEs per month, and one assigned to Metropolitan Port-au-Prince with a constant workload of four SDEs per month. The exact division of the SDEs among supervisors was determined by the location and spread of SDEs in the sample; in fact, there were times when a supervisor's work area crossed over the region's boundary for operational convenience. Unlike the interviewers, the supervisors spent the entire year in the field with two exceptions: (1) they traveled once a month to the central office to receive their pay, participate in field debriefing and consultation with the technical staff, turn in their interviewer performance ratings (form EBCM-C9), and pick up materials for the next month; and (2) through the use of a 10th (substitute) supervisor who rotated throughout all supervisory areas, they were permitted to be away from the field during 1 month in which they worked in the central office.

A "Work Assignment for Supervisors and Interviewers" form (EBCM-C9) was issued each month to indicate the sample SDEs, the interviewers' and supervisors' work assignment areas, the supervisors' recommended location and the dates for his visit to the central office. The supervisor/interviewer ratio never exceeded 1 to 4, that is, 1 supervisor per 4 interviewers.

Training of Field Staff

From a total of approximately 700 applicants, 90 were selected to participate in the training sessions on the basis of ratings they received on a preliminary test measuring availability, experience, general knowledge (math, geography, etc.), problem-solving ability, insight and motivation. The final choice of 48 interviewers/office clerks and 10 supervisors was made after training.

Training sessions began September 22 and lasted 4 weeks. They included classroom training as well as field exercises. Based on performance in the classroom portion, a few participants were eliminated and did not participate in the field training session.

The classroom training was based on a thorough coverage of the Interviewer's Manual. Class discussions were encouraged. There were numerous classroom exercises and quizzes. Trainers demonstrated interviewing techniques and participants conducted simulated interviews. Participants also had to practice using the questionnaire at home or with their neighbors as homework. A final examination was administered in two parts. In the first part, participants answered questions about the survey design and methodology. They had to prove their understanding of procedures and concepts. It also included a timed test in which they had to find specific information in the manual rapidly; this test served to evaluate their mastery of the manual as a reference tool. The second part of the written test consisted of filling out the questionnaire based on an interview summary given to them in prose. This test permitted to see if they understood the information given and if they could record it correctly and in the proper place.

The rating they obtained on their test scores determined whether they would be considered further for the job. During field training, participants went out in pairs in the surrounding areas of Port-au-Prince, and conducted interviews under observation of the HECS technical staff. Based on the ratings for their field performance, the best 58 were retained. Using both the classroom and field performances as criteria, 10 supervisors were also selected among the participants. Remedial training sessions were held for all interviewers and supervisors to correct problems found during the field exercise. Separate training was given to the supervisors using the Supervisor's Manual.

Preparation of Field Materials

Each month, during their time away from the field, a team of four to six interviewers worked on the preparation of materials for the following month's fieldwork. This work took place approximately 3 weeks before their departure to the field. The work concerned mainly the preparation of sample materials and was supervised by the sampling statistician as well as the office supervisor.

The sample selection form and the week assignment form completed by the sampling statistician on Lotus 1-2-3 were distributed to these clerks. For each of the 24 sample SDEs, the clerks: circled and renumbered the selected HUs on the listing records and on the maps, transcribed their identification from the listing record on to a draft sample summary sheet which then went to the Data Processing Unit for entry on the mainframe computer, verified the sample summary sheet printed by the mainframe, assembled the different modules to prepare 24 lots of 10 questionnaires, and entered the HU identification in each questionnaire. Each one of the above steps was verified in its entirety by a different clerk.

Interviewer Departure and Arrival

Pre-departure activities took place over a period of 2 days each month. The interviewers left for the field approximately 1 week before their first interview. Before the actual check-out, a refresher training session was held where field operations were reviewed and interviewers were briefed on any changes in procedures. Errors found during editing, coding and conversion were discussed so that interviewers could prevent them in the field. Interviewers had the opportunity to ask questions before they went out in the field.

Following the refresher training, materials were distributed. The interviewers had to ensure that they could retrieve the SDE and HUs with the directions given. They also had to check all materials before leaving as they were held responsible if missing or defective material later affected their ability to work. Before departing, they signed the questionnaire transmittal form (EBCM-C4) and the materials checklist form (EBCM-A5).

Upon returning to the central office, interviewers turned in the sample materials for each SDE as well as all other forms and equipment. They then organized the questionnaires into SDE lots which contained normally 10 questionnaires, except for cases where fewer than 10 interviews were obtained. The interviewers had to fill out a questionnaire log-in form (EBCM-C14) by listing the originally-selected units and indicating the interview completion status for each: for example, "vacant - replaced by unit R1", or "completed", etc., which was later verified by the office supervisor. This form served to obtain a preliminary count of completed interviews (response rate) for the given month. The interviewer also had to sign the questionnaire transmittal form (EBCM-C4) and the materials checklist form (EBCM-A5). The field equipment was examined and then stored for the next field departure; replacement equipment was procured, as necessary.

Once all materials were logged in, a debriefing session was held in which the interviewers provided a summary of their field experience, including problems found and solutions adopted. They were then given feedback on their actions.

General Field Procedures

The interviewers were able to locate the SDE using the directions provided in the listing record. They were expected to arrive in the SDE at least a couple of days before the first interview. This preliminary time was used to find lodging, make public contacts, get acquainted with the local lifestyle and food types, retrieve the sample housing units and make appointments for the visits.

Because of the socio-political tensions present in the country and the hostilities arising from them, the interviewers were instructed not to start working until they had enlisted the support of local authorities as well as other persons of influence. Visits were thus made to the "prefet" (head of the province or department), the "magistrate" (head of the commune), the "chef de section rurale" (head of the local rural section), the local armed forces representative, religious leaders, and sometimes other local administrators to the extent that they were influential members of the community. The interviewer was trained to give appropriate explanations about the survey, stressing its importance, and to answer the questions that are typically asked. The public relation visits were crucial to the success of the interviews and to the physical safety of the interviewer since, in most cases, the community would as a unified body either accept or reject the whole survey. If they accepted it, the interviewer normally did not have great problems in obtaining the cooperation of individual households; if, on the other hand, they rejected it, then very often, the interviewer was ousted.

Once public contacts were established, the interviewer had to take time to familiarize himself/herself with the local milieu; that is, to find out: how local residents distributed their time, what their normal activities were, how many meals a day were common, at what times they ate, whether

eating away from home was common, whether there were times or days when visits would be unwelcome, what the best time for visits was, what days were market days, what products were commonly consumed in the area and their local names, and other useful information.

The interviewers also took the time, using the SDE map and a local guide, to do a general reconnaissance permitting them to locate the housing units and contact the heads of households. They would then set up an appointment in accordance with the interview schedule. If the household was to be absent during its scheduled interview week, the interviewer, following strict guidelines, tried to interchange weeks with another sample household in order to keep it in sample. The interviewers were trained to recognize changes in the occupancy status such as: a new household, an ineligible sample unit, and consolidated or divided housing units. They had to decide, based on strict substitution rules, whether to keep a unit in sample or locate a substitute. Once all this was accomplished, the interviewer was ready for the interviews.

Household Interviews and Market Visits

The interviewers prepared themselves before the interview by reviewing the manual and the questionnaire, organizing the materials needed and checking their appearance. They then made contact with the respondent, and tried to obtain cooperation. They were trained to use a standard introduction, and in procedures for establishing rapport with the respondent and dealing with refusals.

