

PN-ARN-641
81528

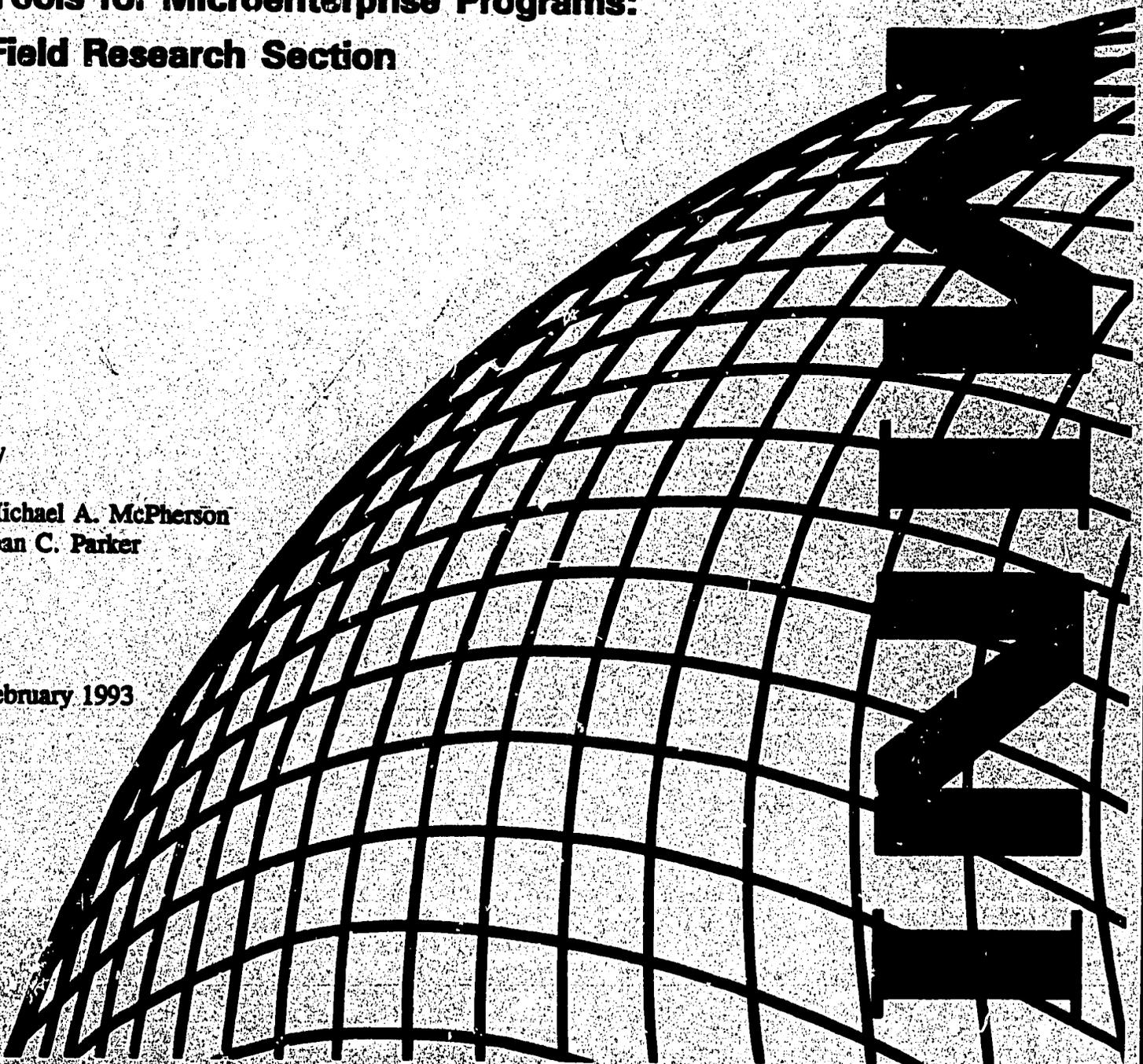
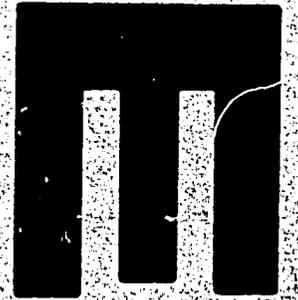
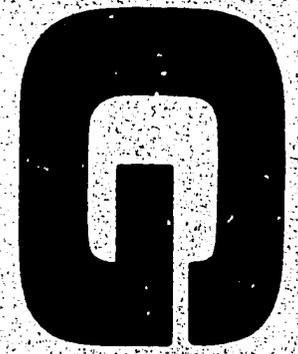
A MANUAL FOR CONDUCTING BASELINE SURVEYS OF MICRO- AND SMALL-SCALE ENTERPRISES

**Tools for Microenterprise Programs:
Field Research Section**

by

**Michael A. McPherson
Joan C. Parker**

February 1993



PREFACE

As the field of microenterprise development becomes more sophisticated, practitioners require greater technical skills to meet increasingly complex challenges. As part of its continuing efforts to meet the needs of microenterprise practitioners, GEMINI is producing a series of technical notes and training materials on topics related to microenterprise development.

The GEMINI technical note series is divided into three sections. The microenterprise finance section includes notes on setting interest rates, the expansion of microenterprise finance programs, and other specific topics. The nonfinancial assistance section contains notes on how to carry out subsector analyses and identify cost-effective ways of enhancing the productivity and competitiveness of microenterprises. The microenterprise research section contains notes on survey methodologies designed to increase understanding of the growth and dynamics of the microenterprise sector. Each note can be used on its own; however, those interested in one note within a section will probably be interested in all of the notes produced in that section.

The notes can be used as reference materials and can also serve as the basis for creating customized training materials for individual organizations. Although these "Tools for Microenterprise Programs" are specifically designed for practitioners, they should also prove useful for donors, evaluators, and others interested in microenterprise development.

For a complete listing of other technical notes and training materials available from GEMINI, as well as a list of all GEMINI Technical Reports and Working Papers, please see the publication list at the end of this document.

The GEMINI Project is the U.S. Agency for International Development's primary technical resource in the field of micro- and small-scale enterprise development. GEMINI explores the latest in microenterprise development and brings new findings to the field through direct work with A.I.D. missions, U.S.-based private voluntary organizations, and local organizations in developing countries. GEMINI offers technical assistance, training, economic research, and information to A.I.D., implementing organizations, resource institutions, national governments, and other practitioners involved with microenterprise development. The project aims to have a catalytic effect on a broad spectrum of efforts to promote the growth of micro- and small-scale enterprises.

ACKNOWLEDGMENTS

The authors gratefully acknowledge the helpful comments of Carl Liedholm, Donald Mead, Lisa Daniels, and Matthew Gamser. In addition, we recognize appreciatively the many discussions over the years with Yacob Fisseha. Finally, we thank the many people who participated in the baseline surveys with which the authors were involved.

TABLE OF CONTENTS

INTRODUCTION	
WHAT IS THE MSE BASELINE SURVEY?	1
WHEN IS A BASELINE SURVEY USEFUL?	1
PURPOSE AND STRUCTURE OF THE MANUAL	3
CHAPTER ONE	
SURVEY MANAGEMENT AND PERSONNEL	5
WHAT NEEDS TO BE DONE?	5
Presurvey Activities	5
Team Training	7
Fieldwork, Coding, and Data Entry	7
Data Cleaning	8
Data Analysis	8
Report Writing	8
WHO DOES WHAT?	8
Enumerators	9
Supervisors	9
Coder	9
Data Enterer	11
Survey Administrator	11
Principal Researcher	12
SELECTION OF PERSONNEL	12
MANAGING THE SURVEY: THE ROLE OF THE	
PRINCIPAL RESEARCHER	13
SURVEY EQUIPMENT	16
BUDGETARY CONSIDERATIONS	17
SURVEY MANAGEMENT: KEY POINTS	19
CHAPTER TWO	
TRAINING	21
BASIC TRAINING	21
Preparation for Basic Training	21
Who Attends Basic Training	22
Illustrative Schedule of Basic Training	23
SUPERVISOR TRAINING OPTIONS	25
SUPPLEMENTAL CODER TRAINING	26

DATA ENTERER TRAINING	27
POST-TRAINING TROUBLE-SHOOTING SESSIONS	28
Trouble-shooting With the Field Teams	28
Trouble-shooting With the Coder	28
Trouble-shooting With the Data Enterer	29
TRAINING: KEY POINTS	29
CHAPTER THREE	
SAMPLING	31
CHOOSING THE SAMPLING METHOD	31
CHOOSING STRATA	32
CHOOSING CLUSTERS	33
SAMPLING: KEY POINTS	38
CHAPTER FOUR	
INSTRUMENT DESIGN	41
DETERMINING THE SURVEY'S CONTENT	41
THE BASIC QUESTIONNAIRE	42
Structure of the Basic Questionnaire	42
Content of the Basic Questionnaire	44
Coding	45
Conventions for Completing the Questionnaire	46
Choosing the Unit of Analysis	49
EXTENSIONS OF THE BASELINE SURVEY	50
The Supplementary Questionnaire	50
The Closed Enterprise Questionnaire	53
INSTRUMENT DESIGN: KEY POINTS	55
CHAPTER FIVE	
DATA MANAGEMENT AND ANALYSIS	57
THE PAPER TRAIL	57
DOCUMENTING AND MANAGING COMPUTER FILES	59
CLEANING DATA	60
ANALYZING THE DATA	61
WEIGHTING DATA FOR COUNTRYWIDE EXTRAPOLATION	61
DATA MANAGEMENT: KEY POINTS	63

CHAPTER SIX	
DISSEMINATION OF RESULTS	65
WRITTEN PRODUCTS	65
MAKING THE DATA AVAILABLE	67
OTHER USES FOR BASELINE SURVEY DATA	67
RESULTS: KEY POINTS	68
CONCLUSIONS	71
REFERENCES	73
APPENDIX 1: THE HISTORY OF BASELINE SURVEYS	75
APPENDIX 2: ILLUSTRATIVE BUDGET FOR A COUNTRYWIDE BASELINE SURVEY	77
APPENDIX 3: CONTENT OF BASIC TRAINING	81
APPENDIX 4: ILLUSTRATIVE TRAINING EXERCISE FOR THE BASELINE SURVEY	87
APPENDIX 5: COLUMN EXPLANATIONS FOR BASIC QUESTIONNAIRE	89
APPENDIX 6: COLUMN EXPLANATIONS FOR CLOSED ENTERPRISE QUESTIONNAIRE	107
APPENDIX 7: ILLUSTRATIVE BUSINESS CODE SHEET	117
APPENDIX 8: INTERNATIONAL STANDARD INDUSTRIAL CLASSIFICATION CODES	119
APPENDIX 9: INFORMATION ABOUT HOUSEHOLDS	123
APPENDIX 10: FORMULAE FOR SAMPLE SIZE	125

LIST OF FIGURES AND TABLES

Figure 1	Survey Time Line	6
Figure 2	Illustrative Basic Questionnaire	43
Figure 3	Illustrative Supplementary Questionnaire	51
Figure 4	Illustrative Closed Enterprise Questionnaire	54
Figure 5	The Paper Trail	58
Table 1	Proportion of Sample Involved in MSE Activity	62

INTRODUCTION

WHAT IS THE MSE BASELINE SURVEY?

The Micro- and Small-Scale Enterprise (MSE) Baseline Survey is a door-to-door enumeration of nonagricultural income-earning activities. The survey counts all activities in manufacturing, trade, and services, regardless of size, location, or legal status. In so doing, it uncovers the smallest and most invisible part of the private sector, the MSE sector.¹ The survey not only counts activities, but it also describes them in static and dynamic terms. In short, it provides a detailed and comprehensive picture of the MSE sector, a sector too often missed by conventional survey methods.²

The MSE Baseline Survey counts all activities in manufacturing, trade, and services.

Low costs and fast results (usually no more than 18 weeks from presurvey activities to the completion of the final report) complement the baseline survey's comprehensiveness. The survey is strikingly simple by design. A short, precoded questionnaire allows easy recording of interviews and immediate data entry. The standardized approach allows fast survey start-up, quick results, and comparability of results between surveys.

WHEN IS A BASELINE SURVEY USEFUL?

Because the MSE Baseline Survey is known for its broad coverage and its ability to gather information on even the most invisible businesses, it has become a popular instrument to meet a variety of goals. The most frequently invoked reason for doing a baseline survey is to provide general information for developing a countrywide MSE strategy or for designing targeted programs or projects. The baseline survey is well suited to provide such general information, but it is not well equipped to provide other pieces of information (Box 1). The information the survey does not provide includes flow variables, such as profits, sales, or costs, and the contribution of the MSE to family income. Requests for the kinds of information that can be provided by a baseline survey have come from national governments, international donor agencies such as the U.S. Agency for International Development, and organizations that provide small enterprise assistance.

The survey is known for its broad coverage and its ability to gather information on even the most invisible businesses.

On occasion, a baseline survey is conducted after programs or projects have been in place for some time, often because those efforts are not meeting their stated objectives. In such cases, the baseline survey can be a diagnostic tool in uncovering more information about the target population than was previously available. In some cases, the survey reveals that the structure of the MSE sector is different than had been assumed. In other cases, it highlights where the project has or has not been effective at reaching given segment of the target population.

First implemented in 1974, the MSE Baseline Survey has evolved over time in goals and content. A brief history of its evolution as a method is contained in Appendix 1.

BOX 1

WHAT CAN THE BASELINE SURVEY DO?

The baseline survey is ideal for gathering information regarding:

The baseline survey is not well suited for gathering information regarding:

Numbers of MSEs in a region or a country;

Product or service prices;

Location of MSEs;

Input costs or sources;

Types of activities in which MSEs are engaged;

Contribution to family income;

Number of workers in the MSE;

Characteristics of the household;

Types of workers in the MSE;

Specific constraints and problems;

Obvious proprietor characteristics, such as gender;

Product marketing; and

Sales, profits, and value added;

Seasonality of MSEs.

Growth of MSEs in terms of workers; and

Age profiles of MSEs.

PURPOSE AND STRUCTURE OF THE MANUAL

This manual is a practical guide on the way to conduct an MSE Baseline Survey using the methods designed by Michigan State University. It is designed to be used primarily by practitioners interested in learning how to carry out such a survey, but may also be helpful for decision makers when deciding whether a baseline survey can provide them with the information they need and, if so, how to approach the task.

The manual is designed for practitioners interested in learning how to carry out a baseline survey.

The manual is divided into six chapters, each of which covers an important aspect of survey preparation, implementation, and follow-up. The chapters are:

- Chapter One: Survey Management and Personnel
- Chapter Two: Training
- Chapter Three: Sampling
- Chapter Four: Instrument Design
- Chapter Five: Data Management and Analysis
- Chapter Six: Dissemination of Results

Each chapter is somewhat self-contained. Readers with experience in conducting other sorts of surveys may wish to look quickly at Chapters One and Two, and concentrate their attention on Chapters Three and Four. Each chapter concludes with a list of key points to remember when carrying out the survey. The 10 appendices contain technical information and examples of survey documents. Notes are found at the end of each chapter.

Notes

1. For the purposes of these surveys, an MSE is defined as any income-earning activity with 1-50 workers. Primary industries, such as agriculture and mining, are excluded from the definition as are activities in which more than 50 percent of the output is not marketed, but consumed instead within the household.
2. The MSE Baseline Survey counts businesses of all sizes, even those larger than defined by "micro" and "small." However, the vast majority of businesses uncovered in such surveys in developing countries fall within the micro category of 1-10 workers or the

small category of 11-50 workers, and only a handful fall within the medium- and large-size category of more than 50 workers. Although these larger businesses are included in the size profile of the business sector, the survey's focus is on the nature of the micro- and small-scale enterprise population.

CHAPTER ONE

SURVEY MANAGEMENT AND PERSONNEL

To the uninitiated, the responsibility of designing and implementing a baseline survey can seem bewilderingly complicated and onerous. The process by which such surveys are brought to fruition does contain a large number of complex tasks. This chapter is intended to alleviate some of these worries by outlining exactly what these tasks are, how long each takes to complete, and the personnel and equipment necessary for their successful completion.

WHAT NEEDS TO BE DONE?

A typical baseline survey — from planning through to the final report — requires 14-18 weeks from the start of the actual preparatory work, after the budget has been approved, the institutional set-up agreed upon, and the principal staff selected. Figure 1 is a graphical presentation of the way this time is allocated to survey activities.¹ The specific tasks in the eight categories of activity are outlined below.

A typical survey requires 14-18 weeks from the start of the actual preparatory work.

Presurvey Activities

(begin two weeks before fieldwork starts)

... for developing a research context:

Collect relevant background documents.

... for developing survey structure and content:

Finalize administrative arrangements.

Finalize scope of work with client.