The respondent was usually the husband and/or the wife or other head of household; proxy respondents were also accepted, using strict guidelines. Third parties were usually not allowed during interviews except with the respondent's express consent. This was to discourage any influential effect on the respondent's answers as well as to protect the confidentiality of the interview.

On an average, the visits lasted approximately 1 hour each, although, the first interview usually required additional time to convince the respondent to participate. In addition to completing the data collection questionnaire (EBCM-2) at the household, the interviewer went to market each week to collect price and weight information for food items which he recorded in a separate form (EBCM-6). He used food scales to weigh the products. The information obtained was later transcribed to the data collection questionnaire for each product.

Field Quality Control

Quality control was achieved through careful planning and rigorous supervision of the field and office work. The thorough training given to the survey personnel helped prevent the occurrence of errors. Moreover, the HECS data were subject to three types of edit: field edit (by both interviewers and supervisors), manual edit, and machine edit. The office operations (edit, coding and conversion) were subject to 100% verification by a different clerk.

Both interviewers and supervisors were given guidelines for quality control in the field. These included instructions for preventing errors and specific error checklists. Other quality control measures included scheduled and surprise visits to the interviewer, interview observation where a detailed observation form (EBCM-C6) was filled, interview verification (partial reinterviews), independent remeasurement of food weights and prices, and of children's weight and height. Central office staff made surprise visits to both the supervisors and interviewers in the field. As far as the listing operation, one tactic used was to have the supervisor perform an independent listing of a portion of the SDE and compare it with the enumerator's work.

Another form of quality control was the continuous evaluation of the interviewers' work. Their field performance was rated throughout the survey by the supervisors. The supervisor turned in every month Form EBCM-C9 "Interviewers' Work Report" which summarized for each interviewer the number of completed questionnaires, the number of substitutions made, the number of questionnaires received late and also provided a verbal account of each interviewer's performance. It also included comments by the supervisor on the interviewer's performance. In addition, the number of corrections made (that is, errors found) in the questionnaire during manual edit was tallied in form EBCM-C11. Interviewers who showed problems either in their performance evaluation report or the summary of questionnaire errors were counseled or given remedial training. Some were terminated when their performance did not improve. Interviewers who had remarkable performances were given letters of commendation to encourage them to continue on that track.

CHAPTER V

DATA PROCESSING OPERATIONS

The data processing operations can be divided into pre-machine and machine activities. The pre-machine processing, or "office" activities, included the pre-machine (or manual) edit, coding and conversion of items in the completed questionnaires and are documented in detail in the Office Operations Manual. Computer processing activities included: data entry, machine edit and imputation, and tabulations. The machine edit activities are documented in the Data Verification and Correction Manual.

A. OFFICE OPERATIONS

Office Calendar and Organization

The office activities took place each month during the 3 weeks following data collection. A complete calendar is presented in Figure 3 (Form EBCM-A9). Following check-in, a 2-day refresher training session was held in which office procedures were reviewed and eventual changes in the procedures were discussed before work sessions actually began. The pre-machine edit operation lasted about 1 week; the coding and conversion began after the pre-machine edit was completed and lasted also approximately 1 week. The last week was devoted to field preparations for the next round of interviews.

The office operations staff consisted of 24 clerks and 3 supervisors. As discussed in the previous chapter, an arrangement was adopted whereby the field interviewers rotated in and out of the field, working as office clerks during the time they were not in the field. Their experience with the questionnaire made them well suited for the editing task. In effect, only 16 to 20 interviewers actually worked on the edit and coding given that the others worked on the preparation of sample materials for the following month or served as substitute interviewers in the field.

Each clerk had a total workload of about 6 lots: 3 for editing and 3 for coding/conversion (1 lot corresponded to 1 enumeration area (SDE), i.e., 10 questionnaires). The average number of questionnaires per day was thus 6: 60 questionnaires in 10 days. The work lots were assigned by the supervisor in such a way that the 3 lots the clerk was assigned for edit were different from the lot for which he collected the data. The lots he received for coding were furthermore different from the edit lots and from the data collection lot. Since the edit and coding operations were subject to 100% verification, some of the lots the clerk received had already been edited or coded. However, he was never allowed to verify a lot that he had edited or coded himself.

Figure 3

EECM-A9

CALENDRIER DES ACTIVITES DE BUREAU

| Periode Mensuelle (au BC) | ARRIVEE ET FORMATION | | | CRITIQUE ET CODIFICATION | | | | | PREPARATION MATERIEL | | PREPARATIONS ENQUETEURS POUR DEPART | | |
|---------------------------------|---|---|-----------|------------------------------|-------------------------------|----------------------------|--------------------------|--|----------------------|-------------------|---|-----------|--------------------------------------|
| | Controle de l'Arrivee au BC (mercredi) | Recyclage de Bureau (jeudi et vendredi) | Formation | Debut Critique (lundi) | Fin Critique (vendredi) | Debut Codif. (lundi) | Fin Codif. (jeudi) | Transmission a la Saisie (vendredi) | Debut (lundi) | Fin (vendredi) | Recyclage de Terrain (lundi et mardi) | Formation | Organisation Depart (mercredi) |
| | | | | | | | | | | | | | |
| 1 | 03-Dec-86 | 04-Dec-86 | 05-Dec-86 | 04-Dec-86 | 12-Dec-86 | 15-Dec-86 | 16-Dec-86 | 19-Dec-86 | 13-Oct-86 | 24-Oct-86 | - | - | 29-Oct-86 |
| 2 | 31-Dec-86 | 01-Jan-87 | 02-Jan-87 | 03-Jan-87 | 09-Jan-87 | 12-Jan-87 | 15-Jan-87 | 16-Jan-87 | 10-Nov-86 | 21-Nov-86 | 24-Nov-86 | 25-Nov-86 | 26-Nov-86 |
| 3 | 28-Jan-87 | 29-Jan-87 | 30-Jan-87 | 02-Feb-87 | 06-Feb-87 | 09-Feb-87 | 12-Feb-87 | 13-Feb-87 | 08-Dec-86 | 19-Dec-86 | 22-Dec-86 | 23-Dec-86 | 24-Dec-86 |
| 4 | 25-Feb-87 | 26-Feb-87 | 27-Feb-87 | 02-Mar-87 | 06-Mar-87 | 09-Mar-87 | 12-Mar-87 | 13-Mar-87 | 05-Jan-87 | 16-Jan-87 | 19-Jan-87 | 20-Jan-87 | 21-Jan-87 |
| 5 | 25-Mar-87 | 26-Mar-87 | 27-Mar-87 | 30-Mar-87 | 03-Apr-87 | 06-Apr-87 | 09-Apr-87 | 10-Apr-87 | 02-Feb-87 | 13-Feb-87 | 16-Feb-87 | 17-Feb-87 | 18-Feb-87 |
| 6 | 22-Apr-87 | 23-Apr-87 | 24-Apr-87 | 27-Apr-87 | 01-May-87 | 04-May-87 | 07-May-87 | 08-May-87 | 02-Mar-87 | 13-Mar-87 | 16-Mar-87 | 17-Mar-87 | 18-Mar-87 |
| 7 | 20-May-87 | 21-May-87 | 22-May-87 | 25-May-87 | 29-May-87 | 01-Jun-87 | 04-Jun-87 | 05-Jun-87 | 30-Mar-87 | 10-Apr-87 | 13-Apr-87 | 14-Apr-87 | 15-Apr-87 |
| 8 | 17-Jun-87 | 18-Jun-87 | 19-Jun-87 | 22-Jun-87 | 26-Jun-87 | 29-Jun-87 | 02-Jul-87 | 03-Jul-87 | 27-Apr-87 | 08-May-87 | 11-May-87 | 12-May-87 | 13-May-87 |
| 9 | 15-Jul-87 | 16-Jul-87 | 17-Jul-87 | 20-Jul-87 | 24-Jul-87 | 27-Jul-87 | 30-Jul-87 | 31-Jul-87 | 29-May-87 | 05-Jun-87 | 08-Jun-87 | 09-Jun-87 | 10-Jun-87 |
| 10 | 12-Aug-87 | 13-Aug-87 | 14-Aug-87 | 17-Aug-87 | 21-Aug-87 | 24-Aug-87 | 27-Aug-87 | 28-Aug-87 | 22-Jun-87 | 03-Jul-87 | 06-Jul-87 | 07-Jul-87 | 08-Jul-87 |
| 11 | 09-Sep-87 | 10-Sep-87 | 11-Sep-87 | 14-Sep-87 | 18-Sep-87 | 21-Sep-87 | 24-Sep-87 | 25-Sep-87 | 20-Jul-87 | 31-Jul-87 | 03-Aug-87 | 04-Aug-87 | 05-Aug-87 |
| 12 | 07-Oct-87 | 08-Oct-87 | 09-Oct-87 | 12-Oct-87 | 16-Oct-87 | 19-Oct-87 | 22-Oct-87 | 23-Oct-87 | 17-Aug-87 | 28-Aug-87 | 31-Aug-87 | 01-Sep-87 | 02-Sep-87 |
| 13 | 04-Nov-87 | 05-Nov-87 | 06-Nov-87 | 09-Nov-87 | 13-Nov-87 | 16-Nov-87 | 19-Nov-87 | 20-Nov-87 | 14-Sep-87 | 25-Sep-87 | 28-Sep-87 | 29-Sep-87 | 30-Sep-87 |

N. B. Quand les dates projetees tombent sur des jours de conge, des ajustements se feront afin de respecter les delais.

Pre-Machine Edit, Coding and Conversion

The pre-machine edit operation consisted of reviewing the questionnaires to prepare them for accurate data entry. Consistency checks, range checks and imputation were not performed at this stage, since this is usually handled by computer. The main objectives of the pre-machine edit were to: ensure that the questionnaire identification was correct, ensure that all entries were legible, verify that numeric fields were properly recorded, change literal responses into numeric values, control the number of responses provided for each question, eliminate unnecessary entries, check that responses were recorded in appropriate spaces, convert fractions into decimal units, nullify blank questionnaire pages, and keep a record of pages to be keyed for control purposes.

The coding consisted of replacing literal answers into numeric codes. The main questionnaire items requiring coding were: interview completion status, educational level, occupation and industry of household members, food products, categories of goods and services. Coding tables were provided in the Office Manual for this purpose. The conversion work consisted of changing the measurement unit for certain items into uniform standard units; a table was provided which indicated for every food, goods or service item the standard unit of measurement chosen for analytical purposes. For instance, milk was to be measured in liters; if it was recorded in ounces, then the coding/conversion clerk converted the ounces into liters. Conversion was usually not necessary when the product reported was one of the pre-printed items in the questionnaire because the reserved space indicated the choice of measurement unit to be recorded.

B. COMPUTER PROCESSING

Data Entry and Creation of Data Files

Following office operations, questionnaires were transmitted to the Data Processing (DP) Unit for data entry with 100% verification. This work was performed by eight operators and overseen by one supervisor. It took approximately 45 minutes to enter one questionnaire and 2-3 weeks to enter the monthly workload of 480 questionnaires (240 keyed twice). An additional time of 1 week was spent matching the results of the two independent data entries and reconciling differences.

The data entry was performed on remote terminals connected to a Wang mainframe computer. The operator was prompted first for questionnaire identification data and the number of the page to be keyed; at that point, a screen appeared showing all the source codes and the operator then entered the information for each field. The program had certain built-in edit specifications such as the identification of the specific SDEs in sample each month and the response codes possible for each question. This permitted certain edit checks during the data entry itself. For instance, if the operator entered an incorrect SDE identification number or housing

answer into a 3-digit field, the machine rejected it. Likewise, if a response code other than the pre-printed ones in the questionnaire was entered, the machine did not accept it.

The data file emerged from the reconciliation of the two data entry executions. It was then run through a data file/control file match program which identified a responding questionnaire for each sample unit. This control was extremely important to ensure that the achieved sample matched the selected sample and that there were no lost questionnaires. It also served to adjust the estimated response rate and the nonresponse adjustment factors, since the preliminary calculations were based on simple counts as the questionnaires arrived from the field. Another control that the data file was subject to was the verification of the interview completion status. By verifying certain responses, the program listed all questionnaires which had ineligible, consolidated or divided units. This listing was used to refine weight adjustments based on changes in occupancy status.

Machine Edit and Imputation

Two computer programs were developed for the HECS at the Center for Agricultural and Rural Development (CARD) to detect range and consistency errors. The first program uses a parameter file as reference to determine whether or not entries fall within specified ranges. This parameter file contains specifications for virtually all the fields in the questionnaire which were divided into two groups: those for which the possible entries were elements of a discrete set of values and those for which the possible values constituted a continuous range bounded by a lower limit and an upper limit.

Outputs from these programs were reviewed by a team of data verifiers; appropriate corrections were identified for specific fields on the form EBCM-INF03 "Transmission of Corrections for the Data File" and transmitted to the DP Unit. The corrections were entered at first on a separate file, which was validated by the data verifiers, before being applied to the data file. Edit programs were run on the data file several times until an output free of resolvable errors was obtained. The unresolved errors were flagged on the data file for imputation purposes.

As mentioned in Chapter I, the machine edit operation was not completed. Only 3 months of data had been edited when the survey activities were canceled due to the political unrest in the country.

Tabulations

The survey plans called for the tabulation and publication of survey estimates and selected variance estimates. A draft tabulation plan was prepared by IHSI. But with the discontinuation of the survey, work on the finalization of this plan and on the development and programming of tabulation specifications was not completed.

CHAPTER VI

SAMPLE DESIGN AND ESTIMATION PROCEDURES

A. SAMPLING PLAN

Universe and Unit of Analysis

The unit of analysis for which data was collected in the HECS was the household. The universe covered was all households in the Republic of Haiti. Excluded from this universe was the population in hospitals, prisons, military quarters and other group quarters.

Sampling Frames

A two-stage probability sample design was used for the HECS. The primary sampling units (PSUs) were selected from the list of the 4,730 sections d'enumeration (SDEs) or enumeration areas defined for the 1982 census, taking advantage of the existing maps. The frame also included zone of residence (urban or rural) as a classification variable and the number of households tabulated from the 1982 census as a measure of size.

The secondary sampling units (SSUs) were housing units (HUs). In the HECS, the HU is defined as the space occupied by one household. It is used as a sampling unit instead of the household itself because it is more readily identifiable in the field.

The sampling frame for the second stage was constructed through a mapping/listing operation within SDEs selected in the first stage. This operation was needed to ensure that this sampling frame: was up-to-date, complete and accurate in its coverage of the target population; contained sufficient elements to permit the identification of the sample units; and most importantly, gave knowledge of the exact probability of selection of each unit. The fieldwork for this operation was discussed in Chapter IV.

Together the maps and listing records provided a complete account of all HUs within the SDE. While the maps facilitated the spatial identification and retrieval of the SDEs and HUs, the list provided such detailed information as: name of head of household, street address and occupancy status for each HU. The latter was used to screen for permanently occupied HUs, the only units eligible for selection under the survey requirements.