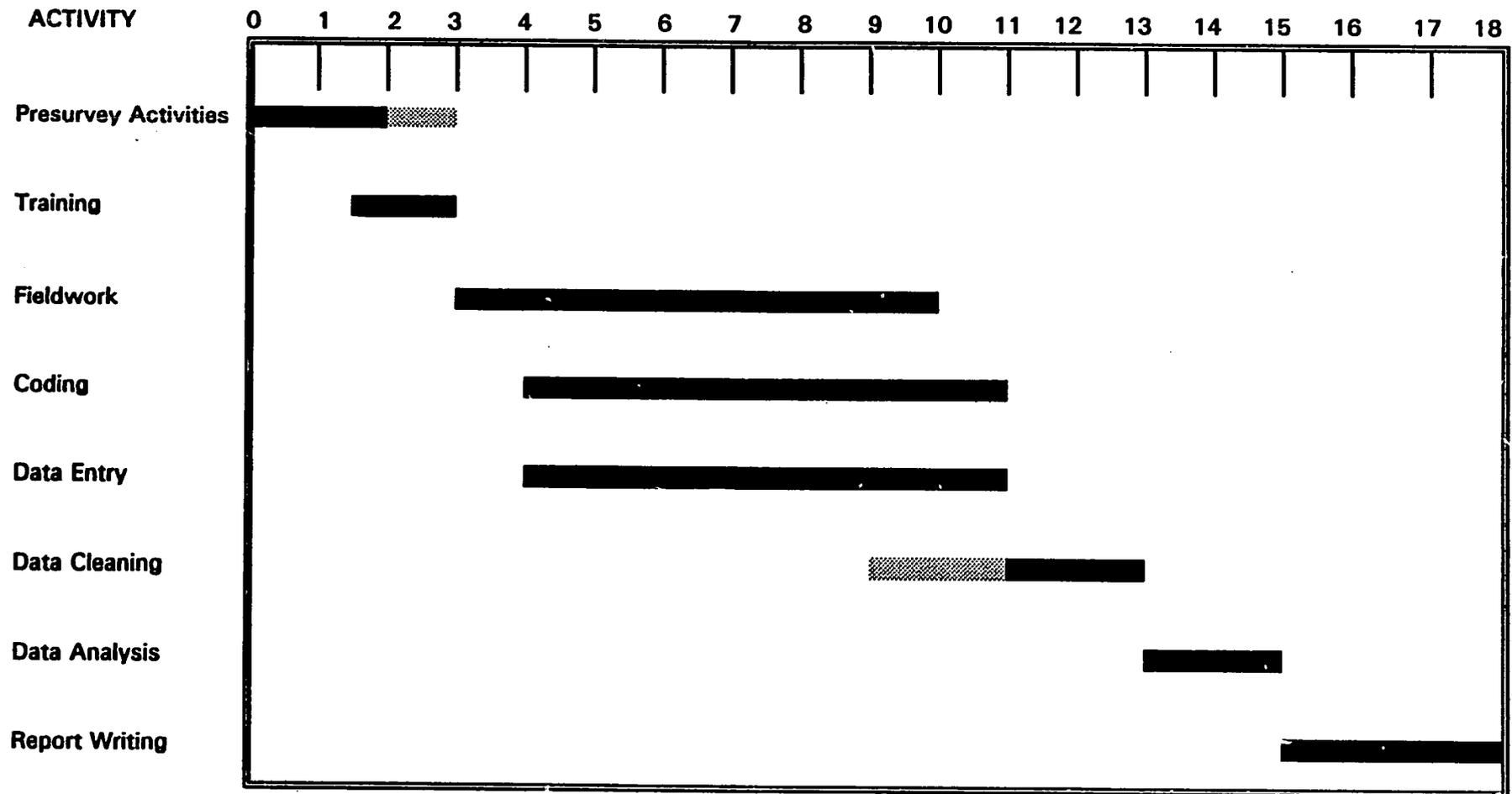
Make contacts at relevant government offices, including the Central Statistics Office.

... for sampling:

Get census lists for the country.

Choose strata to be used for sampling.

FIGURE 1
SURVEY TIME LINE



Generate sampling clusters, if necessary.
Draw random sample of clusters from each strata.
Get maps that identify sampled areas and their boundaries.

... for instrument design:

Finalize questionnaire(s) and codes with input from client.
Pretest questionnaires in the field; revise as necessary.
Reproduce questionnaires.

... for training and fieldwork preparation:

Finalize arrangements for training.
Develop training materials.
Identify and hire enumerators, supervisors, data enterer.
Develop computer entry program for each questionnaire.
Arrange for access to computer in survey headquarters.
Buy field supplies.
Set up fieldwork schedule, including number of field teams and order of survey coverage.
Arrange for vehicles as required by fieldwork schedule.
Set up survey financial system.

Team Training

(begins one week before fieldwork starts)

Train supervisors (2 days).
Train enumerators and coder (3-4 days).
Field training and debriefing of enumerators and supervisor (1 day).
Train data enterer (1/2 day).

Fieldwork, Coding, and Data Entry

Transport teams to sampled areas.
Make contact with appropriate local officials.
Identify boundaries of each sampled area.
Conduct interviews using door-to-door method.
Collect and transport completed questionnaires to survey headquarters.
Check codes.
Enter data into computer.
Verify data in computer.

Data Cleaning

(begins after fieldwork has been under way for a few weeks)

- Run an error search on the data.
- Run consistency checks on the data.

Data Analysis

- Run preliminary analyses on data midway through fieldwork.
- Run final analyses after all data are cleaned.

Report Writing

- Write up results of data analysis.
- Share draft report with clients.
- Finalize report.
- Reproduce and disseminate report.
- Finalize report.
- Reproduce and disseminate report.

Pitfalls abound in the few weeks preceding fieldwork.

As the list illustrates, the most difficult period of the entire survey is often the few weeks preceding fieldwork, when tasks often occur concurrently and when pitfalls abound. This is a good place to give the principal researcher an extra week if resources allow. Whenever possible, tasks should be completed before the principal researcher arrives in the country. Such preliminary tasks might include setting up the administrative arrangements and advertising for personnel. If preliminary planning is good and if a strong team is identified, trained, and fielded, the survey will develop its own momentum once fieldwork starts.

WHO DOES WHAT?

Most surveys divide the labor force into six groups.

The typical MSE countrywide Baseline Survey enumerates several thousand enterprises and, in most cases, no more than 10 weeks are allocated for all fieldwork. To complete the many tasks noted in the previous section, the survey must employ the appropriate number of people and use their talents and energies efficiently. Although there are many variations, most MSE surveys divide the survey labor force into six

groups. Each of these groups is responsible for the completion of one or more of the survey tasks. Keeping in mind that baseline surveys can be flexible, we now consider the responsibilities of each group.

Enumerators

Enumerators are the personnel who actually collect the data from the proprietors of MSEs. Each enumerator is responsible for the interview process: rapport must be established with the respondent, the purpose of questionnaire must be explained, and the questions asked clearly. The enumerator must take down the respondent's answers accurately and completely, and ensure that completed questionnaires are delivered to the supervisor. In addition, the enumerator must cover every household and shop in the area to which he is assigned, and report any difficulties to the supervisor.²

Supervisors

Each team of enumerators is managed by a supervisor. In many ways, the supervisors are the keys to the success or failure of the survey exercise. The supervisor is responsible for getting each team to the correct area each day and for keeping the team working. Another important responsibility of the supervisor is to identify precisely the boundaries of the area to be enumerated. This difficult task is discussed in the example box on the following page. The supervisor also ensures quality control. She makes certain that each and every household and shop in the assigned area is visited and that the interviews are conducted in an appropriate manner. The supervisor ensures that the survey conventions, discussed in detail in Chapter Four, are followed at all times. She also checks all of the questionnaires produced by the team for accuracy and completeness, and passes them on in a timely manner to the coder in the main office. A final important job of the supervisor is to ensure the safety of the team.

The supervisor is the key to the success of the survey.

Coder

The baseline survey questionnaire combines a large amount of information on several enterprises on a single sheet of paper.³ In spite of the vigilance of the enumerators and supervisors, questionnaires that are incomplete or inconsistent will find their way to the main office. It is the responsibility of the coder to check carefully each piece of information on

BOX 2

IDENTIFYING BOUNDARIES: IT'S HARDER THAN IT LOOKS

From the standpoint of statistical accuracy, it is crucial that every household and shop in the areas selected be covered. This means that it is imperative that the boundaries of the enumeration area (EA) be clear to all enumerators. This seems simple enough: how hard could it be to recognize boundaries? As it turns out, it can be very difficult, especially in rural areas. In urban areas, most EAs are defined by streets. With a good-quality street map of the city, it is not difficult to keep the team within the area, and to make sure that the whole area has been visited. This procedure is seldom so simple in rural areas.

In most countries, topographic maps of all rural EAs are available. In some countries, aerial photographs are available as well. However, frequently the boundaries of these areas are natural physical characteristics: rivers, streams, mountains, and hills, for example. Occasionally, man-made characteristics such as dirt roads or fences will be the boundary. In some instances, the boundary will be unmistakable. One enumeration area in Swaziland, for example, bordered on the sizeable Ngwempisi River. In other cases, the landmarks will be much more obscure. Trying to find which dirt road is the boundary can be difficult if it is one among many in an area. Sometimes the creeks that form boundaries have been dry for decades. Although every situation is unique, here are some general tips for identifying boundaries:

1. **Get maps with as much detail as possible.** Acquire maps with at least a 1:50,000 scale. Any less detail will make the EA impossible to identify.
2. **Learn to read the details on the maps.** Make sure the supervisor knows the symbols for bridges, churches, schools, and similar features. Be sure she takes advantage of them. For example, if she knows that the dirt road that forms one part of the boundary runs by a church, she can orient herself by locating the church.
3. **Use the scale on the map.** Every map has a scale. If the odometer in the survey vehicle is operating properly, this scale can be invaluable. For example, if the supervisor knows from the map that the EA begins 6.7 kilometers after the turn-off from the paved road, then the odometer can be used to measure that distance.
4. **Let the local people help.** In almost every case, one of the local people can help your team. The supervisor should ask them questions to establish how near the team is to the EA. Although local residents may not be able to read the map, they almost certainly will know whether the team is near the landmark the supervisor is looking for.

each questionnaire. When problems are found, the coder must discuss them with the supervisor and return the suspect questionnaires to her to be corrected. Once the coder is satisfied with the quality, he assigns a unique identification number to each MSE and entrepreneur, and then passes the questionnaires along to the data enterer.⁴

Data Enterer

Almost every bit of information on the questionnaire in an MSE Baseline Survey is entered into a computer. This makes the analysis of the data fairly simple, and allows the data to be easily distributed to other interested researchers. The task of logging these data into a computer falls on the data enterer. To reduce the incidence of keypunch errors, all data are entered twice. Some data entry programs, such as SPSS/PC+, have such verification routines built into them.

Survey Administrator

One of the key tasks of presurvey planning is careful attention to setting up the administrative arrangements that will support the survey. Typically, one (or more) person is brought on to act as the survey administrator. Sometimes, the survey initiator contracts with an individual for this job, but recent MSE surveys have arranged with local consulting firms to provide this person.

A qualified, full-time person should be assigned as survey administrator.

When choosing an administrative arrangement for the survey, it is important to decide how extensive a role that person or organization will play in the survey. Three levels of administrative support have been tried. At the lowest level of effort, the administrator identifies survey personnel (supervisors, enumerators, coder, and data enterer) and makes initial contacts in key government offices. At the medium level of effort, he is responsible for a wider range of administrative services: personnel contracts, reproducing questionnaires, providing computer access, purchasing supplies, renting vehicles, and managing the budget. And at the highest level of effort, the administrator becomes the key overseer of fieldwork in addition to the tasks listed above. In these cases, while the principal researcher is responsible for sampling, instrument design, training, data cleaning, analysis, and report writing, the administrator is responsible for day-to-day management of the field teams, plus the coding and data entry activities. The more responsibilities that the administrator assumes, the more important it is to ensure that the arrangements are

carefully understood in advance, and that a qualified, full-time person is assigned to the job.

The logistics of keeping a countrywide survey in operation are complicated.

Whether the logistical arrangements and the financial management of the survey are to be done by the administrator, the principal researcher, or someone else, it is imperative that the assignment of tasks be carefully thought out in the design phase of the survey. The logistics of keeping a countrywide survey of any sort in operation are complicated, to say the least. The personnel must be paid in a timely fashion, and, depending on the particular arrangements, the teams and the office personnel may need to be fed lunch each day. When teams travel away from the urban centers, arrangements must be made for sleeping and eating accommodations. In addition, the vehicles used by survey personnel must be serviced and kept fueled. Because the survey involves the disbursement of tens of thousands of dollars, books must be kept and receipts turned in for every purchase.

Principal Researcher

The ultimate responsibility for the success of the survey rests with the principal researcher. The job requires tremendous versatility and hard work at every phase of the survey process. So important and complicated is this job that it will be discussed in length below.

SELECTION OF PERSONNEL

Enumerators should be intelligent, articulate, and personable.

Having discussed the tasks involved in the running of a baseline survey and the personnel responsible for each task, it is important to consider the issues surrounding the selection of the survey team. Specifically, what are the attributes of the personnel in successful baseline surveys, and how many of each sort of worker is required? Experience has led to the conclusion that field teams should consist of between six and eight enumerators and one supervisor. Enumerators should be intelligent, articulate, and personable. It is usually necessary for them to be bilingual, because although training is conducted typically in English, interviews are given in the local language. Because enumerators must cover a lot of ground each day on foot, it is helpful if they are reasonably physically fit. The authors believe that a mix of male and female enumerators works best.

Supervisors, as noted above, are one of the most important parts of the team. It is difficult to have a good survey with bad supervisors (or a bad survey with good supervisors). Above all, supervisors must be strong leaders. This involves strength of personality, dedication to the quality of the survey, and intimate familiarity with the survey instruments. Although some recommend hiring supervisors who are older than the enumerators, selecting supervisors from the pool of enumerators has also worked well in MSE surveys. It is also important that each supervisor have a valid driver's license, because otherwise a driver must be hired for each team, adding expense and reducing space for enumerators in the survey vehicles. If the survey uses drivers, the supervisors are freed from such concerns as servicing the vehicle, finding accommodation for the team, and arranging lunch. Finally, it is important that the supervisor be able to read and interpret maps, a point made plain in Box 2.

What are the necessary attributes of the office personnel? The coder must know the survey instruments exceedingly well. Typically, the coder is identified in training as the enumerator with the best grasp of the concepts and details of the questionnaires. The data enterer need not have vast computer experience. Rather, this person should be careful and meticulous in her work. It is even possible to use someone with no experience with computers, as long as the data entry software is user-friendly, and the data enterer is well trained.⁵ Only one coder and one data enterer are necessary for even a large baseline survey.

The survey administrator should be an experienced and responsible person, because he or she may be responsible for large sums of money. Although a person with some experience in accounting or bookkeeping would be ideal, these qualifications are not necessary. Under the supervision of the principal researcher, the administrator will have major responsibility for many aspects of survey implementation.

MANAGING THE SURVEY: THE ROLE OF THE PRINCIPAL RESEARCHER

The principal researcher is responsible for the proper operation of all parts of the process, and for carrying out many of the survey tasks. Researchers are generally shocked at the volume of work with which they are confronted upon entering a country to carry out a baseline survey. This section spells out some of the tasks involved.

The principal researcher is responsible for the proper operation of all parts of the process.

The principal researcher is responsible for developing relationships with key government offices. One reason for this is that government planning offices may be very interested in the results of the baseline survey, and, if informed about what the survey can do for them, can be invaluable allies in opening other government doors from the national to the local level. An additional reason for cultivating close links with key government agencies is that in many countries government permission for surveys is required. In such cases, liaising with the research approval office should be initiated well before the survey is to begin. Even when it is not a requirement, however, a letter of introduction from the appropriate government office can open many doors for a supervisor or enumerator out in the field. A good relationship with the Central Statistics Office is critical, and contact should be made early. This office usually has the census lists and the maps that make sampling possible.

Another important task handled by the principal researcher involves developing a good relationship with the client. Everyone has different experience in research activities, and therefore different expectations about what information a survey can provide. Although the basic purpose of the baseline survey is known to the client, extra time should be allocated during the planning stage for discussing the time frame, administration, detailed content of the questionnaire, and output of the survey.

It is also the responsibility of the principal researcher to design the sampling scheme. The intricacies of this process will be described in Chapter Three; it is sufficient at this point to realize that this is a strenuous, time-intensive process. Another task the principal researcher must complete before the data collection begins is the design of a data entry package. Most data entry software packages, such as SPSS/PC+, are easy to adapt to the particular circumstances of a survey. Nonetheless, some time must be spent carrying out this customization.

The principal researcher also designs and implements the training of all personnel. Chapter Two is devoted to this topic. What is important to note here is that the principal researcher must prepare this session well in advance: training sessions cannot be carried out "off the cuff."

Once training has been completed and the enumerators and supervisors have been chosen, data collection can begin. At this point, the principal researcher's field management duties become important. It is advisable for the researcher to spend as much time in the field as possible. This is particularly true in the early days of the fieldwork, when

enumerators and supervisors are "learning the ropes." During this period, the researcher should make sure that enumerators and supervisors are performing their functions correctly and efficiently. Many questions about the questionnaires and about survey procedure surface at this time, and it is important for the researcher to be on hand to answer them. Perhaps most importantly, the principal researcher should be working closely with the supervisors to ensure that they learn how to manage the survey when the researcher is elsewhere.

As data collection proceeds, the researcher must make certain the main office is operating smoothly. Particularly in the beginning, the researcher should work closely with the coder, answering any questions that arise, and correcting any mistakes. It is also important to check the data that have been entered for keypunch errors. If errors are found, the data enterer must correct them immediately. Finally, the researcher must examine the paper trail that completed questionnaires follow from the field to storage after data entry. This system must work well, and each step along the trail must be documented. Setting up a paper trail is discussed further in Chapter Five.

Nothing can cause a survey to come to a halt as abruptly as a breakdown in the financial and logistical arrangements. Thus, it is crucial that the principal researcher consult regularly with the survey administrator if the administrator is involved in these arrangements. Because the administrator may have little or no contact with the teams in the field, the researcher acts as liaison between the two. Problems of a logistical sort identified by supervisors in the field must be reported to the survey administrator in a timely fashion, so that appropriate changes can be made. Similarly, the financial and other arrangements settled on by the survey administrator must be communicated to the teams. The researcher is the conduit through which these communications flow.

Nothing can cause a survey to come to a halt as abruptly as a breakdown in the financial and logistical arrangements.

Once data from all areas in the sample have been collected and entered into the computer, it is time for the principal researcher to begin the process of data analysis. In this process, the first step is to clean the data. Cleaning involves checking the data for completeness, accuracy, and consistency. Once the data are clean, the full-blown analysis can proceed. The topics of cleaning and analyzing data are covered in greater detail in Chapter Five.

The final tasks of the principal researcher involve writing up the results of the analysis, and disseminating the findings and the data to others who might have an interest in them. One useful exercise, made possible

by the streamlined nature of the MSE survey process, is a presentation of preliminary results to interested parties before the researcher leaves the country. Although the results are not based on clean data and have not yet been weighted to reflect properly the entire country, clients often appreciate hearing about the basic parameters of the MSE sector. Naturally, the shape of the final report depends on the particular preferences of the client. The data should be stored on disk, and sent to anyone who requests them. The details of these tasks are more fully considered in Chapter Six.

SURVEY EQUIPMENT

The survey will not be a success without adequate equipment.

Even with top-flight personnel, without adequate equipment the survey will not be a success. Exactly what are the requirements of a typical baseline survey with respect to equipment? These requirements can be divided into three main groups: office supplies, vehicles, and computer equipment. In the office supply category, all questionnaires, coding sheets, and other survey paperwork must be photocopied and, when necessary, collated and stapled. It is also necessary to provide a clipboard, pencils, erasers, pencil sharpeners, and a rain-proof pouch to each enumerator. Supervisors must be provided with envelopes and forms for sending questionnaires back to survey headquarters. In addition to these, good-quality maps of the areas to be covered should be given to supervisors.