Stratification

To increase efficiency in the sample design, the universe was divided into homogeneous areas or strata. The first level of stratification corresponded to nine major geographic domains for which separate survey estimates were to be tabulated: the four planning regions divided into urban and rural areas and the metropolitan area of Port-au-Prince as a separate stratum. Census SDEs were classified as urban, rural or "quartier". For HECS purposes, "quartiers" were included in the rural stratum, since they were considered more rural than urban in nature.

In order to improve the efficiency of the sample design, the nine major geographic strata were subdivided into substrata with socioeconomic homogeneity. Hence, metropolitan Port-au-Prince was divided into three (low, middle and high) socioeconomic substrata while the other major cities, Cap-Haitien, Gonives and Les Cayes, were divided into two (low and middle/high) such substrata. The basis for this substratification was field observation guided with maps; one very useful reference was a map prepared by George Anglade of the University of Quebec in Montreal which showed general boundaries for the different Port-au-Prince socioeconomic areas.

Since ecological zone is a variable which correlates with socio-economic variables (discussed for the case of Haiti in particular in George Anglade's work: Atlas Critique d'Haiti), the substrata in the rural areas were formed by dividing the regions into "plain" and "mountain" areas. A topographic map of Haiti was used for this work.

It should be noted that the substratification was designed to increase the precision of the stratum estimates and not for the purpose of producing reliable substratum level estimates. To achieve even greater efficiency in the sample design, the SDEs were ordered geographically in a serpentine fashion within a substratum to obtain implicit stratification during the systematic selection process.

Sample Size and Allocation

The sample size was determined by the precision required for the survey estimates for each domain as well as the resource and operational constraints. In the absence of current data for estimating variances for key survey variables, experiences from other surveys of this type in developing countries were used. Nonsampling error considerations were also taken into account in view of the fact that quality and operational controls are inversely affected by sample size. The total sample size required was thus determined to be around 3,000 households; this size was judged sufficient for obtaining reliable urban and rural estimates at the national level as well as estimates of predominant characteristics at the regional level (with coefficients of variation mostly within 15 percent). It was recommended, however, that estimates of more rare characteristics only be made at the national and urban/rural levels as such estimates for smaller domains would suffer from high sampling errors.

Because the same precision was needed for the estimates from each major geographic domain, a similar sample size had to be allocated to each. At the same time, it was necessary to consider the differential variability of socioeconomic characteristics and cost of fieldwork in the urban and rural areas. Variability was higher in the urban areas than in the rural areas, and even higher in metropolitan Port-au-Prince; field costs were higher in the rural areas. Taking into account these various considerations as well as the data collection methodology, the actual sample size was determined to be 3,120 HUs: 312 SDEs and 10 HUs per SDE, allocated to the strata as indicated in Table 3. The number of SSUs per PSU was limited to 10 because certain socioeconomic characteristics may have a high intraclass correlation. As shown in studies of intraclass correlation for similar surveys in other countries the range of optimum values for the number of sample households per cluster (here, the SDE) normally includes 10.

Table 3. SAMPLE SIZE AND ALLOCATION BY STRATUM

| Region | Total | | Urban | | Rural | |
|--------------|-------|------|-------|------|-------|------|
| | SDEs* | HUs* | SDEs | HUs | SDEs | HUs |
| North | 65 | 650 | 39 | 390 | 26 | 260 |
| Transversale | 65 | 650 | 39 | 390 | 26 | 260 |
| South | 65 | 650 | 39 | 390 | 26 | 260 |
| West | | | | | | |
| Metro PaP | 52 | 520 | 52 | 520 | -- | -- |
| Remainder | 65 | 650 | 39 | 390 | 26 | 260 |
| Total Nation | 312 | 3120 | 208 | 2080 | 104 | 1040 |

*Number of Sample SDEs and Sample HUs

The sample allocation specified in Table 3 resulted in different sampling weights for each stratum; however, the differential weights did not present a problem for estimation, since they were calculated by computer. Within each stratum, the sample was allocated to the substrata in proportion with the (1982 census) number of households in the substrata resulting in an approximately self-weighting sample within stratum.

Taking into account seasonal variability in consumption and expenditures, data collection was designed to take place throughout the year. To ensure that each of the 52 weeks of the year was properly represented, the sample SDEs were divided into 13 national subsamples, each randomly assigned to 1 month (4 weeks) of the year. A monthly subsample for each region consisted of three urban SDEs and two rural SDEs, while the monthly subsample for metropolitan Port-au-Prince consisted of four SDEs. In this manner, a national probability sample of 240 HUs in 24 SDEs was interviewed each month.

Sample Selection Procedures

1. First Sampling Stage

The sample of SDEs was selected systematically with probability proportional to size. Within the rural strata (by region), each SDE was assigned to either the plains or mountain substratum. Within metropolitan Port-au-Prince, each SDE was assigned to one of the three substrata: low, medium or high. Within the other major cities (Gonaives, Cap-Haitien and Les Cayes), SDEs were assigned to either low or medium/high substrata.

Within each such formed substratum (h), the following steps were performed. The example in Figure 4 illustrates the selection method.

- a. the SDEs were ordered geographically in a serpentine manner, to provide implicit stratification;
- b. the number of households from the 1982 census was recorded for each SDE to serve as measure of size (M_{hi});
- c. the measures of size were then cumulated down the list of SDEs; their sum was the total number of households in the substratum (M_h);
- d. the number of sample SDEs in the substratum (n_h) was determined by proportional allocation using the following formula:

$$n_h = \frac{M_h}{M} \times n$$

where,

- M = total number of households in the stratum, from '82 census;
 M_h = number of households in substratum h, from '82 census; and
n = number of sample SDEs allocated to the stratum, from Table 3.

- e. the sampling interval (I_h) was determined by:

$$I_h = \frac{M_h}{n_h}$$

- f. A random start (R_h) between 0 and I_h was obtained from a random numbers table, and the sample "hits" were obtained as follows:

$$S_{hi} = R_h + (i-1) \times I_h, \text{ for } i=1, \dots, n_h$$

the i-th selected SDE was the one with the cumulated measure nearest to S_{hi} without exceeding it.

Figure 4. Illustration of First-Stage Sample Selection

Allocation¹

Region: Sud $n_h = \frac{M_h}{M} n = \frac{2610}{16965} (26) = 4$

Stratum: Rural

Sub-stratum: Plain $I_h = \frac{M_h}{n_h} = \frac{2610}{4} = 652.50$

$R_h \in (0.00, I_h] = 276.91$

| SDE i | Number of Households M_{hi} | Cumulated No. of Households | Selection Interval | S_h^i |
|----------|-------------------------------------|-----------------------------------|-----------------------|---------|
| 1 | 250 | 250 | 1-250 | |
| 2 | 140 | 390 | 251-390 | |
| 3 | 195 | 585 | 391-585 | 277 |
| 4 | 55 | 640 | 586-640 | |
| 5 | 200 | 840 | 641-840 | |
| 6 | 390 | 1230 | 841-1230 | 930 |
| 7 | 265 | 1495 | 1231-1495 | |
| 8 | 105 | 1600 | 1496-1600 | 1582 |
| 9 | 180 | 1780 | 1601-1780 | |
| 10 | 290 | 2070 | 1781-2070 | |
| 11 | 80 | 2150 | 2071-2150 | |
| 12 | 460 | 2610 | 2151-2610 | 2235 |

Total Sub-stratum 2610 = M_h

¹The allocation is based on the following fictitious numbers:

$M = 16,965$ (number of households in rural stratum of south region)

$M_h = 2,610$ (number of households in "plain" substratum)

$n = 26$ (number of SDEs allocated to the stratum from Table 3)

2. Second Sampling Stage

The second-stage selection procedure is documented in detail in Part B of the HECS manual titled "Instructions for the Supervision of Mapping and Listing and Procedures for Housing Unit Sample Selection".

As noted earlier, the listing operation provided the records which served as sampling frames for the second stage of selection. A listing record was obtained for each sample SDE. Using as input information the total number of eligible HUs in the SDE obtained during the validation of the listing record, a Lotus 1-2-3 worksheet (EBCM-3) was used to draw a systematic random sample of 15 HUs from each SDE. An example of this worksheet is furnished in Figure 5.

The j -th selected unit in the i -th SDE was given by:

$$R_{hi} + (j-1) \times I_{hi}, \text{ for } j=1, \dots, 15$$

rounded up to the next integer

where,

$$I_{hi} = M_{hi}/m_{hi} = \text{sampling interval within the } i\text{-th SDE}$$

$$R_{hi} = \text{a random number between 0 and } I_{hi}$$

By randomly extracting five of these HUs for use as eventual replacement units, a random sample of 10 HUs was left for interviewing purposes.

Once the sample HUs were selected, a second Lotus worksheet (EBCM-4) was used to assign to each a random number between 1 and 4 representing the interview week. An example of this worksheet is included in Figure 6. In order for the sample to be representative over time, randomness in the week assignment was a requirement; only this way could the week measured be considered a representative cross-section of the household's socio-economic year.

Selected HUs were then serially numbered, circled both on the map and on the listing record and recorded on a Sample Summary Sheet (EBCM-C3) for control purposes.

B. ESTIMATION PROCEDURES

As mentioned in Chapter I, the HECS estimation phase has not been completed because of the termination of financial and technical assistance to the IHSI. Estimation weights were calculated; but estimates and variances have not been produced. Still, for documentation purposes, a summary of the procedures is presented below.

Figure 5

| | | | |
|--|--|---|---|
| Formulaire ERC#-3 | | 1. IDENTIFICATION DE LA SDE | |
| FICHE DE TRAVAIL POUR LE TIRAGE DES UNITES DE LOGEMENT | | 1. Departement: | Nord |
| | | 2. Commune: | Grande Riv. du Nord |
| | | 3. Ville: | Grande Riv. du Nord |
| | | Section Kurale: | |
| | | 4. Code Geographique: | 2331-90 01-03 |
| | | 5. No. d'Ordre SDE | |
| | | Selectionnee: | 01-01-19-19 |
| II. BASE DE TIRAGE | | | |
| a. Nombre d'ULs Eligibles (m): | 432 | | |
| b. Nombre d'ULs a Choisir (n): | 15 | | |
| c. Intervalle de Sondage (O): | 28.80 | | |
| d. Point de Depart Aleatoire (k): | 14.29 | | |
| III. TIRAGE DES UNITES DE LOGEMENT | | | |
| (1) | (2) | (3) | (4) |
| No. de ligne (1) | Dumul. Systematique $[R+(-1)*O]$ | Logement Eligible Correspondant (annonci par exces) | No. d'Ordre du Logement Selectionne |
| 1 | 14.29 | 15 | _____ |
| 2 | 43.09 | 44 | _____ |
| 3 | 71.89 | 72 | _____ |
| 4 | 100.69 | 101 | _____ |
| 5 | 129.49 | 130 | _____ |
| 6 | 158.29 | 159 | _____ |
| 7 | 187.09 | 188 | _____ |
| 8 | 215.89 | 216 | _____ |
| 9 | 244.69 | 245 | _____ |
| 10 | 273.49 | 274 | _____ |
| 11 | 302.29 | 303 | _____ |
| 12 | 331.09 | 332 | _____ |
| 13 | 359.89 | 360 | _____ |
| 14 | 388.69 | 389 | _____ |
| 15 | 417.49 | 418 | _____ |
| IV. ED-ENTRILLOA RESERVE | | | |
| (prendre les 5 premiers chiffres de la colonne a cote et encercler le numero de ligne correspondant dans la section III ci-dessus. Les noter sur la Fiche d'Inventaire, ERC#-1, comme R1,...,R5. en respectant l'ordre dans lequel ils apparaissent.) | 15 4 9 14 1 9 5 4 12 13 | | |

Figure 6

| Formulaire EBCM-4 | | 1. IDENTIFICATION DE LA SDE | |
|---|------------------------------|----------------------------------|---|
| ASSIGNATION DE LA SEMAINE DE VISITE AU LOGEMENT | | | |
| | | 1. Departement: | Ouest |
| | | 2. Commune: | |
| | | 3. Ville: | Carrefour |
| | | Section Rurale: | |
| Periode Mensuelle: | 1 | 4. Code Geographique | 1111-8: 69-0: |
| Commerçant: | 12-Oct-85 | 5. No. Ordre SDE | |
| Terminant: | 06-Nov-86 | Selectionnee: | 01-24-92-42 |
| II. ASSIGNATION ALÉATOIRE DU NOMBRE DE LOGEMENTS ET DE LA SEMAINE AU GROUPE | | | |
| Groupe | Nombre de Logements Assignés | Semaine Assignée | Date |
| G1 | 3 | 2 | 19-Oct-86 |
| G2 | 2 | 3 | 26-Oct-86 |
| G3 | 3 | 4 | 02-Nov-86 |
| G4 | 2 | 1 | 12-Oct-86 |
| III. ASSIGNATION DU GROUPE ET DE LA SEMAINE AU LOGEMENT | | | |
| Groupe | Semaine de Visite | No. d'Ordre Logement Selectionné | (Instructions: Assigner a la main le numero d'ordre du logement selectionné en suivant: d'abord, l'ordre de la semaine et a l'interieur de celle-ci, l'ordre de serie du logement. Ainsi, le 1o. logement dans la 1o. semaine aurait le numero d'ordre 0! et le dernier logement dans la dernière semaine aurait le numero 10.) |
| 1 | 19-Oct-86 | _____ | |
| 1 | 19-Oct-86 | _____ | |
| 1 | 19-Oct-86 | _____ | |
| 2 | 26-Oct-86 | _____ | |
| 2 | 26-Oct-86 | _____ | |
| 3 | 02-Nov-86 | _____ | |
| 3 | 02-Nov-86 | _____ | |
| 3 | 02-Nov-86 | _____ | |
| 4 | 12-Oct-86 | _____ | |
| 4 | 12-Oct-86 | _____ | |

Basic Weighting Procedures

In order to maintain the representativeness of the sample and obtain unbiased survey estimates, a weight or expansion factor was assigned to each sample household record on the data file; all the data from each record can then be multiplied by the corresponding final weight before summarizing to higher levels at the tabulation stage.

The final weight for each household was calculated by multiplying its basic sampling weight by the appropriate adjustment factor(s) used to compensate for nonresponse and for changes in the status of HUs between the listing and the interview operations. (Weight adjustments are discussed in the next section.) A Lotus 1-2-3 worksheet (EBCM-5) was used to calculate initial and partially-adjusted weights for each SDE in the survey. An example of this worksheet is included in Figure 7 for illustration purposes.