Vehicles form the second grouping of survey equipment. Each team will need a vehicle large enough to carry it, and rugged enough to go to the most remote areas. In some countries, vans are adequate. For example, vans were used in Swaziland, because that country's roads are relatively well developed and the countryside is not terribly rugged. In other places, or if the survey is carried out during the rainy season, it may be necessary for the teams to use four-wheel drive vehicles. This was important in Lesotho, where most of the country is mountainous. One team was even caught in a snow storm in Lesotho and was fortunate to be in an all-terrain vehicle. It is important to decide before the fieldwork begins what sort of vehicles the teams are to use, because rental vehicles are often in short supply in developing countries. Occasionally, only smaller four-wheel drive vehicles are available, and decisions about splitting up the team must be made. Again, being prepared for this eventuality in advance is indispensable. Finally, a car is necessary for the principal researcher. This is not an optional item, given that the researcher

is in perpetual motion visiting the client, the census bureau, and other local experts; picking up supplies and maps; supervising teams in the field; and undertaking many other activities.

A final category of equipment involves the computer. First, the computer itself must be rented. There is no need for an advanced machine, but it should be IBM-compatible, and have at least a 20 megabyte hard drive and a floppy-disk drive. The survey also requires access to a printer. It is preferable to rent a printer along with the computer, but, if this is not feasible, some arrangement should be made whereby documents can be printed from disk fairly easily. It is necessary as well to have software to load onto the computer. First and foremost, some data entry package is needed. One of the most powerful and flexible packages is SPSS/PC+, although there are many other equally suitable programs. It is also important to have access to a word processing program. The researcher should ascertain whether the client can provide these programs in-country; otherwise he or she should bring them on disk.⁶ Finally, it should be possible for the principal researcher to have access to the computer and printers in the evening; otherwise the survey should rent a lap-top computer for her use.

BUDGETARY CONSIDERATIONS

So far, this discussion of survey management has not included financial considerations, except in general terms. This section provides some detail on this important matter, including budget-writing suggestions, information on who should control the survey's pursestrings, and some ideas about building performance incentives into the payment scheme.

The budget is often decided at a level above the principal researcher. It is important, however, for the researcher to be involved in this process, because she must live with the consequences of the budget. As with any type of survey, there will be pressure to reduce some or all of the costs involved in a baseline exercise. It is crucial that the researcher have an idea of the minimum amount of money necessary to complete a survey of reasonable quality, and not allow the budget to be reduced below that point. One line item frequently underbudgeted is the transportation of survey personnel. Information on local rental rates, per-kilometer charges, insurance rates, and petrol prices should be obtained early on and be built into the budget. It will be necessary to estimate the number of

It is crucial that the researcher have an idea of the minimum amount of money necessary to complete a survey of reasonable quality.

kilometers likely to be travelled by all vehicles over the life of the survey. Experience on this issue is presented in Box 3. The researcher must also make sure that all the necessary personnel are budgeted for, and that the appropriate number of person-days is included in the budget, as well as the expenses the teams will incur while on the road. To guide future researchers, an illustrative budget representative of those used in MSE surveys is included in Appendix 2.

Whichever of the administrative arrangements mentioned above is used, the researcher needs to involve herself in how the personnel are paid. Two points are of particular interest. First, the authors have found that it is better to pay the survey teams relatively well. In many developing countries, it is easy to find people willing to work for low wages. Paying more than the individuals on the team are used to getting ensures that higher quality personnel are hired in the first place, and also provides an incentive to continue to perform well throughout the survey and to stay with the survey through completion. The exact wage will, of course, depend on local conditions.

BOX 3

ESTIMATING VEHICLE USAGE

To help the prospective researcher, below are data on the kilometers travelled during the Swaziland and Zimbabwe surveys:

COUNTRY	SIZE OF COUNTRY	DAYS IN FIELD	VEHICLES	TOTAL KILOMETERS
Swaziland	17,000 km ²	30	2 Vans	5,800
			1 Car	800
			TOTAL	6,600
Zimbabwe	391,000 km ²	40	2 Vans	8,000
			1 Car	3,600
			TOTAL	11,600

A second point is that although paying the enumerators and supervisors well provides incentive to work hard, on some MSE surveys additional incentives have been necessary or desirable, and have been implemented with some success. One example involves the payment schedule. Specifically, it may be useful to pay field personnel monthly, making sure that the last month's wages are not paid until all work is complete. If the enumerators are told that they forfeit this pay should they be terminated or leave before the end of the survey, the temptation to slack off during the final days of the survey will be removed. Other incentive plans are more positive. In Zimbabwe, there was an award every two weeks for the most effective team, with effectiveness measured in terms of quality of questionnaires completed.⁷ In addition, every two weeks the top enumerator was selected from each team, with quality again being the operative criterion. Each award was a substantial amount of money, and was effective in fostering a healthy competition between and within the teams. If an incentive scheme is to be used, it must be included in the budget.

SURVEY MANAGEMENT: KEY POINTS



- **Do** be aware of the lists of tasks that need to be done, keeping in mind that the two weeks of presurvey activity are especially difficult.
- **Do** keep in mind the skills necessary for each job when selecting survey personnel.
- **Don't** underestimate the time and skill necessary to do a good job as the principal researcher.
- **Do** arrange to have all necessary supplies and equipment on hand early in the survey process.
- **Don't** underestimate budgetary needs; survey transportation is often underbudgeted.
- **Do** think about monetary incentives for team members, but **don't** pay them by number of interviews produced.

Notes

1. As additional questions or questionnaires are added, more time is needed at every stage of the process.
2. The authors recognize that survey positions may be filled by either men or women; in this manual, the convention has been adopted to refer to enumerators, coders, and survey administrators as "he," and to supervisors, data enterers, and principal researchers as "she."
3. The questionnaire is presented in Chapter Four.
4. The process of assigning unique identification numbers is described in detail in Appendix 5.
5. It is possible to have an overqualified data enterer. A person with extensive computer experience may find this task tedious.
6. The researcher should not take any chances with software. Once the researcher was told that the data entry package would be provided, only to discover upon arrival that there had been a misunderstanding. Had not the researcher brought along copies of the program anyway, the survey would have been delayed until the software could be procured.
7. It is dangerous to reward teams or individuals for quantity of interviews produced, as this may provide incentive for fudging interviews.

CHAPTER TWO

TRAINING

Training is an especially important part of the baseline survey exercise because of the unique questionnaire design and the fast pace that each member of the team must maintain. Training on the questionnaires is required for supervisors, enumerators, and the coder. These personnel are usually trained together during a three- to four-day period of basic training. Additional training is then required for supervisors and coders to explain their job-specific tasks. Specialized training is given to the data enterer on computer operation and the specific design of the programs she will use. Finally, if the survey administrator will be in charge of overseeing the fieldwork, he should also be fully trained on all aspects of the survey work, with special emphasis on how to detect problems.

Training is an especially important part of the baseline survey exercise because of the unique questionnaire design and the fast pace that each team member must maintain.

This chapter covers the content and methods of basic training, as well as specialized training for supervisors, coders, and data enterers. It also discusses post-training review sessions that have been used to identify and fill gaps in the enumerators' understanding. Methods of training vary somewhat by principal researcher. The approach presented here is a synthesis of methods used by the two authors.

BASIC TRAINING

Basic training gives survey personnel an understanding of the baseline survey's purpose, content, and method by introducing trainees to small enterprise research in general, to the matrix-style precoded questionnaire, and to the door-to-door field method. It includes lecture-type sessions, classroom exercises, and field visits.

Basic training introduces trainees to small enterprise research, to the matrix-style precoded questionnaire, and to the door-to-door field method.

Preparation for Basic Training

Basic training requires a physical space that can hold up to 35 people and that is available for the entire week of training. The ideal room configuration includes a single large table around which all of the participants can fit or several long tables set up in a rectangular or "U" shape. This seating arrangement provides a more participatory atmo-

sphere, which is important for classroom exercises, as well as a flat surface for writing. All participants should be able to see the principal researcher comfortably. A large writing stand with paper and markers or a blackboard should be set up near the principal researcher's seat.

Before training begins, the principal researcher should make sure that the following items are on hand for each participant: name tags, blank questionnaires, coding sheets, pencils, sharpeners, erasers, clipboards, exercise sheets, and blank writing paper for note-taking. Many of these supplies are then used during fieldwork.

Prior to training, the decision should also be made whether to pay participants for training time. Experience has shown that some payment (even if only for a percentage of their fieldwork rate) generates better attendance and a better attitude among participants. Other small perks, such as coffee breaks or group lunches, give the participants a feeling of being part of a professional exercise and a chance to get to know each other and the survey leaders informally. Such social amenities also build a more cohesive team spirit, which is important once fieldwork is under way.

Who Attends Basic Training

At least 30 percent more enumerators should be trained than are needed.

Enumerators are the largest group of those attending basic training. As a rule of thumb, the group trained as enumerators should be at least 30 percent larger than that actually needed for fieldwork. This safety margin has proved necessary in nearly every survey for three reasons. First, the job of enumerator is tougher than many applicants expect, and some may drop out. Second, competition within the group for enumerator positions enhances performance. Therefore, if the training group is larger than the number required for the survey, the trainee has strong motivation to prove that he deserves a space on the team. Finally, if someone quits once fieldwork is under way, a ready pool of trained replacements exists to step into the gap.

All enumerators must participate in basic training in its entirety, including arriving on time in the morning. Experience has shown that when this rule has been broken, the enumerator has not been adequately prepared for fieldwork. In addition, this is good practice for fieldwork, when no one can start work until the entire team has arrived, and when absenteeism can wreak havoc on a tight schedule.

Supervisors also need the information that is provided in basic training, and should therefore be required to attend all sessions. Although many supervisors have significant previous survey experience (and often think that they do not need training), it is highly unlikely that they have ever participated in a survey using a matrix-style questionnaire or the cluster sampling field methods, topics that are covered in basic training. In addition, experience has shown that supervisors not participating fully in basic training have not been able to undertake quality control activities in the field, or to assist enumerators with difficulties on the questionnaires.

The coder must also attend basic training to understand the questionnaires and codes completely. The coder is the only person who examines every questionnaire for completeness and accuracy; therefore, he serves as the principal researcher's eyes for quality control. The coder's level of understanding is so important that he is often drawn from the pool of enumerators after the training is complete on the basis of competency in the questionnaires and codes.

Finally, if the survey administrator is to have a role in overseeing the fieldwork along with the principal researcher, it is important that he attend basic training in its entirety. If he will be directing enumerators and supervisors, he will need training on field techniques; and if he will be overseeing data coding and entry, he will also need to know the questionnaires and codes.

Illustrative Schedule of Basic Training

The content and specific methods of basic training are provided in Appendix 3. Here is a hypothetical basic training schedule for a survey with two questionnaires: the basic questionnaire and the closed business questionnaire, a sister survey often administered with the basic questionnaire. Because the survey includes two questionnaires, three and one-half days of classroom training are needed, plus a half day for the market visit and one day for the pretest.

DAY/TIME: TOPIC/ACTIVITY:

Day 1, AM: Introduce those in attendance.
Cover administrative matters.
Give substantive introduction to the subject matter and background to the survey.

- Day 1, PM:** Introduce basic questionnaire structure.
Introduce basic questionnaire content.
Introduce basic questionnaire codes.
- Day 2, AM:** Review basic questionnaire content and codes.
Do mock interviews of basic questionnaire in English with roundtable discussion of difficulties.
- Day 2, PM:** Discuss on-site interviewing and survey conventions.
Do more mock interviews of basic questionnaire in English.
Give written exercise on basic questionnaire.
- Day 3, AM:** Discuss problems found in written exercise.
Do more mock interviews of basic questionnaire in English.
Introduce closed business questionnaire purpose.
Introduce closed business questionnaire content.
Introduce closed business questionnaire codes.
- Day 3, PM:** Review closed business questionnaire.
Do mock interviews of closed business questionnaire in English.
Do mock interviews using both questionnaires together.
Give second written exercise.
- Day 4, AM:** Discuss problems found in written exercise.
Break into small groups for translation exercise.
Reconvene entire group to discuss translations.
Do mock interviews on both questionnaires in survey languages.
- Day 4, PM:** Make field visit to marketplace for observation.
Return to classroom for debriefing.
- Day 5, AM:** Discuss how to conduct fieldwork.
Look at maps in reference to determining survey areas and boundaries.
Travel to field and conduct pretest.
- Day 5, PM:** After completing pretest, have debriefing session back at headquarters.
Finalize arrangements for time and place to meet for fieldwork.

Had the closed business questionnaire not been included in the survey, the training plus pretest could have been condensed into a four-day period. Adding a third instrument to the survey, such as the supplementary questionnaire, would have increased the required training time by another half day.

SUPERVISOR TRAINING OPTIONS

Although basic training covers much of the information supervisors need, additional training will also be required to discuss supervisory responsibilities and provide them with information they need to carry out their tasks. In the past, supervisor training has too often been carried out on coffee breaks during basic training. Given the importance of the supervisors' role, it is important to dedicate time for their specific information needs. If time and resources allow, the best time for supervisor training is just before basic training begins, and should be allocated two days. The supervisors then attend basic training as part of their duties, where they serve as resource people as well as strengthening their own grasp of the questionnaires. A brief description of the optimal early training schedule is below.¹

Additional time should be allotted to training needs of supervisors.

The first day of supervisor training, an intensive introduction to the survey instruments and the field methods is given. All questionnaires are covered in the morning session, and supervisors are provided with the questionnaires, coding sheets, and a written description of the questions.² That afternoon, training covers the information they will need to lead a team in the field: administrative arrangements (driving, expenses, lodging, and so forth); the survey schedule; map-reading and boundary definition; protocol in the local areas; questionnaire handling procedures; and enumerator oversight.³

On the second day, the supervisors meet to discuss any difficulties with the questionnaires. Then the supervisors take the questionnaires to the field and pretest them, specifically checking that questions fall in a logical order and that the coding systems are complete. Any information the supervisors generate in the pretest can then be used to make revisions on the questionnaires before basic training begins.

When in the throes of presurvey planning, setting up a separate supervisor training session seems like an overwhelming task. However, it is well worth the effort and can save the principal researcher time in the

long run. As an immediate benefit, the participation of pretrained supervisors improves the quality of basic training — they can serve as resources on survey content and method and as leaders in classroom and field exercises.

SUPPLEMENTAL CODER TRAINING

The coder attends all of basic training, during which time he learns about questionnaire layout, content, and codes. In addition, however, he needs to learn techniques to spot and correct errors on the completed questionnaires. This mainly requires common sense plus the information provided in basic training, but it is useful for the principal researcher to allocate some time to explicitly go over a sheet listing the coding rules for each question, which the coder can use to refresh his memory as needed.

The coder should also be shown some of the most common mistakes and told how to handle them. The principal researcher should decide which errors can be corrected by the coder and which errors need to be flagged for the principal researcher's attention. These rules need to be communicated to the coder.

When teaching the coder how to spot problems, it is helpful to develop a mock completed questionnaire in which some entries have errors. The coder can then practice checking the codes, and deciding which errors he can correct himself and which he must record as missing information or seek guidance on.

The coder should also be trained on how to assign unique identification numbers to both the proprietor and the enterprise, a process which is more complex when the survey has multiple questionnaires. The method of assigning identification numbers is described in Appendix 5.

Supplemental coder training requires only about two hours, and should be carried out after the basic training is completed. In addition, it is important for the principal researcher to stay in close touch with the coder during the first weeks of fieldwork to answer questions and to do quality checks on both the coding and the assignment of identification numbers.

DATA ENTERER TRAINING

The data enterer is the only team member who does not attend basic training, because she does not need to know the content of the questionnaires or the coding system. What she does need to know is the computer program used for data entry, and how data entry files correspond to the questionnaires.

The data enterer is the only team member who does not attend basic training.

Data entry training begins with an introduction to the computer, and with instruction on how to access the necessary files and how to store and back up files. Once the data enterer is familiar with the computer as an instrument, she is then introduced to the specific software package to be used. Training should focus on that part of the software package that she will use. In SPSS/PC+, the most commonly used package, there is a data entry module complete with a training manual that suffices. The principal researcher will be well served by writing out a list of commands for the data enterer to follow and refer to.

Once the data enterer is familiar with the software package, the questionnaires can be introduced. For each questionnaire, the information to be entered into the computer should be highlighted, and variable names assigned. These variable names should be written in the appropriate space on blank questionnaires as a guide to connect the questionnaire information to the variables found in the computer program, and should be posted near the computer terminal.

At this point, it is helpful to do a data entry exercise. For the exercise, the data enterer will need empty data files for each questionnaire, as well as mock completed questionnaires. She can then practice booting up the computer, entering the data entry module of the software program, accessing the correct data file, entering the data from the questionnaires into the computer, and, finally, saving her work and exiting the program.

One important item to cover in training is the system of backing up the data, which is discussed in Chapter Five. Rules on when to back up data, how to label different generations of back-ups, and where to store the files should be clearly spelled out. As with the coder, it is a good idea for the principal researcher to be on hand early in the data entry process to answer any questions that arise.

POST-TRAINING TROUBLE-SHOOTING SESSIONS

The most effective time for identifying and clearing up questions is often not in training itself, but after a few days in the field, at the coding desk, or at the computer.

Not surprisingly, questions on all aspects of the survey arise once the enumerators and supervisors have been in the field for a few days, the coder has weeded through dozens of questionnaire forms, and the data enterer has entered a few hundred cases into the computer. Therefore, the most effective time for identifying and clearing up questions is often not in training itself, but after a few days in the field, a few days at the coding desk, or a few days at the computer. For this reason, the authors advocate three post-training trouble-shooting sessions led by the principal researcher: one for the field teams, one for the coder, and one for the data enterer.

Trouble-shooting With the Field Teams

The best time for the meeting with the field teams is near the end of the first week of fieldwork, once the principal researcher has carefully examined each enumerator's first completed questionnaires for coding and completeness. It is important not to wait too long for this session, because any corrections or changes should be made as early in the survey as possible.

Before the entire field team meets, it is helpful for the principal researcher to sit down with the supervisors and discuss any matters of concern that have emerged during fieldwork. This is also an opportunity for supervisors to report problems with specific enumerators, either with respect to their work or the way they deal with their teammates or survey respondents.

Trouble-shooting With the Coder

The principal researcher needs to check the quality of the coder's work early in the survey. The coder's job is complicated, and involves innumerable details to remember; it is only natural that he may overlook some of those details. Only by actually doing coding checks of questionnaires already completed by the coder can the principal researcher identify where the coder needs further instruction. After checking the questionnaires, the principal researcher can then meet with the coder to share concerns.