The basic sampling weight, as determined by the sample design, was calculated as the reciprocal of the final probability of selection, which is the product of the probabilities of selection at each sampling stage. Given that the SDEs were selected with probability proportional to size within each substratum, and that a fixed number of HUs (10) was selected within each sample SDE, the probability of selection of the households within each SDE is:

$$P_{hi} = \frac{n_h M_{hi}}{M_h} \times \frac{10}{M'_{hi}}$$

where,

P_{hi} = probability of selection of each sample household the i -th sample SDE of substratum h

n_h , M_{hi} and M_h are as defined in the previous section

M'_{hi} = total number of eligible HUs (households) listed in the i -th sample SDE of substratum h

In actuality, 15 sample HUs were selected within each sample SDE, with a random subsample of 5 HUs set aside as possible substitutes and the complementary subsample of 10 HUs designated to be interviewed. It can be seen from the following equation that the resulting probability of selection is equivalent to that based on a direct systematic selection of the 10 HUs:

$$P_{hi} = \frac{n_h M_{hi}}{M_h} \times \frac{15}{M'_{hi}} \times \frac{10}{15} = \frac{n_h M_{hi}}{M_h} \times \frac{10}{M'_{hi}}$$

Figure 7

| A L'USAGE DE L'INFORMATIQUE | | A L'USAGE DES STATISTICIENS | | | | | | |
|------------------------------|--|----------------------------------|--------------------------------------|------------------------------|-----------------------|--------------------------|------------------------|--------------------------|
| (1) | (2) | (3) | (4) | (5) | (6) | (7) | (8) | (9) |
| No. d'Ordre SDE Selectionnee | Ponderation pour Extrapolation (7*8*9) | Nombre SDEs Choisies Sous-strate | Nombre Menages Sous-strate (Rec. B2) | Nombre Menages SDE (Rec. B2) | Nombre Log. Elig. SDE | Ponderation Initiale SDE | Ajustement Non-Reponse | Ajustement Non-Eligibles |
| 01 01 19 19 | 75.38 | 24 | 20,915 | 348 | 301 | 75.38 | 1.00 | 1.00 |
| 01 02 19 32 | 83.16 | 24 | 20,915 | 358 | 334 | 83.16 | 1.00 | 1.00 |
| 01 03 11 06 | 74.72 | 6 | 4,876 | 459 | 422 | 74.72 | 1.00 | 1.00 |
| 01 04 21 07 | 645.78 | 8 | 39,325 | 204 | 268 | 645.78 | 1.00 | 1.00 |
| 01 05 22 20 | 258.66 | 18 | 85,290 | 85 | 58 | 323.32 | 1.00 | 0.80 |
| 01 06 39 21 | 161.20 | 31 | 36,481 | 500 | 761 | 179.11 | 1.00 | 0.90 |
| 01 07 39 34 | 119.17 | 31 | 36,481 | 317 | 321 | 119.17 | 1.00 | 1.00 |
| 01 08 33 08 | 167.67 | 3 | 4,297 | 422 | 494 | 167.67 | 1.00 | 1.50 |
| 01 09 41 04 | 906.39 | 10 | 106,661 | 233 | 220 | 1007.18 | 1.00 | 0.90 |
| 01 10 42 17 | 1185.82 | 16 | 178,162 | 77 | 82 | 1185.82 | 1.00 | 1.00 |
| 01 11 59 09 | 33.19 | 39 | 16,091 | 348 | 311 | 36.87 | 1.00 | 0.90 |
| 02 13 59 36 * | 55.01 | 39 | 16,091 | 258 | 344 | 55.01 | 1.00 | 1.00 |
| 01 13 59 35 | 40.10 | 39 | 16,091 | 213 | 207 | 40.10 | 1.00 | 1.00 |
| 01 14 62 13 | 744.64 | 17 | 166,227 | 312 | 297 | 930.80 | 1.00 | 0.80 |
| 03 15 62 15 * | 585.55 | 17 | 166,227 | 172 | 103 | 585.55 | 1.00 | 1.00 |
| 01 16 79 13 | 78.92 | 28 | 16,308 | 138 | 187 | 78.92 | 1.00 | 1.00 |
| 01 17 79 26 | 124.90 | 28 | 16,308 | 346 | 742 | 124.90 | 1.00 | 1.00 |
| 01 18 79 39 | 55.39 | 28 | 16,308 | 194 | 205 | 61.55 | 1.00 | 0.90 |
| 04 19 81 09 * | 675.34 | 9 | 64,372 | 233 | 220 | 675.34 | 1.00 | 1.00 |
| 01 20 82 19 | 692.06 | 17 | 125,626 | 126 | 118 | 692.06 | 1.00 | 1.00 |
| 01 21 91 03 | 457.14 | 30 | 83,967 | 362 | 473 | 365.71 | 1.25 | 1.00 |
| 01 22 91 16 | 316.94 | 30 | 83,967 | 491 | 556 | 316.94 | 1.00 | 1.00 |
| 01 23 91 29 | 286.01 | 30 | 83,967 | 320 | 327 | 286.01 | 1.00 | 1.00 |
| 01 24 92 42 | 357.93 | 15 | 41,657 | 287 | 411 | 357.70 | 1.00 | 0.90 |

* = substitution

Ponderation Initiale
donnee par:

$$w_{hi} = \frac{M_h}{m_h * M_{hi}} * \frac{M_{hi}}{10}$$

The basic sampling weight for each sample HU is the reciprocal of its probability of selection, expressed as follows:

$$w_{hi} = \frac{1}{p_{hi}} = \frac{M_h M'_{hi}}{10 n_h M_{hi}}$$

It can be seen that the sample is approximately self-weighting within each substratum. Since the weights within a substratum vary by a factor of M'_{hi}/M_{hi} , their variability depends on how well the measure of size for each SDE approximated the actual number of households listed. Since the sample within each stratum was allocated proportionately among the substrata, the sample is also approximately self-weighting within each stratum. However, the weights vary considerably among strata, as the sample was allocated almost equally among the strata (which differ considerably in size).

Weight Adjustment Factors

In order to maintain an exact workload of 10 interviews in each SDE, a random substitute was interviewed whenever a sample household could not be interviewed. This procedure automatically adjusted the weight for non-interviews although it did not eliminate non-interview bias. Still, when the substitute units were exhausted, it was necessary to proceed with the non-interview weight adjustment since the achieved sample had fewer than 10 households. Likewise other weight adjustments were necessary because of certain changes in the status of HUs between the time of the listing and the interview. Final adjustments to the weights were transmitted to the Data Processing Unit by means of the "Transmission of Corrections Form" (EBCM-INF03). The weight adjustments are discussed below.

1. Non-Response Adjustments

Use of the base sampling weight alone would presume the existence of 10 completed interviews. Yet in certain cases, the number of completed interviews was less than 10 despite the availability of a maximum of 5 replacement units. In order to diminish the bias resulting from missing data from a sector of the population, the weights for the responding units were inflated to account for non-respondents; this adjustment was performed within SDE where a certain level of homogeneity with regard to socioeconomic variables can be assumed. The non-response adjustment factor was:

$$NRAF_{hi} = \frac{10}{r_{hi}}$$

where,

r_{hi} = number of completed interviews in i-th SDE of substratum h

$r_{hi} \geq 10$

This procedure assumed that the responding and non-responding households did not, on the average, differ significantly with respect to socio-economic characteristics and that the non-response rate within the SDE was small. To the extent these assumptions were false, an unquantified amount of bias was introduced in the data. Nevertheless, to control this bias, a constraint was imposed so that if the NRAF exceeded 2.00, that is to say, if the SDE response rate was less than 50%, the non-response adjustment was to be performed at the substratum rather than SDE level, collapsing together the SDEs in the substratum. Within the achieved SDEs, the response rate was never less than 70%. However, there were four SDEs in which no interviews were obtained. For these, it will be necessary to collapse within substratum.