Trouble-shooting With the Data Enterer

The principal researcher also needs to check the quality of the data enterer's work early in the survey. The only way the principal researcher can identify where the data enterer needs further instruction is by sitting down at the computer and looking at the data. Ideally, the principal researcher should do some of the early data verification, which involves re-entering data from a given set of questionnaires in search of errors from the first round of data entry. In addition, the principal researcher can run some diagnostic tests on the entered data, looking for numbers that fall outside of allowed ranges or that violate some rules — problems that suggest either a keystroke error or a coding error. After checking the data, the principal researcher can then meet with the data enterer to share concerns.

TRAINING: KEY POINTS



- **Do prepare lesson plans and exercises for all training sessions in advance.**
- **Do begin training with completed and pretested questionnaires.**
- **Do have preliminary training for supervisors if possible.**
- **Do require complete attendance in training as a prerequisite for participation in fieldwork.**
- **Do use participatory exercises as part of training.**
- **Do train more enumerators than necessary for fieldwork.**
- **Do schedule time for trouble-shooting sessions.**

Notes

1. **Even if it is impossible to have separate supervisor training, it is still necessary to find time to brief the supervisors completely on the topics mentioned in this section.**
2. **The supervisors are expected to use these documents to do any further review needed to make themselves familiar with all the questionnaires before the following day.**
3. **The oversight task is not insignificant. Supervisors are responsible for ensuring that the team starts work promptly, that enumerators work individually and do not skip sites, that the team does not fall behind their schedule, that interviews are not forged by enumerators, that questionnaires are complete, and that team morale remains high.**

CHAPTER THREE

SAMPLING

CHOOSING THE SAMPLING METHOD

Even in the smallest of countries, there are an enormous number of micro- and small-scale enterprises. This complicates the task of the researcher, because it is not possible given the resources of the typical survey to enumerate every MSE in the country. Some method must therefore be used to select a feasible number of enterprises that are as representative as possible of the population of all MSEs in that country.

Each of the countrywide MSE Baseline Surveys identifies enterprises to be interviewed by means of a technique called stratified cluster sampling.¹ This method consists of two steps. The first involves dividing the country into strata. This designation is usually based on demographic characteristics of the country in question. For example, most MSE surveys contain urban, rural, and smaller-town strata. The second step entails the selection of small groups, or clusters, of enterprises within each stratum. Once the choice of strata has been made and the clusters have been identified, enumerators go door to door within the chosen clusters in search of enterprises.

Why do these surveys employ stratified cluster sampling? Are there alternative sampling frameworks? One obvious way to draw a sample would be to select MSEs randomly from the country as a whole. As the sample size increases, the accuracy with which the sample describes the population increases. Unfortunately, this is not possible: no list, or sampling frame, of MSEs exists from which to draw a sample.

The method of choice is stratified cluster sampling.

A more feasible alternative to stratified cluster sampling might involve randomly selecting localities from a list of all such areas in the country. This option is not palatable for two reasons. First, even small countries are diverse. For example, some people live in urban areas, some in rural areas, and some live in smaller communities. It seems reasonable that this distribution influences the characteristics of MSEs in each area. There are many other ways in which countries are internally diverse. To capture this diversity effectively in a random sample, a very large number of units would need to be chosen. Stratification, properly done, reduces the sample size necessary for a given level of precision by grouping

together similar MSEs.² A second reason why drawing a random sample of localities is not preferred is that stratified cluster sampling is likely to be more economical. As mentioned above, stratification reduces the sample size needed for a given level of precision. This obviously has cost implications. The use of clusters also is likely to be less costly. Instead of traveling to a large number of widely dispersed enterprises selected in a random draw, cluster sampling enables a large number of interviews to be conducted in a concentrated area.

CHOOSING STRATA

MSE surveys often stratify by population size.

It was noted above that selecting strata properly would reduce the necessary sample size. How can the researcher make this decision? Clearly, there is an infinite number of ways a researcher could split up any country. One could stratify by geographic or climatic regions, along tribal lines, or by demographic characteristics, to name only a few. MSE surveys have often chosen variations on the demographic criterion, dividing enterprises according to the population size of the areas in which they are located. This seems to be a natural way to group MSEs, since enterprises within such strata face similar demand and supply conditions. This stratification scheme has the additional advantage of being simple to put into operation. What matters, however, is that MSEs within strata be as alike as possible, and that MSEs be as different as possible across strata. In other words, the within-stratum variance should be small, but the between-strata variance large. Unfortunately, these variances are not known beforehand. Of course, one can make these calculations *ex post*, but this only informs as to whether the stratification actually chosen was appropriate. Further information on this topic, in the form of a case study, is available in Box 4 at the end of this chapter.

An additional wrinkle is that each country has unique features that may be important to include in the stratification. For example, in Zimbabwe, it was necessary to separate small towns from "growth points," which are towns of similar size but which operate under special investment and tax incentives from the government. The two types of towns were judged to be sufficiently different that they warranted separate strata. Such special cases arise in every countrywide survey.

Given that the variances between and within potential strata are unknown, and that each country has special circumstances, it is important for the researcher to learn as much about the country as possible before

the exercise begins. Perhaps more importantly, it is highly desirable to consult with several local experts in the design phase of the survey. In the past, MSE surveys have relied on experts from census bureaus, members of the business community, academic scholars, and others familiar with the country including personnel from the U.S. Agency for International Development or other donor organizations.

A final point on the topic of the selection of strata involves the number of strata to include in a given survey. As the researcher consults with local experts, pressure will mount to increase the number of strata, to be certain that some particular MSE feature is included. Resources for conducting the survey, however, will typically not increase. This means that as the number of strata increases, the number of enterprises enumerated in each goes down. Naturally, this makes the information gathered about MSEs in each stratum increasingly imprecise. Often it is the client who is applying this pressure, and the researcher must point out the costs involved in adding strata. Although as many as 20 strata have been used in MSE surveys, experience has shown that a more reasonable number may be 4 or 5. Fewer than 4 strata may mean that the within-stratum variance is higher than the between-strata variance, while more than 5 strata may stretch available resources too thinly.

Chose no more than five strata for the survey.

CHOOSING CLUSTERS

Having settled on the appropriate stratification scheme for the survey, it is necessary to select clusters within each stratum. Although there is no single right way to proceed, one rule of thumb has been condensed from sampling theory. From the standpoint of sample precision, it is better to have a large number of small clusters than a small number of large ones.³ Of course, the desire to sample many tiny clusters must be balanced with the realization that costs increase with the number of clusters.

Equipped with the knowledge that having many small clusters is better from a precision point of view than a smaller number of larger clusters, what is the best way to actually select clusters? Although no list of MSEs within each stratum exists, a sampling frame of sorts is usually available. Most countries have had a population census carried out in the recent past, and many repeat this exercise every several years. The Census Bureau of any country typically has divided that country into small units, usually called enumeration areas, to which they send their enumerators.

These enumeration areas, or EAs, are natural clusters to use in a baseline survey exercise.⁴ However, to increase the precision of the survey, it may be advisable to subdivide EAs so that the average cluster size is smaller.⁵ Lists of all EAs, and the population and number of households in each one, can often be obtained from the Census Bureau or Central Statistics Office.

EAs form a sampling frame for each stratum.

Once the EAs are sorted by strata, they form a sampling frame for each stratum. If the census was carried out fairly recently, the information regarding number of households per EA will be fairly accurate. This knowledge can be used to estimate the expected number of enterprises per EA. The percentage of households engaged in MSE activity is fairly constant across countries, at least in southern Africa. This proportion has been reported to be 29 percent in Kenya, 30 percent in Lesotho, 35 percent in Zimbabwe, and 39 percent in Swaziland.⁶ Thus, if it is known that the average EA in a rural area contains 150 households, it would be reasonable to expect 45-60 enterprises in the EA. This is useful information to have when deciding how many EAs will need to be covered to sample the desired number of MSEs.

Select enumeration areas for interviews.

Once the sampling frame for each stratum is assembled, EAs can be selected. This needs to be done randomly, using either a random number table, or a random number generator on a computer or hand calculator.⁷ It is important to note that once this drawing is completed, the selected areas must all be covered. If resources are so constrained that a certain area, if selected, could not be covered, the researcher should exclude this EA before the random draw.⁸ Similarly, areas of interest that happen not to be selected by the draw cannot be added later. If there are such areas, these can be handled in one of two ways. First, if these special areas are interesting enough, they should be given their own stratum. For example, if the client has a particular interest in fishing villages along the shores of Lake Kariba, these small towns should be separated from all other small towns in the stratification. A second way to handle these is to purposively, rather than randomly, sample a number of the interesting EAs. These cannot be included as part of the sample, but the special information sought by the client can be provided this way. In any case, these decisions must be made in the design phase of the survey and not after data collection has begun. This should be made clear to the client early on.

In addition to the EA sampling frame, it is imperative to be able to procure high-quality maps of each EA selected, with the EA boundaries clearly defined. In even a small country like Swaziland, there are thousands of EAs. Maps are crucial in getting the enumerators to the

correct EAs and in knowing the boundaries of the areas to be enumerated. As noted in Chapter One, most countries have an assortment of maps available for sale from the Census Bureau with specific EAs delineated.⁹ These maps should be as detailed as possible, with the preferable scale being 1:50,000. Occasionally, as in Swaziland, aerial photographs are available. These are useful, especially in conjunction with physical maps of the same areas.

BOX 4

SAMPLE SIZE IN ZIMBABWE

This chapter has offered general advice on the selection of strata and clusters. It has not been specific on the overall sample size that should be the target of the researcher when the survey is designed, or the size for each stratum. These issues are complex and require some information not available to the researcher. In a survey with unlimited resources, the researcher would use the smallest-sized clusters possible. Surveys, however, always have resource constraints. As a result, the researcher must trade precision for cost in deciding on the cluster size to use. Similarly, the allocation of resources between the strata is troublesome. Clusters in urban strata are relatively cheap to enumerate, since the clusters tend to be close to one another, and each is in a concentrated area. The opposite is true in rural areas. Because the cost of enumerating an additional MSE in each stratum is not known to the researcher, and because the variances across strata are unknown, the decision about allocating time in the field per stratum is not an easy one. In the past, MSE surveys have aimed at covering at least 1 percent of the enterprises in the least-covered stratum. This is a useful departure point, but has not always been possible, given the size of the country and the resources available.

To shed light on these difficult decisions, we present a case study involving a recent MSE baseline survey in Zimbabwe. Table A presents some basic information about this survey. As can be seen, only a small fraction of the clusters, or of the total number of MSEs in each stratum, were enumerated. Such small sample proportions are not necessarily any cause for alarm, however. If the variances within each stratum are low, one does not need to sample very many enterprises to get a reasonably accurate picture of the MSEs operating there. Unfortunately, one knows this information only after the survey is completed. An *ex post* analysis, then, can be useful in examining the validity of the sampling approach used, and can shed light on the design of subsequent MSE surveys in the same or in other similar countries. It is in this vein that this case study is presented.

To assess the appropriateness of the sample size of each stratum, the Zimbabwe survey was carefully scrutinized *ex post* using statistical sampling theory. The outcome showed the survey to be basically sound, but that several strata were either over- or under-sampled. The exercise calculated the appropriate sample size in a stratified design with a margin of error of 20 percent, and then adjusted this for the imprecision of cluster sampling. The formulae account for differences in variances across strata, as well as the average cluster size per stratum. The "correct" sample size as well as the actual number of MSEs enumerated is presented in Table B, where "correct" refers to the required number to attain the given level of precision of results, given what we know now (after the survey) about variances within strata.

BOX 4 -- Continued

TABLE A

ZIMBABWE BASELINE SURVEY:
BASIC SAMPLING PARAMETERS

Stratum	No. of Clusters in Stratum	No. of Clusters Sampled	% of Clusters Sampled	Est'd. No. of MSEs	No. of MSEs Sampled	% of MSEs Sampled
Urban Areas						
High Density	312	9	2.9	225,032	2,754	1.2
Low Density	469	12	2.6	35,883	294	0.8
Commercial Districts	120	9	7.5	6,884	518	7.5
Industrial Areas	107	8	7.5	2,583	194	7.5
Rural Areas						
District Councils	2,372	7	0.3	398,177	358	0.1
Rural Councils	719	4	0.6	90,047	139	0.2
Smaller Towns	55	4	7.3	60,224	765	1.3
Growth Points	69	5	7.3	26,604	553	2.1

TABLE B

ZIMBABWE BASELINE SURVEY:
ACTUAL AND APPROPRIATE SAMPLE SIZE BY STRATUM*

Stratum	Actual Sample Size	Distribution of Sample	Appropriate Sample Size	Appropriate Distribution of Sample
Urban Areas				
High Density	2,754	49.4%	2,864	41.1%
Low Density	294	5.3%	145	2.1%
Commercial Districts	518	9.3%	47	0.7%
Industrial Areas	194	3.5%	11	0.2%
Rural Areas				
District Councils	358	6.4%	2,008	28.8%
Rural Councils	139	2.5%	193	2.8%
Smaller Towns	765	13.7%	1,313	18.8%
Growth Points	553	9.9%	387	5.6%
TOTAL	5,575	100.0%	6,968	100.0%

* Assuming a stratified cluster sampling design and a 20% margin of error.

BOX 4 -- Continued

Table B reveals that district councils and smaller towns were severely underenumerated, while too many resources were spent in urban commercial and industrial districts, and in rural growth points. Still, all of the urban strata and two of the rural strata were covered adequately. The table also indicates what the appropriate distribution of the sample across strata should have been. Although the set of circumstances in Zimbabwe was unique, this distribution gives the prospective researcher a rough idea of how to allocate her resources when designing a survey elsewhere.

Every survey has shortcomings, and those of the Zimbabwe survey are probably not excessive. Nevertheless, it is instructive to consider how this situation arose, and how it might have been possible to improve the design of the survey. A large part of the problem was that there were too many strata for the resources available. Although the principal researcher had intended to include only five strata, he quickly came under pressure to add more. Most of these additions were resisted, but three new strata were added. These additions were interesting and important, but this caused the survey resources to be stretched thinly, particularly in the rural areas. As noted above, four or five strata is probably a good balance between enough and too much stratification.

A second lesson that emerges from the exercise involves cluster size. Too few clusters were chosen, each of which was too large. Accuracy would have been higher with more and smaller clusters in each stratum. Whenever possible, then, care should be taken to keep the average cluster size as low as is feasible.

In addition to these lessons, this case study demonstrates that too many resources were spent in urban areas and not enough in the rural strata. A reallocation of resources within the sampling design would have improved precision, even without reducing the number of strata or the average cluster size.

The Zimbabwe baseline survey was designed under a unique confluence of circumstances. Nevertheless, the strengths and weaknesses that were exposed by the postsurvey design examination may be useful to researchers designing future surveys in other countries. Such ex post statistical scrutiny should become a routine part of the baseline survey process.



SAMPLING: KEY POINTS

- **Do** choose the sampling scheme as early as possible.
- **Don't** have too many strata: four or five is a reasonable number.
- **Do** choose the smallest feasible cluster size.
- **Do** discuss the sampling scheme with the client before starting fieldwork.

Notes

1. The exceptions to this are two urban baseline surveys (in Kibera, Kenya, and in the South African townships), which were censuses and therefore did not involve sampling.
2. For algebraic discussions of the relative precision of stratified sampling and simple random sampling, see Cochran (1977).
3. This is because the enterprises in a given cluster will tend to be more alike than those in different clusters. A small number of clusters is less likely to capture the diversity within the stratum than is a large number of smaller clusters. For more detail on the mathematics of cluster and overall sample size, see Appendix 10. Further information regarding appropriate cluster sizes is presented in Box 4 at the end of this chapter.
4. An important exception can be industrial and commercial areas. Occasionally, these areas are not divided into EAs by the census bureau, because no people live in this areas. Other times EAs exist, but they are very large. One solution is for the principal researcher, having clearly established the boundaries of these areas, to divide industrial and commercial sites into equal areas on the map. These areas can then be randomly selected and enumerated.

-
5. This is a straightforward exercise, as long as good maps of the EAs are available. For example, if a river or stream runs through an EA, this provides an obvious way to divide the area into two clusters.
 6. See Parker and Dondo (1991), Fisseha (1991), McPherson (1991), and Fisseha and McPherson (1991).
 7. By drawing EAs at random, the representativeness of the choice is enhanced. Nonrandom selection is likely to introduce bias.
 8. An anecdote from the recently completed Malawi baseline survey sheds light on this topic. A small island in the middle of Lake Malawi was selected as an EA to be covered. Thought was given to excluding this area because of the costs involved in covering it, and because some felt that MSEs on the island might not be representative of MSEs countrywide. Because this area was not excluded beforehand, the principal researcher correctly insisted that this EA be covered.
 9. Other government agencies may sell maps, such as the Central Statistics Office or the General Surveyor's Office.

CHAPTER FOUR

INSTRUMENT DESIGN

One of the attractive aspects of choosing the MSE Baseline Survey is that it follows a standardized approach, which makes instrument design a straightforward exercise. Despite its apparent simplicity, however, instrument design still begins with the difficult task of defining the exact content of the baseline survey questionnaire. This chapter opens with a discussion on determining the scope of the survey. It then describes the standard questionnaire design, and closes by covering two common extensions to the basic questionnaire that make it more responsive to the particular needs of a given situation.

DETERMINING THE SURVEY'S CONTENT

Perhaps the most important task at the planning stage is to decide on the content of the baseline survey, in part because this is where conflict can easily arise between the client and the researcher. There are typically several rounds of discussion on questionnaire content, which winnow out a list of critical information to be collected and develop a list of optional questions to be addressed if space allows. It is extremely important that the client be integrally involved in this process, so that the researchers understand what is important to the client, and the client understands what is possible in a survey of this type.

At the planning stage, winnow out a list of critical information to be collected.

One important point for the researchers to convey to the client is that the greatest strength of the baseline survey comes from its leanness. Although this means that the baseline survey does not waste time and resources collecting unnecessary data, it also means that it is not suitable for collecting complicated data on such matters as sales, income, or capital stock, or data on such matters as policy impact.¹ This is often a point of frustration for the client, which feels the need for detailed data on many variables for planning and monitoring purposes.

The greatest strength of the baseline survey comes from its leanness.

If it is essential that more information be collected than can fit on the basic survey form, a supplementary questionnaire can be added that collects targeted information on specific subjects from a smaller sample of respondents. Again, the client and researcher need to work together to decide what information can and should be collected in any extension of

Previous Page Blank

the baseline survey. If the baseline survey is extended to include additional questionnaires, the client should realize that their inclusion will reduce the precision of the basic questionnaire unless more resources and more field time are provided.

THE BASIC QUESTIONNAIRE

The basic questionnaire is the core of the baseline survey. It contains the most critical information of the survey, and, if used with other questionnaires, is always administered first. An illustrative basic questionnaire is provided in Figure 2.

Structure of the Basic Questionnaire

Each column is a question; each row a complete interview with a business.

When first handed the basic questionnaire, people are usually surprised by its appearance. It is a matrix formed by a series of columns and rows. Each column is labeled by a word or phrase, and there is a list of codes in tiny print at the bottom of the page. How is this matrix used?

First, each column is a question. The heading on each column is a key word or phrase to remind the enumerator what question to ask. The typical survey has 30 columns, and therefore 30 questions. Some of the columns are wide enough to allow the enumerator to write down detailed information, such as where the business is located or the exact nature of the business. Most of the columns, however, are only wide enough to allow a code or number to be recorded in them. Second, each row on the page is a complete interview with a single business. The typical page has five to six rows, which means that five to six full interviews can be recorded on a single page. Below the last row, codes are listed to refresh the enumerator's memory as needed.

The advantages of the matrix format are no page turning, no shuffling with coding sheets, and no delays.

What are the advantages of this matrix format? During the interview, there is no page turning, no shuffling with coding sheets, and no delays while the enumerator records answers in longhand. And best of all, each page that an enumerator turns in provides data on several businesses that are completely precoded and ready to be entered into the computer. One of the challenges of using a matrix format is that the enumerators must know the questionnaire intimately — they must be familiar with the order of questions, they must all use identical phrasing of questions despite the fact that the questions are not written out, and they

ILLUSTRATIVE BASIC QUESTIONNAIRE

) Region: _____ Date Completed: _____ Page #: _____ Size With No Answer: _____
) District: _____ Date Filled: _____ Business: _____
) Enumeration Area: _____ Date Entered: _____ Signature: _____ Sex: _____
 Date Valid: _____

1	PROPRIETOR SEQUENCE #	
2	ENTERPRISE SEQUENCE #	
3	PROPRIETOR ID #	
4	ENTERPRISE ID #	
5	ENTERPRISE TYPE	
6	ENTERPRISE CODE	
7	# FAMILY INCOME (a)	
8	LOCATION TYPE (b)	
9	STRUCTURE TYPE (c)	
10	# MONTHS PER YEAR	
11	# WORKING PROPRIET.	
12	# UNPAID FAMILY	
13	# PAID WORKERS	
14	# APPRENTICES	
15	TOTAL # OF WORKERS	
16	# FEMALES	
17	# PART-TIME	
18	# CHILDREN	
19	PAY OF BEST PAID WORKER	
20	PAY PERIOD (d)	
21	YEAR BUSINESS WAS ACQUIRED	
22	# WORKERS AT START	
23	SELLS TO WHOM? (e)	
24	NATURE OF INPUTS (f)	
25	CREDIT SOURCE (g)	
26	PROPRIETOR GENDER (h)	
27	PROPRIETOR CATEGORY	
28	TOTAL # OF EH ENTERPRISES	
29	# IN EH IN PAID EMPLOYMENT	
30	PROPRIETOR'S NAME, NICKNAME, ENTERPRISE NAME, PHYSICAL ADDRESS	
1c	SUPPLEMENTARY QUEST.? (1=YES, 0=NO)	
2c	ANY PAST ENTERPRISES? (1=YES, 0=NO)	

Code
 of 1. Provides all of household income 2. Provides more than half of income 3. Provides less than half of income 4. Provides about half of income
 M 1. Non-commercial 2. Traditional market 3. Commercial district 4. Residential, not in commercial district 5. Mobile
 of 1. Permanent structure 2. Temporary structure 3. Road Only 4. Open air 5. Other
 of 1. Per hour 2. Per day 3. Per week 4. Per month 5. Per year 6. Per piece 7. Other
 of 1. Individuals 2. Urban commercial enterprise 3. Urban manufacturing enterprise 4. Rural commercial enterprise 5. Rural manufacturing enterprise 6. Export 7. Other
 of 1. Makes/purchases own inputs 2. Buys unprocessed inputs 3. Buys semi-processed inputs 4. Buys finished products for resale 5. Other
 of 0 Never received loans 1. Loans from family/friends 2. Loans from entrepreneur 3. Loans from bank 4. Loans from elsewhere
 M 1. Female, one proprietor 2. Male, one proprietor 3. Female, more than one proprietor 4. Male, more than one proprietor 5. Mixed-gender joint proprietorship

FIGURE 2

must know the codes. These requirements justify an entire week for training, as discussed in Chapter Two.

Content of the Basic Questionnaire

What information can possibly be collected on such a tight questionnaire? The basic questionnaire usually collects the following information:

- Type of business activity;
- Importance of business to family income;
- Location of business;
- Seasonality of business;
- Number of workers currently in business, with a break down by type;
- Number of workers in the business when it first opened;
- Year in which the business was established;
- Proprietor's gender;
- Level of specialization in the business;
- Whether proprietor has previous business experience; and
- Name and address of the proprietor.

Appendix 5 includes detailed notes on how each variable included on the illustrative questionnaire has been defined and measured.

These variables show the number of enterprises in different activities, where different groups of enterprises can be found, who operates and is employed in the enterprises, and how important the enterprises are to household income. In addition, from these variables one can develop a profile on the size and growth of enterprises, both for the entire population and for subgroups of the population broken down, for example, by type or location of activity or by the gender of the proprietor. These are often the kinds of information that policy makers and project designers are looking for in deciding how and where to target small enterprise assistance.

Including the name and address of the respondent takes both time and questionnaire space, and is often difficult to record accurately. However, to use the baseline survey as a sampling frame for future studies, this information will be required to locate the respondents when needed. It is placed near the end of the questionnaire in case it causes respondents to falsify information or terminate the interview. Assuring respondents of confidentiality can help alleviate these problems.

The basic questionnaire usually has two or three free columns that can be used for questions that change from survey to survey. These open columns are often used to collect information on questions of particular interest to the client. For example, in Swaziland, the client was interested in what percentage of MSEs belong to business support groups, so one of the free columns was used to record respondents' answers to this question. However, if an additional questionnaire is being administered with the basic questionnaire, one of the free columns must be used to tie the two questionnaires together.² Similarly, if two questionnaires are added, two columns must be used to tie the questionnaires together.

The questionnaire has free columns that can be used for questions that change from survey to survey.

Coding

As mentioned above, most codes appear in small print at the bottom of the questionnaire. The goal is to minimize paper shuffling that can distract both the enumerator and respondent. Codes must follow two rules: they must be all-inclusive and not overlap. The all-inclusive rule means that there must be a code for every conceivable answer. To address this, once the most common answers are assigned a code of their own, an "other" category should be included to capture all other responses. To obviate overlapping, codes must be structured so that an answer can fit under only one code. Even with the most careful planning, these two rules are occasionally violated, but these cases can usually be uncovered during the pretest and corrected before the survey gets under way. More frequently, the researcher runs into a third problem — that a major code has been omitted, and all the answers in that category are falling in the "other" category. This problem may only appear once the survey has begun and the problem is identified by enumerators. If discovered early, an additional code can be added to capture that omitted category.

Codes must follow two rules: they must be all-inclusive and not overlap.

The easiest code is for questions framed in "yes/no" terms. "Yes" is usually coded as "1," while "no" is coded as "0." No matter what codes are given to "yes" and "no," however, they should be consistent across all "yes/no" questions to avoid enumerator confusion and miscoding. Similarly, if "yes/no" is coded as 1/0, then all other dichotomous variables (such as "male/female" or "member of X/nonmember of X") should be coded as "1/0" as well, rather than as "1/2" or any other choice. Otherwise, some enumerators are likely to start coding "yes/no" as "1/2" as well.

Some columns do not require coding. These are the columns in which information is recorded longhand (for example, respondent's name

In some columns, the enumerator records actual values.

and address), and those that record an actual value. There are several columns that ask for values rather than codes. For the year the business was established, for example, enumerators record the last two digits of the actual year in the column ("86" for a business started in 1986). Also, actual members are requested in response to other questions such as number of workers. Distinguishing between the columns that require codes and those that require actual values is something that enumerators are taught during training.

One set of codes — the enterprise code list — does not appear at the bottom of the questionnaire, because it includes up to 100 specific activities. Each survey's enterprise code list is unique, reflecting the most common activities in that country's context. An illustrative business code sheet can be found in Appendix 7. The one commonality across all countries is that the list is broken down into three main categories: manufacturing, trade, and service activities. In some surveys, the trade category is further divided among vending, retailing, and wholesaling activities; however, this introduces some subtleties in definition that can be difficult to convey to enumerators.³ During the data analysis phase, the codes are often converted to the International Standardized Industrial Classification (ISIC) one-, two-, and four-digit codes (provided in Appendix 8) to make the categories more standardized for cross-country and cross-survey comparisons.⁴

Conventions for Completing the Questionnaire

It is surprising just how many complexities the baseline survey includes. To deal with some of the more puzzling questions, eight conventions have been designed to help enumerators decide how to record information.

Convention 1

Collect information about a business only at its place of operation.

As was noted above, enumerators go door to door to both homes and businesses, as well as interview businesses operating out in the open, either in a fixed location or mobile. As such, it is possible that multiple members of a single household could be interviewed at different locations. As an example, one enumerator could interview a woman found in the home, while another enumerator interviews the woman's husband in the marketplace. If both respondents discuss the same business, that business would be counted twice, or double-counted. Magnify this error for the entire survey, and survey results would be biased. To avoid double-

counting, the first survey convention is invoked: *collect information about a business only at its place of operation*, be it in a home, along a road, in an open-air marketplace, or in a business premise. Because each site is visited only once, each enterprise will be included in the survey only once, and double-counting of businesses will be avoided.

Here is another example of how on-site interviewing works. If an enumerator knocks on a door and finds a woman running a business in that home, he collects data on that business since it is operating at that site. If the woman mentions that her husband also has a business in the home, he collects information on that business as well because he is also at the site of the husband's business. If, however, the woman says that her husband has a business in the marketplace, the enumerator assumes that someone will collect information on that business when they visit his business site at the marketplace.

Convention 2

Record information about a site even when no one at that site has a business. A space is provided on the questionnaire to keep a tally of the number of sites where no one has a businesses. This information is collected for an important reason. The number of sites with enterprises will be compared to the number of sites without enterprises to come up with the percentage of sites with a small enterprise activity. This figure will be used to approximate the percentage of households with MSE activities and to weight data for extrapolating survey results (see Chapter Five and Appendix 9).

Record information about a site even when no one at that site has a business.

Convention 3

How do enumerators deal with a closed site? There are two varieties of closed sites. The first variety is a household where no adult is home to answer questions. It cannot be marked as a "site with no activity," because the enumerator cannot ask whether they have any businesses or not. The second type of "closed site" is a location where the respondent refuses to be interviewed. For these situations, there is a second space provided in the upper right-hand corner of the questionnaire to *tally the number of sites with no response*, usually titled "closed sites." This count is important at the end of the survey when the researcher calculates the number of enterprises expected for the entire population. For example, if enumerators could only interview half of the sites in a given area, than it can be expected that they undercounted the true number of MSEs in that area by half.

Tally the number of sites with no response.

Convention 4

If the business is closed but the enumerator can tell what kind of a business it is, fill in the questionnaire as completely as possible.

How do enumerators record that they found something that they know to be a business, but the business was closed so they could not conduct an interview? *If the enumerator can tell what kind of a business it is, he fills in the questionnaire as completely as possible.* He will at least be able to fill in business location and business type. If, however, he cannot identify the type of business, he again puts a tick in the space for "closed sites" in the upper right-hand corner of the questionnaire, as in the third convention.

Convention 5

How to code an entrepreneur who both makes dresses and sells groceries.

Suppose the enumerator finds a woman who both makes dresses and sells basic groceries from her home, working on the garments between grocery customers. How does the enumerator decide whether this is one diversified business or two separate businesses? The rule here is based on time and place. *If two activities are undertaken at the same time and in the same place, as in this case, then the person has one diversified business, and only one line should be filled in on the survey. Had the two activities been undertaken in the same place but at different times (say the woman only made dresses when the grocery was closed), then they would have been considered two businesses and two lines would have been filled in on the survey. Had the two activities been undertaken at the same time but in two different places (say the woman ran the grocery in the neighbor's home and the dressmaking from her home), then they would also have been considered two businesses and two lines would have been filled in on the survey.*

Convention 6

In a multiactivity business, determine the code by sales value.

If it is decided that this is a single business with more than one activity, which code does the enumerator give it to identify "enterprise type"? The convention is to give it the code of the activity that is more important in sales value. For the grocer-dressmaker mentioned above, dressmaking brought in fewer customers but generated more money in sales. Therefore, the business would be coded as "dressmaking," and would be considered a manufacturing activity.

Convention 7

If a business has both a production and a sales branch (for instance, the making and retailing of shoes), should the business be considered

manufacturing or trade? The convention is that *if the person only sells items she produced herself, then the business is considered a manufacturing activity*. If she also sells some items she has purchased, the convention shifts. *If purchased items dominate in value of sales, it is a commerce activity; if self-produced items dominate in sales value, it is a manufacturing activity.*

How to distinguish trade from manufacturing.

Convention 8

If the enumerator meets someone who is selling agricultural produce, is this an agricultural activity that should not be enumerated, or a trade activity that should be enumerated? Here, the convention is simple. If the person produces the agricultural products himself, it is an agricultural activity and should not be included in the survey. *If the person buys the agricultural goods from the producer for resale, then it is a commerce activity and should be included in the survey.*

To be included in survey, person must buy agricultural goods for resale; agricultural producers are out.

Choosing the Unit of Analysis

The survey's door-to-door method guarantees full coverage of MSEs, and on-site interviewing guarantees that each MSE will only appear once in the survey. This much-venerated result has one drawback, however. With on-site interviewing, it is impossible to link enterprises from the same household that occur at different sites. Therefore, it is impossible to get a composite picture of the various activities a single household undertakes.

The only unit of analysis that is consistent with the on-site interviewing rule is the enterprise. Therefore, each row of the questionnaire is a set of questions about a single MSE in a given site. If there is more than one MSE in a site, then one row of the questionnaire is filled in for each MSE. A column is added to the questionnaire where each enterprise is given a unique enterprise identification number.

The survey also collects information about the proprietor, often so that questions relating to women-owned businesses can be specifically addressed. Therefore, it is important that MSEs with the same proprietor be identified together on the questionnaires and in the computer. When the closed enterprise questionnaire is used, it is also important to be able to link a given proprietor's open and closed enterprises for analysis. Therefore, another column is added that gives the proprietor a unique identification number as well. Details on how to use these columns appear in Appendix 5.

EXTENSIONS OF THE BASELINE SURVEY

Some of the basic baseline questionnaire's greatest strengths are its careful choice of variables, limited size, and consistency from survey to survey. However, these three characteristics mean that the survey is not as adaptable to the diverse information needs of its clients as a more open-ended survey approach. To make the baseline survey more adaptable to the client's interests, supplemental survey instruments have been added to the survey that provide more information without altering the content or approach of the basic baseline questionnaire. Two types of survey extensions are discussed below.

The Supplementary Questionnaire

One extension of the basic baseline questionnaire is the supplementary questionnaire, which is given to a subset of baseline survey respondents. It uses the more traditional questionnaire format of questions written out in full, and blank spaces in which to record answers. An illustrative supplementary questionnaire can be found in Figure 3.

The supplementary questionnaire's greatest advantage is that it provides an opportunity to focus on the particular information needs of the client. In South Africa, for example, the client was in the process of re-evaluating its credit programs, and requested a supplementary questionnaire that focused on credit needs and uses. In other cases, the supplementary questionnaire included general questions about business conditions and constraints to help identify opportunities for small enterprise assistance. Although these questions could be relegated to subsequent research efforts, the supplementary questionnaire takes advantage of the fact that a survey team is already in the field.

Supplementary questionnaires take resources away from the basic baseline survey.