2. Adjustments Due to Change in Occupancy Status

Given that the HU selection was based on a listing conducted 1 or 2 months before the interview, the status of some units changed during this interim period. As the field procedures required maintaining a workload of exactly 10 interviews in each sample SDE, certain types of changes in occupancy status came to require an adjustment to the base sampling weight. Likewise, weight adjustments were required when, at the interviewing stage, a one-to-one correspondence between the selected unit and the reporting unit failed to hold. This correspondence had been established at the listing stage by the definition of the HU (logement) used in the HECS: space occupied by one household.

At the time of interviewing, there were five possible changes related to occupancy status:

- the HU was occupied by a new household;
 - the HU was occupied, but the household members were temporarily absent (or refused to cooperate);
 - the HU had become ineligible, i.e., vacant, seasonal, destroyed, or non-residential;
 - there was a consolidation of two or more HUs;
 - an HU was divided into two or more.
- a. In the case of a new household, the interviewer completed the questionnaire for the new household. No adjustments were necessary since the probability of selection did not change.

- b. In the case of an absent (or refusing) household, the interviewer completed the questionnaire for a substitute household. No adjustments were necessary for this case either.
- c. In the case of an ineligible unit, the interviewer, following field procedures, replaced it with a substitute. However, given that an ineligible HU theoretically should not be included in the sample, it was necessary to deflate the weight by multiplying all household weights within the SDE by the following factor:

$$F_{2hi} = \frac{10 - d_{hi}}{10}$$

where,

F_{2hi} = weight adjustment factor for all interviewed households in the i-th sample SDE of substratum h containing one or more ineligible sample HUs

d_{hi} = number of ineligible sample HUs in the i-th SDE of substratum h

- d. In the case of consolidation, there were three possibilities. To illustrate, let us suppose that there were two HUs, HU-01 and HU-02, in a given housing structure and that HU-01 was selected. At the time of the interview, the interviewer finds out that HU-01 and HU-02 now form a single HU, that is, they are now occupied by a single household.

If the household originally from HU-01 is the one now occupying the consolidated space, the interviewer simply proceeded with the interview. No weight adjustments were required for this case.

If instead it was the HU-02 household which was now occupying the consolidated space, then the interviewer considered the selected unit HU-01 as no longer existing, i. e. ineligible, and replaced it with a substitute. In this case, the F_{2hi} adjustment factor previously described was applied.

If a third (new) household moved into the consolidated space, the interviewer went ahead with the interview since he could not distinguish the first case from the second case. In this case, the weight for the interviewed household was multiplied by 1/2, to correct for the fact that its probability of being interviewed was doubled. The adjustment factor for this case which was applied individually to the affected household was:

$$F_{3hij} = \frac{1}{m''_{hij}}$$

where:

F_{3hij} = weight adjustment factor for the j-th interviewed household in the i-th SDE of substratum h with indistinguishable predecessor units

m''_{hij} = number of predecessor HUs that were listed separately for the now consolidated unit

- e. When an HU was split into two or more households, the interviewer selected one at random for interviewing. Since this action modified the probability of selection for the household, the weight for the responding household was multiplied by the following adjustment factor:

$$F_{1hij} = m'_{hij}$$

where:

F_{1hij} = weight adjustment factor for the j-th sample HU in the i-th SDE of substratum h having more than one household

m'_{hij} = number of households found in the j-th HU.

Estimation Formulas

Most of the estimates from the survey data are in the form of totals and ratios (including means and proportions). The survey estimate of a total of a particular variable X for a certain group A would be calculated as follows:

$$\hat{X}_A = \sum_{j \in A} W'_j X_j$$

where:

\hat{X}_A = estimated total of variable X for group A
 A = subset of records belonging to group A
 W'_j = final weight for the j-th record
 X_j = value of variable X for the j-th record

The survey estimate of a ratio for group A is defined as \hat{Y}_A / \hat{X}_A , where \hat{Y}_A and \hat{X}_A are corresponding total estimates. Means and proportions are special types of ratios. In the case of a mean, the variable X in the ratio is equal to 1 for each record in group A, so that the denominator equals the sum of the weights for group A. In the case of a proportion,

the variable Y in the numerator of the ratio is either 1 or 0 depending on the presence or absence of a specified characteristic.

Other Estimation Considerations

Since different reference periods were used in the questionnaire for distinct types of expenditures, it is important to standardize the reference period at the tabulation stage when combining expenses from the various sections. For tables specifying yearly expenses, the following factors are to be used in converting other reference periods to a year:

| Reference Period | Conversion Factor |
|---------------------|----------------------|
| ----- | ----- |
| 1 week | 52 |
| 2 weeks | 26 |
| 1 month | 12 |
| 1 quarter | 4 |

The daily expenditures on food items are to be summed across the four visits during the week before applying the conversion factor of 52 to obtain yearly expenditures.

Calculation of Variances

It is important to estimate sampling errors for the principal survey estimates in order to determine the corresponding level of precision. For this purpose, a representative group of different types of survey estimates at various levels of disaggregation was selected from the preliminary survey tabulation plan.

Given the complexity of developing customized computer programs to calculate variances and the large amount of time this would require, it was decided to use an existing generalized variance software package called SUPER CARP (Cluster Analysis and Regression Program), developed by Iowa State University. This package offers a variance estimator which takes into account the two-stage sample design used by the HECS. It provides for the calculation of variances for estimates of totals, means, proportions and other ratios. SUPER CARP requires that separate data files be created: a file for variables related to household characteristics, with one record per questionnaire, and a file for variables related to person characteristics, with one record per person within a questionnaire.

The package SUPER CARP uses an ultimate cluster type of variance estimator based on the squared differences between weighted segment totals. The variance estimator for ratios is based on a Taylor series expansion. The variance formulas are presented in the SUPER CARP manual in the form of matrices. The following formulas, presented in a simpler form, are used by SUPER CARP to calculate the variance of totals and ratios (without a finite population correction factor (fpc)).

(1) Variance of a Total Estimate (X)

$$\text{Var}(\hat{\lambda}) = \sum_h \frac{n_h}{n_h - 1} \sum_{i=1}^{n_h} \left(X_{hi} - \frac{\hat{X}_h}{n_h} \right)^2$$

where:

- n_h = number of SDEs selected in substratum h
- \hat{X}_{hi} = $\sum_j W_{hij} X_{hij}$ = weighted total of variable X for the i-th sample SDE in substratum h
- W_{hij} = final weight for the j-th sample HU in the i-th sample SDE of substratum h
- X_{hij} = value of variable X for the j-th sample HU in the i-th sample SDE of substratum h
- \hat{X}_h = $\sum_{i=1}^{n_h} \hat{X}_{hi}$ = estimated total of X for substratum h
- X = $\sum_h \hat{X}_h$ = estimated total of X for all strata

(2) Variance of a Ratio Estimate (\hat{Y}/\hat{X})

$$\begin{aligned} \text{Var} \left(\frac{\hat{Y}}{\hat{X}} \right) = & \sum_h \frac{n_h}{(n_h - 1)} \frac{1}{(\hat{X})^2} \sum_{i=1}^{n_h} \left[\left(\hat{Y}_{hi} - \frac{\hat{Y}_h}{n_h} \right)^2 + \left(\frac{\hat{Y}}{\hat{X}} \right)^2 \left(\hat{X}_{hi} - \frac{\hat{X}_h}{n_h} \right)^2 \right. \\ & \left. - 2 \left(\frac{\hat{Y}}{\hat{X}} \right) \left(\hat{Y}_{hi} - \frac{\hat{Y}_h}{n_h} \right) \left(\hat{X}_{hi} - \frac{\hat{X}_h}{n_h} \right) \right] \end{aligned}$$

where \hat{Y} , \hat{Y}_h , \hat{Y}_{hi} and Y_{hij} are defined for variable Y in the same manner as \hat{X} , \hat{X}_h , \hat{X}_{hi} and X_{hij} , respectively.