A disadvantage of a supplementary questionnaire is that sampling cannot be random, because prior to completion of the baseline survey, there is no sampling frame from which to draw a random sample. This means that the results cannot be accepted as representative of the population. Another disadvantage is that supplementary questionnaires take resources away from the basic baseline survey. In the Lesotho case, it was estimated that for every two-page supplementary questionnaire that was administered, the number of basic questionnaire interviews went down by five. So for every 200 supplementary questionnaires administered, basic questionnaire coverage dropped by 1,000 observations. This

Figure 3

ILLUSTRATIVE SUPPLEMENTARY QUESTIONNAIRE

Enumerator: _____ Page: _____ Proprietor Sequence #: _____
 Supervisor: _____ Date: _____ Enterprise Sequence #: _____
 Enumeration Area: _____ ID #:  Proprietor Name: _____

I. CHARACTERISTICS OF THE ENTREPRENEUR AND HOUSEHOLD

1. What is your age? _____ years ()
2. What is the highest level of school that you have completed? _____ ()
 Codes: 1) No school; 2) Primary school; 3) Secondary school or High School;
 4) University; 5) Other _____
3. What did you do before you started this business? _____ ()
 Codes: 1) Ran another business; 2) Worked in another business; 3) Was Unemployed;
 4) I was too young to work before I started this business; 5) Other _____
4. For how many years have you been engaged in this type of business, including this one? _____ ()

II. BUSINESS HISTORY

5. a) What was the largest number of workers your business has ever had, including:
- | | | | | | |
|---------------|-------|-------------|-------|-----|-----|
| Proprietor(s) | _____ | Family | _____ | () | () |
| Hired | _____ | Apprentices | _____ | () | () |
- b) In what year did this occur? 19 _____ ()
6. What was the principal source of the money you used to start this business? _____ ()
 Codes: 1) Family/Personal savings; 2) Loans/gifts from family or friends; 3) Formal credit institutions;
 4) Moneylenders; 5) Risk Capital Funds; 6) Other _____

III. BUSINESS OPERATION

7. a) Does your business have a high and a low season? 1)Yes 2)No ()
- b) IF YES: When is the high season and how many days/month and hours/day does your business operate? What about during the low season?
- | | Which Months? | Days/Month | Hours/day | () | () |
|--------------|---------------|------------|-----------|-----|-----|
| High Season: | _____ | _____ | _____ | () | () |
| Low Season: | _____ | _____ | _____ | () | () |
- c) IF NO: How many hours per day does this business usually operate? _____ ()

Figure 3 -- Continued

IV. PROBLEMS AND CONSTRAINTS

8. Who are your major business competitors? _____ ()
 Codes: 1) No competitors; 2) Businesses located nearby; 3) Businesses located elsewhere; 4) Public Enterprises; 5) Other _____
9. Did you face any problems when you first acquired this business? ()
 1) Yes 2) No (IF YES: LIST MAJOR TWO IN ORDER OF IMPORTANCE)
 1st _____ ()
 2nd _____ ()
10. a) Have you ever experienced a period of major growth in your business? 1) Yes 2) No ()
 b) IF YES: in what year did it occur? 19 _____ ()
 Did you face any problems during this time? 1) Yes 2) No ()
 (IF PROBLEMS: LIST MAJOR TWO IN ORDER OF IMPORTANCE)
 1st _____ ()
 2nd _____ ()
11. Are you currently facing any problems in this business? 1) Yes 2) No ()
 (IF YES: LIST MAJOR TWO IN ORDER OF IMPORTANCE)
 1st _____ ()
 2nd _____ ()
12. In your perception, how have the following changed over the last five years?
 a) The overall demand for products like yours? _____ ()
 b) The number of businesses just like yours in your locality? _____ ()
 c) The volume of your own business? _____ ()
 Codes: 1) Much increase; 2) Little increase; 3) No change; 4) Little decrease; 5) Much decrease; 6) Do not know
13. a) Have you had any organized training or non-financial assistance for your business activities? 1) Yes 2) No ()
 b) IF YES:
Source of training Type of training Length of training

14. a) Have you ever received loans for your business? 1) Yes 2) No ()
 b) IF YES:
Source of credit What was credit used for? Amount of loan

V. INCOME

15. What part of your household's total income comes from agriculture? ()
 Codes: 1) More than half; 2) Less than half; 3) About half; 4) None

--THANK YOU FOR YOUR ASSISTANCE!--

meant that fewer clusters were included in the survey, which in turn reduced survey precision. To limit this problem, the supplementary questionnaire is usually restricted to two pages and to no more than 600 interviews. Finally, adding a supplementary questionnaire to the baseline survey exercise requires additional time for the task of questionnaire design and pretesting, enumerator training, data coding and entry, data cleaning, and analysis.

The Closed Enterprise Questionnaire

Another extension of the baseline survey is the closed enterprise questionnaire. This questionnaire collects information on previous businesses undertaken by people in the sampled areas. It was first used in the 1990 Kenya survey to analyze business growth and longevity and reasons for business closure — topics that extend the focus on business dynamics beyond that found in the basic questionnaire.⁵ Specifically, the closed business questionnaire collects the following information:

- Size of business at start (in number of workers);
- Size of business at close (in number of workers);
- Peak size of business (in number of workers);
- Age of business at closure;
- Reason the proprietor closed the business;
- What the proprietor is doing now; and
- Location of business.

The closed enterprise questionnaire uses the same matrix format as the basic questionnaire, in which each column is a question and each row is an interview. Codes are again put on the bottom of the questionnaire. Using the matrix format keeps the questionnaire short and coding and data entry simple, and the similarity to the basic questionnaire makes enumerator training on the closed enterprise questionnaire easy. An illustrative closed enterprise questionnaire appears in Figure 4; a description of its contents can be found in Appendix 6.

The closed enterprise questionnaire is administered at every site, whether or not the respondents have a current enterprise. This avoids any problem with random sample selection that appeared with the supplementary questionnaire, and allows the data to be considered representative of the entire population. The closed business survey does, however, have the disadvantage of requiring more survey resources, since it requires more time in the field. If additional resources are not available, the closed business questionnaire will draw resources away from the basic questionnaire.

The closed enterprise questionnaire is administered at every site.

ILLUSTRATIVE CLOSED ENTERPRISE QUESTIONNAIRE

Region: _____ Date Completed: _____ Page # _____
 District: _____ Date Filled: _____ Enumerator: _____
 Administrative Area: _____ Date Entered: _____ Supervisor: _____
 Date Verified: _____

1	PROPRIETOR SEQUENCE #	
2	ENTERPRISE SEQUENCE #	
3	PROPRIETOR ID #	
4	CLOSED ENTERPRISE ID #	
5	CLOSED ENTERPRISE TYPE	
6	CLOSED ENTERPRISE CODE	
7	LOCATION TYPE (a)	
8	YEAR ENTERPRISE STARTED	
9	# WORKERS AT START (incl. proprietor, family, paid, and apprentices)	
10	YEAR ENTERPRISE CLOSED	
11	# WORKERS AT CLOSE (incl. proprietor, family, paid, and apprentices)	
12	HIGHEST NUMBER OF WORKERS (incl. proprietor, family, paid, and apprentices)	
13	SOLD PRODUCT TO WHOM (b)	
14	NATURE OF INPUTS (c)	
15	REASON ENTERPRISE CLOSED	
16	CLOSURE CODE (d)	
17	WHAT ARE YOU DOING NOW?	
18	CURRENT ACTIVITY CODE (e)	
19	PROPRIETOR GENDER (f)	
20	PROPRIETOR CATEGORY	
21	PROPRIETOR'S NAME, NICKNAME, CONTACT'S ADDRESS	
22	HOW-0 / SEX=1) JOB NI TILL FOR DID (WORKER/OWNER)	

Code
 a) 1. Homebased 2. Traditional market 3. Commercial district 4. Roadside, not in commercial district 5. Mall
 b) 1. Individuals 2. Urban commercial enterprises 3. Urban manufacturing enterprises 4. Rural commercial enterprises 5. Rural manufacturing enterprises 6. Export 7. Other
 c) 1. Multiphased own inputs 2. Bought unprocessed inputs 3. Bought semi-processed inputs 4. Bought finished products for resale 5. Other
 d) 1. Market problems 2. Personal reasons 3. Shortage or excess of working funds 4. Shortage or excess of stock or raw materials 5. Got a job 6. Started another MSE 7. Legal trouble/Over-indebtedness 8. Other
 e) 1. New MSE 2. Works for someone else 3. Nothing 4. Other
 f) 1. Female, one proprietor 2. Male, one proprietor 3. Female, more than one proprietor 4. Male, more than one proprietor 5. Mixed-gender joint proprietorship

Figure 4

INSTRUMENT DESIGN: KEY POINTS

- **Don't foster false expectations about what information the survey can provide.**
- **Don't overload the basic questionnaire and never go to a two-page basic questionnaire.**
- **Don't add supplementary instruments unless additional resources are forthcoming.**
- **Don't try to capture flow variables on the basic questionnaire.**
- **Do follow the on-site interviewing convention to guarantee maximum coverage without double-counting.**
- **Do include the necessary columns for data management on all survey instruments, particularly space for identification numbers.**
- **Do make sure codes are all-inclusive and do not overlap.**

Notes

1. **Box 1 provides information on what the baseline survey can and cannot do.**
2. **The information collected on an additional questionnaire does not duplicate that collected by the basic questionnaire. Therefore, it is important to have a way to tie the two questionnaires together so that the data from one can be combined with data from the other for analysis. The simplest method is as follows. First, the same proprietor identification number is assigned to both questionnaires administered on the same site. Second, a column (usually placed at the end of the basic questionnaire) is coded as "0" if the additional questionnaire is NOT administered, and as "1" if it IS administered. After all the basic questionnaire data is entered into the computer, the analyst can select only those cases with "1" in the last column. This will produce a data file including only those proprietors who answered both questionnaires. Those cases are**

then merged with the file of data from the additional questionnaire, matching the cases by proprietor identification number. In SPSS/PC+, the software program most commonly used for the baseline survey data, the procedures are "SELECT IF," followed by "JOIN MATCH."

3. The general convention defines vending, retailing, and wholesaling on the basis of the quantity of goods held for sale. If the proprietor has a limited display of goods for sale but no stock to replenish the display as people buy, the person is said to be vending. If the person has enough stock to both display and replenish the display as customers buy, then the person is retailing. If the person has sufficient stock to supply businesses engaged in retailing, then the person is wholesaling. This quantity-based definition differs from that typically used, which is based on permanence of structure or mobility of the business.
4. ISIC codes are not used as the questionnaire's enterprise codes because two-digit ISIC codes are too broad to be useful for later analysis, while the more specific four-digit ISIC codes are too cumbersome for enumerators to use. Ideally, the enterprise coding system used by the enumerators should bear some semblance to the ISIC codes so that conversion at the analysis stage can be easier and more logical.
5. The closed enterprise questionnaire was originally designed to get around the problem of "sample censoring" that is inherent in the basic questionnaire on issues of MSE dynamics. Specifically, the basic questionnaire data cannot measure precisely the longevity or lifetime growth of an enterprise since the final period of the enterprise's life is not yet observed. Collecting data on closed businesses overcomes the censoring problem by capturing information on enterprises from birth through closure.

CHAPTER FIVE

DATA MANAGEMENT AND ANALYSIS

To this point, the manual has not dealt with the process by which questionnaires go from being filled out in an interview to being entered into a computer file and ultimately analyzed to provide answers to questions of interest. In this chapter, the proper handling and storage of questionnaires and questionnaire data will be discussed, followed by an explanation of the proper documentation and management of computer files. After these issues are considered, data cleaning and analysis will be taken up, and a final section deals with the important matter of assigning weights to the data so that the sample describes accurately the population of MSEs in the country.

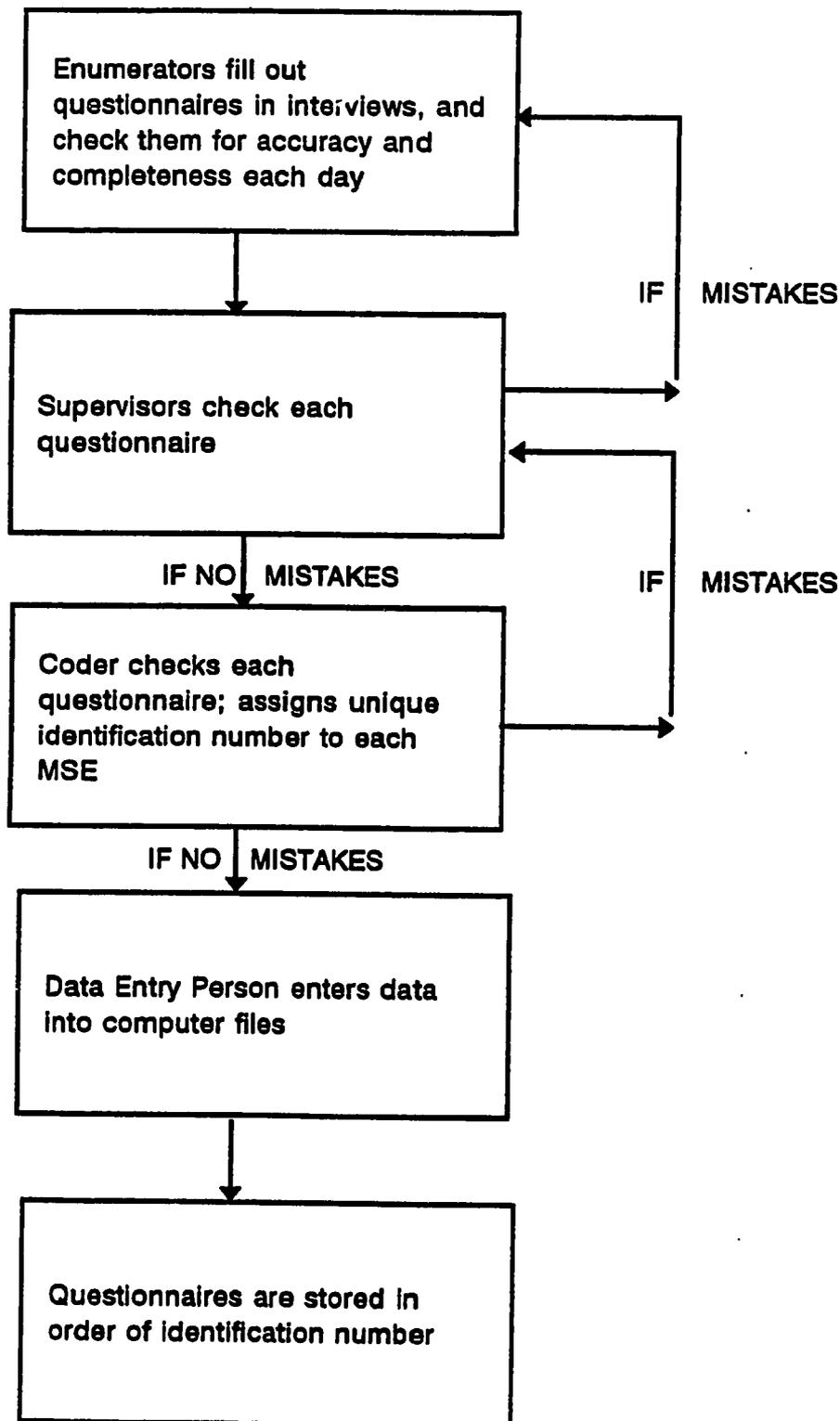
THE PAPER TRAIL

Over the course of a typical baseline survey, thousands of interviews will take place. When the smoke clears, there will be a huge number of pieces of paper, especially when most surveys include at least two and sometimes three survey instruments. It is imperative that some system exist for dealing with this blizzard of paper. One possibility for such a system is illustrated in Figure 5. This system, or variations on it, has been used in all MSE surveys. As interviews are conducted, questionnaires are completed. At the end of each day, or before the next begins, each enumerator should examine his or her own questionnaires, making sure that all spaces have been filled and that the information for each interview is consistent. Having done this, the enumerators turn in all questionnaires to the supervisor. It is a responsibility of the supervisor to scrutinize each piece of information on each questionnaire. Should the supervisor identify any omissions or mistakes, she returns the form to the enumerator for correction. When questionnaires are acceptable to the supervisor, they are sent to the coder, who is usually in the main office. This exchange can be accomplished in person if the teams are working near the main office, or by bus or mail otherwise.¹

A system must be created to deal with the blizzard of paper.

After the coder receives a batch of questionnaires, he carefully checks each one, taking care that there is no missing information, that the proper codes have been used, that the data for each MSE is consistent, and that the writing is legible. The coder then assigns each proprietor and

FIGURE 5
THE PAPER TRAIL



enterprise unique identification numbers, writing these numbers on the questionnaire. If errors are found, the faulty questionnaires are returned to the supervisor of the enumerator who conducted the interviews. The coder should keep track of any persistent problems and report them to the researcher. Questionnaires acceptable to the coder are passed along for data entry. After questionnaires have been entered and verified, the data enterer stores them in a safe place, usually in numerical order by unique identification number.

DOCUMENTING AND MANAGING COMPUTER FILES

The advent of the personal computer has brought many advantages to survey work. Among these are the ability to safely store large quantities of data in a small space, and the capability to document the data accurately so that other researchers can use the data easily in the future. These advantages, however, are not automatic: they must be built into the survey plan.

Regardless of the data entry package used in the survey, it is imperative that all computer files be well documented. Without proper documentation, the data will become useless to all future researchers. Documentation means defining all variables, and providing value labels for any categorical variables. For example, a variable may be entered into the computer as GENDER. To a researcher not familiar with the survey, this could mean gender of the workers in the MSE, gender of the proprietor, or something else. This variable must be defined carefully, either internally in the computer file or on a list that is made available with the data. Suppose further that GENDER is a categorical variable. That is, the two genders are assigned numerical values, such as 1 for females and 0 for males. These value definitions must also be recorded, to avoid future mix-up.

Although documentation is simple in most software packages, it can be time-consuming. It is best if the documentation is complete before any data are entered into the file, so the researcher should allow time in the early days of the survey for this task.

Documentation can be time-consuming.

The researcher also must devote some thought to how the data files are to be managed. A recurring nightmare for many researchers is that somehow all the entered data are erased or lost. There are some simple precautions that, if carefully followed, will prevent most such problems.

Each data file should be stored on at least two disks. Usually, this will include a floppy disk and the computer's hard drive. The data files should be saved to both locations at least twice a day.² Finally, the area in which the computer is situated should be secure.

CLEANING DATA

The paper trail described above includes checks of the data at every step. In addition to this, in the data entry phase, data are entered twice to reduce keypunch errors. In spite of these important measures, there will always be errors in any baseline survey data file. This section discusses ways to rid the data of errors, a process often called cleaning.

In many software packages, it is possible to design the data entry package so that it will identify errors as they are entered. For example, variables can be assigned acceptable value ranges. The computer can be told that the values for the variable for proprietor age, for instance, cannot be less than 10 or more than 85. Should the entered value fall outside this range, the questionable case will be identified. Assigning consistency rules to a file is also straightforward. As an example, refer again to the illustrative basic questionnaire in Figure 2. Note that after the total number of workers (inclusive of the proprietor) is established, the enumerator asks how many of this total are female. It is impossible under any circumstances for the number of females to exceed the total number of workers. At least for the SPSS/PC+ data entry package, when such rules are built into the file, all violations will be reported.

Designing the data entry file with rules and ranges included is an important way to clean a file.

Designing the data entry file with rules and ranges included is an important way to clean a file. However, some mistakes may elude even these measures. A final way to identify mistakes in the file is to compute some descriptive statistics of the data. For continuous variables, such as firm age or number of workers, studying the means of each variable and the minimum and maximum value each takes on is the most efficient way to catch outlying values. For example, if the minimum value the variable for number of workers takes on is zero or negative, there surely was a mistake made. For categorical variables, it is useful to compute a frequency distribution for each variable.

ANALYZING THE DATA

Once the data have been thoroughly cleaned, the task of analyzing the data can begin. Because the specific analyses that are carried out depend on what questions are of interest and who will be reading the survey report, this section will limit the coverage of this topic to a general statement.

Most of the time, it is important to present some of the basic characteristics of the survey data. These descriptive statistics include the means and standard errors of continuous variables, and the frequency distributions of categorical ones. Each of these variables can be stratified in any number of ways. Some examples include splitting up the sample by location of firm (for example, urban versus rural), by sector of activity, or by gender of the proprietor. Various means tests, such as the t-test, are available to ascertain statistically whether the average value of some variable for two groups is the same. Most statistical packages include a dizzying array of options, including cross-tabulation of two variables, analysis of variance, and regression analysis. In general, the simplest analyses are the best places to start as they often provide the most compelling insights.

Start with the simplest analyses — they often provide the most compelling insights.

WEIGHTING DATA FOR COUNTRYWIDE EXTRAPOLATION

All MSE countrywide baseline surveys involve going door to door to identify and interview enterprises in a sample of small areas. These areas are grouped into several strata, as discussed in Chapter Three. However, the final group of enterprises sampled is not necessarily representative of all MSEs in the country. For example, in most developing countries, the vast majority of the populace lives in rural areas. Unless the sample includes a large percentage of rural clusters (which would be prohibitively expensive), any statements made about MSEs nationwide would reflect a bias toward urban enterprises. To avoid such a bias, data from each stratum are weighted by the overall size of the stratum.

Data from each stratum are weighted by the overall size of the stratum.

Coming up with appropriate weights depends mainly on two estimates: the percentage of sites in each stratum engaged in MSE activity, and the total number of sites in existence in each stratum. The more accurate are these estimates, the more likely it is that the weights

calculated are appropriate. What means are available to come up with these estimates? Information regarding the proportion of sites engaged in MSE activity comes directly from the survey. As enumerators cover clusters, they keep track of the status of every household and shop location they visit. There are three possibilities: a site can contain an MSE, can be closed to the interviewer, or can have no MSE activity.³ By dividing the number of sites with MSE activity by the sum of sites with activity and with no activity, an estimate of the proportion of open sites with MSE activity can be calculated.⁴ This estimate of the proportion of sites involved in MSE activity is made for each cluster. To come up with an estimate for the stratum, the cluster proportions are averaged. In Zimbabwe, the information regarding these proportions is presented in Table 1.

TABLE 1

**PROPORTION OF SAMPLE INVOLVED IN MSE ACTIVITY
(Cluster Averages)**

	Urban	Rural
Sites with MSEs	3,760	1,815
Sites with no MSE activity	6,146	2,629
Total sites with information	9,906	4,444
Percentage of sites with MSEs	37.96%	40.84%

Next, the problem of estimating the number of sites in each stratum is tackled. This one is more troublesome, since this information is not generated by the survey, but comes from some outside source. Often, this source is the most recent population census. This is an especially good choice if the clusters were based on census enumeration areas. The drawback is that the census only provides the number of households per stratum, not the number of sites. Still, since the vast majority of sites are households, using the census estimate is the best way of approaching this step. In Zimbabwe, the census figures provided the information that in 1991 there were 837,513 urban and 1,522,125 rural households.

To arrive at an estimate of the total number of MSEs per stratum, the proportion of sites with MSE activity (as calculated in Table 1) is multiplied by the total number of sites. The weights are computed by simply calculating the share each stratum has of the estimated total number

of enterprises. In the Zimbabwe case, these weights turn out to be .338 and .662 for urban and rural strata, respectively. In other words, 33.8 percent of the estimated total number of MSEs are located in urban areas, while the remainder, 66.2 percent, can be found in rural areas.

Once these weights have been calculated, they can be used in the analysis to arrive at figures that reflect what the survey would have found if all MSEs in the country had been enumerated. If one is interested, for example, in the percentage of proprietors made up of females, the percentage of female proprietors in each stratum is multiplied by the above weights and summed. Because 76.2 percent of proprietors in urban areas are women, and the analogous number for rural areas is 62 percent, the weighted average is $(76.2)(.338) + (62.0)(.662)$, or 66.8 percent.

Weighting is a necessary step if the researcher is interested in making statements about the findings of the survey for the country as whole. Although the procedure used to calculate the weights involves estimates and assumptions, as long as care is taken a much better understanding about MSEs in the country is possible.

Weighting is a necessary step to make statements about findings countrywide.

DATA MANAGEMENT: KEY POINTS



- Do invest time setting up a paper trail and seeing that the system functions well.
- Don't fail to document all computer files.
- Do arrange a system whereby computer files are backed up at least once a day.
- Do take advantage of any data cleaning features in the data entry software.
- Don't forget to gather all information necessary for weighting data before leaving the country.

Notes

1. The speed and reliability of the mail should be considered. In Zimbabwe, inexpensive overnight express was available.
2. In some countries, power surges and power failures are common place. In such countries, a surge protector should be bought or rented, and data should be saved to disk much more frequently.
3. That is, no one is home or the respondent refuses to grant an interview.
4. The difficulty lies with guessing what proportion of the closed sites have MSEs. In a survey of Kibera, Kenya, Parker and Dondo (1991) resurveyed households found to be closed when first visited. They found that roughly the same proportion of closed households had MSEs as open households. The weights that are calculated depend on this being the case.

CHAPTER SIX

DISSEMINATION OF RESULTS

As we have made clear, seeing a survey through from the design phase to data collection to analysis is a difficult and time-consuming job. However, the work is not yet complete. One crucial task remains: getting the results of the survey into forms that are useful to the client and to other interested parties. The larger the group that is aware of the survey, the more useful is the exercise. Fortunately, there are two great advantages to the system described in the preceding chapter. The first is that because the process is streamlined and because there are seldom delays in getting the data entered, it is possible to produce preliminary results within a week of the end of field work, and final results within two months. This is a happy surprise for most clients, who are used to waiting for results for many months or even years. A second advantage is that the data generated by the survey can be shared with other researchers as soon as the data are cleaned.

WRITTEN PRODUCTS

The survey results will become known to the greatest number of people by means of written products. These come in two forms: a preliminary description of the findings, and a more extensive final report.¹ What exactly is the nature of these two important documents?

The preliminary report provides a rough idea of survey findings very quickly.

The preliminary report provides a rough idea of survey findings very quickly. Indeed, this report is usually given to the client as an oral briefing by the researcher before she even leaves the country in which the survey took place. Because the data are continually entered as the questionnaires come in from the field, by the time the field work ends almost all of the information is in the computer. This means that some simple features of MSEs in the sample, such as the sectoral distribution, average size, and typical locations, can be ascertained quickly, as can some of the characteristics of the proprietors. It is important, however, to remind the audience of the limitations of these findings. Specifically, the data have not yet been cleaned or weighted. This means that some inaccuracies are still present and that statements about MSEs in a nationwide context are not yet possible. With these important caveats, however, the preliminary report is an important part of the survey process.

Some weeks after the data are cleaned and analyzed, the principal researcher is responsible for producing a report summarizing the findings of the survey. The shape of this document is largely dependent on the preferences of the client. Most clients will want a document aimed at the nontechnical reader. Mead, Fisseha, and McPherson (1991) present a suggested outline for the baseline survey report:²

BOX 5

SAMPLE BASELINE SURVEY REPORT

INTRODUCTION

The national context; MSEs in this country; goals of the survey
Institutional, legal context

SURVEY APPROACH

Coverage
Methodology
Magnitude: staff of survey; timing; number of households and enterprises visited (before extrapolation)

SURVEY FINDINGS

(after extrapolation; all with cross-country comparisons, when possible and appropriate; all by stratum, when possible and appropriate)

Magnitude: numbers of enterprises, levels of employment
Enterprise and employment densities
Size distribution of enterprises
Sectoral breakdown, industrial structure
Labor force characteristics; gender dimensions
Patterns of change
Other MSE characteristics of interest from survey

SUMMARY, IMPLICATIONS, AND CONCLUSIONS

Although the individual researcher must decide with the client on the exact content of the report, this outline provides a reasonable starting point. Examples of recent survey reports include Liedholm and McPherson

(1990), Fisseha (1991), Parker and Dondo (1991), Fisseha and McPherson (1991), and McPherson (1991).

MAKING THE DATA AVAILABLE

After the survey process has wound down and the final report has been written, the principal researcher may still be responsible for the dissemination of survey data. The data technically belong to the client who may distribute them or not according to its wishes. The intentions of the client along these lines should be discussed in the design phase of the survey. Often, the client is the U.S. Government. In this case, the researcher is obliged to share the data with any U.S. citizen. If it is the desire of the client that the data be made generally available, it is the responsibility of the principal researcher to provide complete and well-documented computer files to all interested parties. This sharing of data makes possible a more complete analysis of MSEs, certainly a desirable outcome if the ultimate goal of the survey is to improve the state of knowledge regarding these enterprises.

Baseline surveys have several uses.

OTHER USES FOR BASELINE SURVEY DATA

The MSE Baseline Survey has provided a useful way to identify the number and characteristics of MSEs in a country. But baseline surveys have other potential uses beyond this immediate purpose, five of which are mentioned briefly below.

First, a baseline survey provides a sampling frame for future work, including monitoring and evaluation efforts and also additional intensive studies. As an example of the latter, recent MSE Baseline Surveys in Swaziland and Niger were used to form a sampling frame for a study by the Organization for Economic Cooperation and Development of the impact of the regulatory environment on MSEs.

Another use of basic data is that if the survey is readministered regularly, it can provide macro-level time series data on MSEs. This permits two specific possibilities. First, the data can track the progress of the entire MSE sector when looking for the effects of diffused forces on small businesses, such as when studying the effects of macroeconomic policy changes on MSEs. Second, the data give a periodic look at how the

general MSE population is faring relative to any assisted sectors. In this case, the entire baseline data set serves as a control group, while data on assisted sectors make up the experimental group.

More targeted analyses of the basic data make up the third category of potential uses. These data sets contain a great deal of information that can be examined from other angles by many different researchers. One recent example of this possibility is a study of gender issues by Downing and Daniels (1992), based on baseline surveys from Lesotho, South Africa, Swaziland, and Zimbabwe.

Fourth, as similar surveys are completed in an ever-widening number of countries, they provide an excellent resource for cross-country comparisons. Although this is usually insufficient grounds for undertaking a baseline survey, it is certainly an argument for those already set on studying MSEs to use similar techniques.

Finally, data generated by the baseline survey can be used to check the accuracy of other data sources, such as government statistics on MSEs. Should the other data sources be comparable to the results generated by the baseline survey, it may be possible to use these other data sources for national planning, monitoring and evaluation of projects, and studies of the impacts of policy changes on the MSE sector.



RESULTS: KEY POINTS

- **Don't miss the opportunity to present preliminary findings to the client before leaving the country.**
- **Do consult with the client regarding the nature of the survey report.**
- **Do make available all survey data on disk, subject to the client's approval.**
- **Do remember the many uses of baseline survey data beyond the immediate needs of the client.**

Notes

1. **Recent survey reports are listed in the reference section of this manual.**
2. **If the survey includes the closed enterprise or the supplementary questionnaires, the report could be extended to reflect the information generated by these.**

CONCLUSIONS

Over the last several decades, there has been an increasing awareness of the importance of micro- and small-scale enterprises to the economies of many countries of the world. As a result, surveys of MSEs in developing countries have proved to be valuable tools for policy makers, assistance institutions, and researchers. But, to be valuable, surveys of MSEs must be able to find all businesses regardless of whether they are mobile or home based, registered or not. The MSE Baseline Survey method presented here compiles a complete list of MSEs, penetrating households as well as more traditional places of business, and counting all enterprises regardless of size or registration.

This manual represents a great deal of accumulated experience in MSE Baseline Surveys in developing countries. It seeks to give instruction on such important topics as setting up administrative and financial arrangements, effective survey management, designing an appropriate and feasible sampling procedure, hiring and training of survey personnel, using and adapting the several survey instruments, appropriate methods of data management, and making survey findings generally available. The manual intends to alert future researchers to the pitfalls that inevitably accompany these exercises.

It is hoped that this manual provides a way to share this experience with other interested parties who may be searching for ways to study MSEs around the world. The application of the methods described in this manual, with adjustments made for the particular circumstances in the country at hand, can make this powerful tool available to an ever broader array of countries and institutions.

Previous Page Blank

REFERENCES

Casley, D.J., and Lury, D.A. *Data Collection in Developing Countries*. Clarendon Press, Oxford, 1987.

Cochran, William G. *Sampling Techniques*. John Wiley and Sons, New York, 1977.

Konijn, H.S. *Statistical Theory of Sample Survey Design and Analysis*. North Holland Publishing Company, Amsterdam, 1973.

Mead, Donald C.; Fisseha, Yacob; and McPherson, Michael A. "Baseline Surveys of Micro and Small Enterprises: An Overview." GEMINI Working Paper No. 16, 1991.

Survey Reports

Daniels, Lisa, and Fisseha, Yacob. "Micro- and Small-Scale Enterprises in Botswana: Results of a Nationwide Survey." GEMINI Technical Report No. 46, 1992.

Fisseha, Yacob. "Small-Scale Enterprises in Lesotho: Summary of a Country-Wide Survey," GEMINI Technical Report No. 14, 1991.

Fisseha, Yacob, and McPherson, Michael A. "Small Enterprises in Swaziland: Results of a Country-Wide Survey." GEMINI Technical Report No. 24, 1991.

Liedholm, Carl, and McPherson, Michael A. "Small-Scale Enterprises in Mamelodi and Kwazakhele Townships, South Africa: Survey Findings." GEMINI Technical Report No. 16, 1991.

McPherson, Michael A. "Micro and Small-Scale Enterprises in Zimbabwe: Results of a Country-Wide Survey." GEMINI Technical Report No. 25, 1991.

Parker, Joan, and Dondo, C. Aleke. "Kenya: Kibera's Small Enterprise Sector - Baseline Survey Report." GEMINI Working Paper No. 17, 1991.

Previous Page Blank

APPENDIX 1

THE HISTORY OF BASELINE SURVEYS

Prior to 1974, small enterprise surveys were confined to urban areas, and they counted only those enterprises that could be identified easily. Surveys were usually limited to manufacturing activities. In the rural areas, studies focused on agricultural activities, with little or no attention given to non-agricultural activities.

In 1974, the first MSE Baseline Survey was carried out in Sierra Leone. That survey took several innovative steps. First, it counted all businesses, regardless of their visibility, by going from house to house in search of enterprises. Second, it acknowledged that rural households also had nonfarm activities that provided income and employment. The survey extended the search for enterprises into the rural households, thereby providing a national picture of both urban and rural MSEs. Third, it used a lean questionnaire designed to cover a large sample of enterprises at low resource costs.

Over the rest of the 1970s, the baseline survey was considered a way to provide a sampling frame for more data-intensive research. But, early in the 1980s, the notion that the baseline was primarily a springboard for other research was given less emphasis, and policy makers began requesting baseline surveys for the information they could provide in their own right. At this stage, the baseline survey became an important instrument for identifying the scope and key characteristics of the small enterprise sector, both rural and urban.

Until the mid-1980s, MSE Baseline Surveys focused exclusively on manufacturing activities. In the Zambia baseline survey of the mid-1980s, the survey was extended to count service and trade income-earning activities as well, which dramatically increased the coverage of that survey. In general, however, manufacturing remained the key focus of the MSE baseline surveys until 1989, when the new A.I.D.-funded Growth and Equity through Microenterprises Investments and Institutions (GEMINI) Project made a decision to give equal attention to trade and service activities.

As the 1980s drew to a close, questions about small enterprises began to shift from a focus on the structure of the sector to more complicated questions about the evolution of the sector. In response to these new information needs, the content of the baseline survey was revised in 1990 to add questions on enterprise dynamics, such as growth and longevity. This is the current generation of baseline surveys discussed in this manual.

Although the baseline survey is a particularly good method for studying small enterprises on a countrywide basis, it has also been used to study specific parts of a country of particular concern. In South Africa, for example, the survey was limited to a study of enterprises in black

Previous Page Blank

townships. In Kenya, the survey focused on enterprises in Nairobi's largest slum settlement. These limited surveys were designed primarily to generate specific information for programs operating in the target areas.

To date, some form of baseline survey has been used in 18 countries. The baseline survey has been especially popular in Africa, where it has been used in Sierra Leone, Zambia, Niger, Upper Volta (Burkina Faso), Egypt, Lesotho, South Africa, Kenya, Zimbabwe, Swaziland, Botswana, and Malawi. In Asia, baseline surveys have been carried out in Thailand and Bangladesh. In Latin America, the list includes Jamaica, Haiti, Honduras, and the Dominican Republic.

The one constant feature of the baseline survey has been its ability to evolve as information needs change. As it has proved itself responsive to information needs, it has become an ever more popular method of research. We can expect new generations of small enterprise baseline surveys in the future that will meet information needs in the same flexible way.

APPENDIX 2

ILLUSTRATIVE BUDGET FOR A COUNTRYWIDE BASELINE SURVEY

Coming up with a budget that will not only make possible a complete and accurate survey, but also satisfy the constraints of the client is a complicated but important task. The principal researcher should involve herself heavily in the presurvey budget-writing process, because she will have to live with the budgetary consequences. The following is an illustrative budget. Of course, each survey will have unique circumstances. For example, the arrangement with the survey administrator, the number of person-days necessary, and the types and numbers of vehicles required will all depend on factors particular to each survey. This being the case, the budget presented below is intended to provide some guidance as to the line items that should be considered, and to the particular information the principal researcher must have access to when writing the budget.

SALARIES

Principal researcher	1 person for 16 weeks @ ___/week
Survey administrator	1 person for 11 weeks @ ___/week
Data enterer	1 person for 8 weeks @ ___/week
Coder	1 person for 9 weeks @ ___/week
Supervisors	2 persons for 10 weeks @ ___/week
Enumerators	16 persons for 9 weeks @ ___/week
Drivers	2 persons for 8 weeks @ ___/week

PER DIEMS

Principal researcher	70 days @ international rate
Supervisors	2 persons: 15 days/each @ local rate
Enumerators	16 persons: 15 days/each @ local rate
Drivers	2 persons: 15 days/each @ local rate

INTERNATIONAL TRAVEL

Principal researcher	1 round trip @ ___
----------------------	--------------------

DOMESTIC TRAVEL/TRANSPORT

Vehicle rental

Field team vehicles 2 vehicles: 8 weeks @ ___/wk
 Researcher's car 1 vehicle: 10 weeks @ ___/wk

Per kilometer charges

Field team vehicles 7900 km @ ___/km
 Researcher's car 3600 km @ ___/km

Petrol costs

Field team vehicles 7900 km @ ___ km/liter @ ___
 Researcher's car 3600 km @ ___ kms/liter @ ___

Domestic airfare¹

Researcher travel 16 round trips @ ___/trip

SUPPLIES/EQUIPMENT

Computer rental 10 weeks @ ___/wk
 Printer rental 10 weeks @ ___/wk
 Field supplies ___?___

COMMUNICATION/PHOTOCOPYING

Reproduce questionnaires	2400 pages @ ___/page
Misc. photocopying	500 pages @ ___/page
International fax/phone	___?___
Domestic fax/phone	___?___
Transfer of questionnaires from field to headquarters	___?___

SPACE RENTAL

Training space	1 week @ ___/wk
Headquarter space	10 weeks @ ___/wk

This illustrative budget assumes eight weeks of field work for two teams, each made up of eight enumerators, one supervisor, and one driver. It also assumes that each field team spends 15 nights on the road during the course of the survey.