Even though SUPERCARP has an fpc option, it is based on the assumption that the segments are selected with equal probability. However, because the sample SDEs for the HECS were selected with probability proportional to size, this option was not recommended. It was judged that the fpc factor would only have a small effect on reducing the variance estimates, since the overall sampling rate was relatively small.

CHAPTER VII

EFFECT OF POLITICAL SITUATION ON THE SURVEY

The HECS took place during a period of extreme political instability in Haiti. This factor was more or less constant and caused several delays for the survey, created a hostile climate for fieldwork, and finally forced the data collection to close 8 weeks before its scheduled termination. It further contributed to the termination of U. S. economic assistance to the GOH which, in turn, resulted in the cancellation of financial and technical assistance for the survey before final results could be obtained.

Interaction With Survey Activities

The preparation period for the HECS, in the fall of 1985, coincided with the agitation period whose events led ultimately to the collapse of Haiti's 29-year dictatorship (headed by president-for-life, Jean-Claude Duvalier) in February 1986. The HECS pilot survey originally scheduled for January 1986 had to be postponed until May 1986. A new (provisional) government had taken office in February. Still, unrest continued as manifestations and protests were staged against the new government.

As a result of hostility and unreceptiveness on the part of local populations, tied to the political climate, several enumeration areas had to be replaced in the sample. Throughout the survey, instances of unreceptive local residents threatening and sometimes ousting the field enumerators were experienced. The publicity campaign was stepped up but there were still areas where neither the campaign nor the efforts of the interviewers, supervisors and local authorities were sufficient to overcome the difficulties experienced. Public relation visits from the central office staff were not always possible or effective. The areas involved were usually remote rural zones where normal means of communication (radio, television, press) were not accessible. Other impossible-to-work-in areas included the lower socio-economic peripheral areas of Port-au-Prince where hostility was keen. The above-mentioned areas were the ones which, for the most part, were replaced in the sample.

Throughout the survey period, general strikes were frequent, particularly in Port-au-Prince. Their scale and intensity was usually high, with a fair degree of violence; they usually paralyzed public transportation and created threats to personal safety. As a result, work in the central office was often interrupted, employee absenteeism at IHSI was high (often as high as 50%) and many work days were lost. These strikes brought about interruptions in the survey technical and computer processing activities.

The worst effect of the political situation on the HECS was the closing of the survey before its completion. The IHSI, feeling that the situation was too uncertain and too dangerous for the safety of field enumerators,

terminated the data collection on September 20, 1987, after the 11th survey month. Still, work continued in the central office on data processing and technical activities. In December 1987, however, the suspension of U. S. non-humanitarian aid to Haiti in relation with the failed November 29 elections resulted in the cancellation of all financial and technical support to GOH projects backed by USAID including the HECS. The data processing and analysis phases of the survey were thus discontinued at IHSI.

Effect on Data Obtained

The exact effect of the political situation on the data collected is not known. Nonetheless, it can be expected that expenditure and consumption patterns measured will reflect the actual social, economic and political conditions that were present during the survey period. Analysts will want to study those factors in conjunction with the survey data. An example of the relation between the overall climate and the variables being measured is the following. In the Port-au-Prince area, when strikes were anticipated, many people would buy ahead and stock large amounts of food because during the strikes, they could not go out and make purchases. Whether this is true only of the Port-au-Prince area or whether the effect was more widespread is not known with certainty.

In addition, there is a potential for bias in the estimates resulting from the replacement of enumeration areas which could not be surveyed. Even though the replacement enumeration areas were taken from the same socioeconomic substratum as the original ones, it is not known whether differences in expenditure and consumption patterns existed.

CHAPTER VIII

RECOMMENDATIONS FOR FUTURE SURVEYS

This chapter presents recommendations for the future planning and implementation of similar surveys in Haiti, based on the experiences with the HECS. Their feasibility will depend on the circumstances involving each particular survey and on the resources available.

Recurrence of Surveys

Surveys in Haiti have been taken on an ad hoc basis. This approach is not as effective as continuous data collection for the reason that ad hoc surveys require an almost fresh start every time. The accomplishments of the HECS provide a foundation upon which a continuous household survey program can be built. Therefore, it is recommended that a socio-economic household survey be repeated every year at the IHSI in order to maximize the utility of the investment already made.

A continuous survey program is more economical in the long run because the survey infra-structure can be used repeatedly to address many needs; materials and other resources can be maintained over the time. A regular survey program also offers the possibility of gradually developing an adequate base of trained and experienced personnel. Moreover, it offers regularity in the influx of basic data as well as a capacity for monitoring change over time. When special data need to be generated, adjustments can be incorporated into the survey program without having to lay new grounds.

Finally, a continuous survey program would also offer advantages in timeliness. Since the design of a quality survey demands time for the preparation of different activities (questionnaire design, sample design, reference manuals, field organization, training, data processing system), it would be reasonable to expect that, in the context of an appropriate infra-structure, the time required for the preparing, implementing and processing a particular survey would be minimized.

Survey Management

It is important that, as for the HECS, a technical committee be created expressly to oversee survey operations. However, it is also crucial that arrangements be made for this staff to be relieved of other duties so that they may spend sufficient time on survey activities. In addition, it would be useful, for accountability purposes, to sort technical and administrative responsibilities and assign them to specific persons.

For long-term benefits, a permanent survey staff could be constituted to implement the continuous household survey program discussed above. It would also be beneficial for the survey program if survey personnel and activities were made a part of the IHSI operational budget. The ad-hoc budget allocated to the HECS received low priority and required constant item-by-item re-negotiation to survive. A program budget would improve the timeliness of the disbursements and provide coverage for unexpected expenses not itemized in the budget.

Questionnaire Design

Even though the HECS questionnaire allowed for the collection of many types of data in one survey, the use of lengthy questionnaires should be avoided. Because long questionnaires create a greater response burden for the respondent and increase the overall workload for the survey, they have a greater potential for compromising the quality of the survey data.

In a continuous survey program, information can be collected effectively on many subjects without overburdening a particular survey. This is accomplished through the use of a short, core questionnaire covering basic household characteristics such as demographic variables and employment, supplementing the core questionnaire periodically with specialized modules covering other topics.

Data Processing

It is recommended that resources be invested in creating a permanent capability for efficient electronic survey processing. Ideally, the same programs could be used to process new survey data with a minimum of changes. One of the most important considerations should be to increase rapidity in the computer processing in such a way that tabulated data can be obtained very soon after field collection.

Sample Design

The sample design used in the HECS can be useful to many other socio-economic surveys. For this reason, it is recommended that the sampling frame of census enumeration areas (SDEs), as stratified for the HECS, be maintained on a computer file to serve as a master (first-stage) sampling frame. Second-stage sampling frames of housing units can also be maintained up-to-date through periodic mapping and listing of the 312 SDEs selected for the HECS or a subsample thereof, an operation which would require a permanent trained staff.

The implementation of a master sampling frame should lower the expenses involved in selecting a sample for a particular survey by distributing costs over several surveys. It should, in addition, improve the quality of the lists and maps by providing more time and resources for enumeration and validation.