Notes

1. For a large country, air travel may be required for the principal researcher or survey administrator during training and fieldwork.

APPENDIX 3

CONTENT OF BASIC TRAINING

Chapter Two addressed many of the most important items in designing and implementing a training session for an MSE Baseline Survey exercise. This appendix extends this information by providing details on the content of basic training.

Administrative matters are always points of concern for the enumerators and are the biggest source of questions until addressed. Therefore it is useful to set aside time early in training to go over the survey work week (how many hours a day and days a week people are expected to work), salary (including payment schedule), expense coverage while on the road, and attendance. Details will vary from country to country, but most survey teams have worked long hours, five days a week, leaving weekends free.

Training starts with a substantive introduction to the survey exercise, which answers the following questions: What is meant by small enterprise or income-earning activity? Why do people care about small enterprises? What are the goals of this survey? Where else has such a survey been done? Answering these questions provides a sense of purpose to the exercise, and gives enumerators the background they will need to identify MSEs and to explain the survey to respondents. The second part of the introduction identifies the participants in the survey and their tasks, including the client, the principal researcher, the survey administrator, the supervisors, the enumerators, the coder, and the data enterer. When possible, the individuals serving in these roles should be introduced by name and background during this session.

Training on the basic questionnaire consumes the bulk of the training time, usually about three days, and can be divided into three categories: classroom lectures, classroom exercises, and field exercises. The following sections give a brief overview of the content of each of these categories of training.

CLASSROOM LECTURES

When introducing the trainees to the basic questionnaire, the obvious first step is to explain the matrix format. Once that is understood, it is useful to take each column on its own and discuss the following:¹

- What is the precise information we are looking for in each column?
- Why is the question important?
- How should the question be phrased to ensure consistency between interviews?

Previous Page Blank

- What answers can we expect?
- Are answers recorded as values or codes? and
- If the answer is coded, where are the codes located and what do they mean?

Once the individual questions have been discussed, the principal researcher needs to explain the sequence in which questions are asked, emphasizing that the order should be followed to reduce the number of skipped questions. It is also important to explain which columns are filled in by observation (for instance, it is usually not necessary to ask the proprietor his or her gender!), and which should be filled in after the interview is over.²

Classroom lectures also cover the survey conventions listed in Chapter Four, including the rule of on-site interviewing. For example, enumerators should be trained to ask: "Are there any income-earning activities in this home?" instead of: "Do you have any income-earning activities?" since the latter is not a site-specific question. Using the conventions correctly will keep survey coverage complete and unbiased. The enumerators need to be trained on how to fill in the upper right-hand corner of the questionnaire.³ Enumerators also need to learn how to classify and order their interviews, both by filling in the information in the upper left-hand corner of the questionnaire, and by giving each interview an identification number.⁴

Finally, the lectures include a description of how to conduct a door-to-door survey, including how to identify the boundaries of the entire sampled area, how to divide the area between enumerators, and how to ensure that coverage within the area is complete.⁵ If some of the supervisors and enumerators have participated in the national census, their experience can be shared with others, though with the caution that some census methods may have been less thorough than the small enterprise survey.

If more than one questionnaire is administered in the baseline survey, most of the above topics need to be covered for each questionnaire. It is helpful, however, to finish training on the basic questionnaire before introducing the other survey instruments. If a supplementary questionnaire is administered, some time will be needed to help enumerators make the shift from the matrix-style questionnaire to a more traditional questionnaire. In addition, because a supplementary questionnaire is given only to a subset of respondents, the enumerators and supervisors must be instructed on which respondents should receive that supplementary questionnaire.

CLASSROOM EXERCISES

Although lectures on the basic questionnaire are a necessary starting point, they are most effective when frequently interspersed with practical exercises that make the lectures concrete and help the trainees raise questions.

Exercises provide opportunities to work on key areas of questionnaire delivery, such as:

- Introducing oneself and the survey;
- Wording questions in an identical way in each interview;
- Listening to answers (plus identifying non-answers or answers that show the respondent has not understood the question);
- Recording answers (using codes and values correctly);
- Questionnaire pacing, flow, and completeness;
- Dealing with hostile or unwilling respondents; and
- Translating questions into other languages.

Many methods have been used for classroom exercises, three of which have proved particularly useful and are discussed here.

The first method is the mock interview. One enumerator acts as interviewer while another acts as respondent. The interviewer introduces himself and the survey to the respondent, gets permission for the interview, then conducts the interview, filling in the questionnaire as he goes. The other trainees follow along, filling in blank questionnaires as the interviewer elicits information. After the interview, there is a roundtable discussion of the style, content, order, and wording of the interview, or of any other issues that emerge. This exercise is first carried out in English (or the language of training), until the principal researcher feels comfortable that the enumerators understand the questionnaire completely. After a translation exercise (discussed below), mock interviews can then be carried out in the survey language, breaking the group into smaller units if several languages will be used in the survey. Because the principal researcher is rarely proficient in these languages, this exercise is most useful if the session can be led by someone who understands both the questionnaire and the languages. In Kenya, for example, the national researcher who had participated in questionnaire design led the sessions that were conducted in Kiswahili.

The second classroom activity is a series of short written exercises, which quiz the enumerators on the content of the questionnaire, the meaning of the questions and codes, and the various survey conventions. An example is provided in Appendix 4. These are collected and corrected by the principal researcher, and areas of misunderstanding are then addressed in the next session. These written exercises, along with performance in the mock interviews, provide the principal researcher with a way to decide which trainees should be selected as enumerators, and who should be selected as coder.

Finally, translation sessions are a helpful activity late in training, once the questionnaire is familiar in English. In small groups, the trainees develop translations for each question, an exercise which brings out any ambiguities in or disagreements about meaning. Once each small group has written down their "best" translation for each question, the entire group reconvenes, shares their translations, and agrees on a common translation. This session is best led by a national who understands both the questionnaire's intent and the local languages. Although the principal

researcher is rarely able to understand the subtleties of the discussion, it is still important for her to sit in on the discussion in case questions arise on the intent of the questions that she can resolve.

FIELD EXERCISES

Once the enumerators and supervisors have become proficient with the questionnaires, it is time to get them acclimated to the field. The first field exercise is usually a visit to a marketplace, where enterprises can be observed without knocking on doors. The purpose of this visit is to practice recognizing types of enterprise, then matching types of enterprise with the enterprise codes, identifying boundaries of enterprises (a task that can be difficult in a crowded market), and discovering the locations of enterprises.⁶ This visit must be carried out in an area NOT included in the sample.

The second and more intensive field activity is a survey pretest. Again, it must be carried out in areas not included in the sample. Before the pretest, field team assignments are made, then each field team is designated a pretest area. Once in the field, the supervisor leads the team in identifying the boundaries of the area and dividing the area between enumerators. Enumerators then use the door-to-door method and administer the questionnaires wherever enterprises are found. This gives the enumerators their best opportunity to identify when they may be unsure of how to move through an area, how to introduce themselves or the survey, or how to ask or record specific questions. Enumerators are also likely to have their first encounters with curious bystanders and unwilling respondents, both of whom they will need to be ready to deal with during field work. Enumerators should be encouraged to keep a running list of any difficulties or questions they have. Supervisors, meanwhile, practice keeping track of their enumerators' progress and whereabouts, and make sure that all assigned areas are covered. The principal researcher should also be in the field during this period, spending time with each team and responding to questions as they arise. After the pretest is finished, the entire group reconvenes to discuss the exercise and to deal with questions or concerns.

Notes

1. This information is typically covered more than once during the training period because it is too detailed for trainees to assimilate the first time around. In addition, it is helpful to have this information written down and distributed to the field teams so that the enumerators and supervisors have a source to refer to if questions arise during training or fieldwork.
2. In addition to proprietor's gender, the business address and type of business location should be filled in after the interview is completed, since they can be completed by inspection. Similarly, although the exact type of business is written down on the questionnaire longhand

during the interview, the business code need not be looked up and filled in until after the interview. Finally, there are a few columns that the enumerator never fills in, which are usually shaded, such as those reserved for a unique interview with ID codes assigned by the coder once the questionnaire has been turned in.

3. The upper right-hand corner of the questionnaire has a place for tallying the number of (1) sites where there was no one at home and (2) sites without businesses. These tallies are discussed further in Chapter Four in terms of instrument design, and in Chapter Five in terms of their usefulness for extrapolating survey results to the entire population.
4. The information collected in the upper left-hand corner of the questionnaire includes a listing of (1) enumerator's name, (2) supervisor's name, (3) where the interviews were done, and (4) when the interviews were done. This information allows the principal researcher to keep questionnaires in order and to keep track of which areas have been covered. In addition, the first two columns of the questionnaire give the enumerator a place to number each interview for a given site (starting with "1" each time he goes to a new site) and for each site in a given area (starting with "1" each time he goes to a new area). This numbering system has several important purposes, and needs to be well understood by enumerators. Further details on this system are given in Appendix 5.
5. In brief, the enumerator is instructed how to follow the supervisor's instructions regarding boundaries; how to coordinate coverage with other enumerators (for example, when a street is a boundary between two enumerator's areas, each has a designated side of the street, but one is named to be responsible for anyone in the street itself); and how to most effectively penetrate the area and identify hidden households within it.
6. Enterprises exist along footpaths and side-streets leading to marketplaces as well as within the market's boundaries. It often surprises enumerators how many MSEs are found in a small geographic area, particularly near marketplaces.

6. Using the following information, fill in the baseline survey questionnaire.

Mr. Monareng owns two businesses, a brickmaking concern and a retail hardware shop. The brickmaking concern, which is run from his home, brings in about 30 percent of his household's income. You are interviewing Mr. Monareng in the retail shop, which is in the commercial district. These are the only two sources of income available to the Monareng household. In the brickmaking business, he works with one full-time salaried man. In the retail shop, his wife works part-time keeping the books register but she is not paid a salary. There is also a male full-time salaried shop assistant in the hardware store who is paid Z\$5 per day. The retail shop operates year-round, five days a week, while the brickmaking operation is open only in the summer, and then only on weekends. Mr. Monareng started the hardware store last year, while the brickmaking activity has been going on for nine years. Both businesses have the same number of workers now as when the businesses started. In both enterprises, the major part of the business is selling to individual consumers. To begin his brickmaking concern, Mr. Monareng took out a loan from a moneylender, but he has never borrowed money to support the hardware store. The hardware store is supplied by a wholesaler of tools and other hardware, but the brickmaking concern collects its own inputs.

APPENDIX 5

COLUMN EXPLANATIONS FOR BASIC QUESTIONNAIRE

This appendix provides a description of the information collected on the illustrative basic questionnaire shown in Figure 2, and explains the content, purpose, and coding for each question. At the end of the appendix is a description of the spaces along the top of the questionnaire. Numbering of items in this appendix corresponds to the numbers written in the columns and spaces on the illustrative questionnaire. (These numbers are only for use in training, and are removed before the questionnaire is photocopied for field work.)

The bulk of the questions included on the illustrative basic questionnaire are standard columns that are always included on the questionnaire. For example, columns 1-4 and 31-32 are required for basic data handling, columns 6-10 and 23-26 provide basic descriptive information on the enterprise, while columns 11-18 and 21-22 are required for calculations of business size and growth. A few of the questions are not standard, but have been used on different baseline surveys to address specific questions. For instance, columns 28-29 pick up information about the proprietor's household, a focus which may be important to some clients.

* * * * *

Column 1

"PROPRIETOR SEQUENCE #"

Definition: PROPRIETOR SEQUENCE NUMBER

This column is filled in by the enumerator. It is designed to keep track of the number of proprietors interviewed by an enumerator in a given enumeration area. The first person the enumerator interviews is given a "1," the next "2," then "3," and so on, until the enumerator records the last interview for that enumeration area. If the respondent has more than one business in a given site and is therefore assigned more than one row on the questionnaire, that person is given the same sequence number for all rows containing his or her businesses. The only time the numbering sequence is broken is when the enumerator changes enumeration areas. When the enumerator begins in a new area, he begins again with "1," "2," and so on.

If a supplementary questionnaire is used, a space must appear on the supplementary questionnaire for proprietor sequence number. The same sequence number should appear on the supplementary questionnaire as was given that individual on the basic questionnaire.

If, in addition to the basic questionnaire, the closed enterprise questionnaire is administered, the numbering system becomes more complicated. Individuals are numbered sequentially regardless of which questionnaire they answer, or whether they answer both. If the same proprietor

is interviewed on more than one questionnaire, the same sequence number is given on all questionnaires delivered to that person. If the person appears on only one questionnaire, his sequence number appears on only that form and is skipped on the other questionnaire.

Given this system, it is not uncommon for the column on the basic questionnaire to be numbered "1,2,2,3,5..." while on the closed questionnaire the column reads "2,4,5..." In this example, the first person interviewed had an open business only, the second person interviewed had two open businesses as well as a closed business, the third person interviewed had an open business only, the fourth person interviewed had a closed business only, and the fifth person interviewed had both an open and a closed business.

The sequence numbers are not unique identification numbers: each enumerator has a "1," a "2," and so on for each enumeration area. Because they are not unique, they should not be entered into the computer. They serve only to order the enumerator's field interviews and tie questionnaires together so that the coder can accurately assign unique identification numbers once all the questionnaires are in the head office.

* * * * *

Column 2
"ENTERPRISE SEQUENCE #"
Definition: ENTERPRISE SEQUENCE NUMBER

This column is filled in by the enumerator. It is designed to keep track of the number of enterprises run by each proprietor in a given site. For the first business a proprietor mentions for that site, a "1" is written in the column. For the second business the same proprietor runs in the same site, a "2" is assigned, and so on. For most cases, the column will have a "1" in it, since most proprietors run only one enterprise out of a given site.

A similar system is used on the closed business questionnaire (if administered). The first closed business discussed for a given proprietor receives a "1," the second receives a "2," and so on.

Because the enterprise sequence numbers do not provide a unique identifier for a given enterprise, they are not entered into the computer. They are only useful for the coder who is responsible for tying together multiple businesses of a single proprietor.

* * * * *

Column 3
"PROPRIETOR ID #"

Definition: UNIQUE PROPRIETOR IDENTIFICATION NUMBER

This column is filled in by the coder, not by the enumerator, and is entered into the computer by the data enterer. It contains a unique identification number for each proprietor who was interviewed in the baseline survey.

The coder uses the information in Column 1 to fill in this column. The coder begins with "1" and numbers the individuals sequentially through to the end of the survey. For example, if 8,567 proprietors are interviewed over the course of the survey, numbers in this column will run from 1 through 8,567. The coder must keep track of the last number assigned on any given day, and begin the next day with the subsequent number. By the end of the survey, each proprietor will have a unique number connected with his or her interview in the computer.

The coder assigns a given proprietor the same, unique ID number each time he or she appears on any questionnaire. A person will appear more than once (with the same member) when he or she has more than one business at a given site (therefore, has more than one row on the basic questionnaire) or when the person is interviewed using more than one questionnaire.

* * * * *

Column 4
"ENTERPRISE ID #"

Definition: UNIQUE ENTERPRISE IDENTIFICATION NUMBER

This column is filled in by the coder, not by the enumerator, and is entered into the computer by the data enterer. It contains a unique identification number for each enterprise that is listed in the baseline survey.

The coder begins with "1" and numbers the enterprises sequentially through to the end of the survey. In effect, each case, or row of the questionnaire, has its own number which identifies it as a unique enterprise in the computer. If the 8,567 proprietors discussed above have 10,203 businesses among them, then the unique enterprise identification numbers will run from 1 through 10,203. The coder must keep track of the last number assigned on any given day, and begin the next day with the subsequent number.

If the closed business questionnaire is administered, it has a similar column that identifies each closed enterprise with a unique number. In the computer, and even though these data are stored in different files, this column must be given a different variable name than the corresponding column on the basic questionnaire because it is also numbered from "1" onward, and tracks a different set of business activities. For example, the unique identification number for current enterprises might

be called "ENT_ID", while the unique identification number for closed enterprises might be called "C_ENT_ID". (This differs from the proprietor identification number variable, which is given the same variable name in both data files so that individuals can be linked between data files.)

* * * * *

Column 5

"ENTERPRISE TYPE"

Definition: COMPLETE DESCRIPTION OF ACTIVITY UNDERTAKEN

This column records in words the exact nature of the enterprise. Enumerators are encouraged to be as specific as possible, writing down both the nature of the activity and the nature of the product. For example, if the enumerator writes "shoes" in the column, but does not write whether the person is making, selling, or simply polishing shoes, it is impossible to identify the activity correctly as manufacturing, trade, or services.

This column is not entered into the computer. It is used by the coder to check the accuracy of Column 6, which is entered into the computer.

Enumerators often try to fill in this column by inspection. Many enterprises, however, have hidden activities that may affect how the enterprise is defined. An example from Lesotho can illustrate: a man was found selling candies in a small shop. The enumerator, having observed the stock of candy, had recorded the enterprise as a candy shop. Only later in the interview did the enumerator discover that the proprietor also sold newspapers but had sold out the entire stock of papers for the day. What was more, the enumerator discovered that newspapers accounted for the vast majority of the shop's sales and, therefore, the business should have been categorized as a newspaper shop. The lesson for the enumerator is to always ask about the nature of the business, even if it seems apparent.

* * * * *

Column 6

"ENTERPRISE CODE"

Definition: NUMERIC CODE CORRESPONDING TO ENTERPRISE TYPE

This column is for the numeric code that corresponds to the type of activity listed in column 5. These codes are listed on the sheet entitled "Enterprise Codes."

This column is usually left blank during the interview, in order to minimize paper shuffling and delays while the enumerator looks for the correct code. It is filled in by the enumerator before the questionnaire is turned in to the supervisor.

* * * * *

Column 7**"% FAMILY INCOME (a)"****Definition: PERCENTAGE OF FAMILY INCOME FROM THIS BUSINESS**

This column records the importance of the enterprise with respect to total household income. Does the enterprise contribute all, more than half, less than half, or just half of household income? The codes (1-4) are listed at the bottom of the questionnaire next to the letter (a).

To get the correct answer from the respondent, enumerators need to draw the respondent's attention to all possible sources of income into the household, including:

- employment income from any household member
- remittances from household members living away from home
- other business income
- agricultural income
- gift income

and then ask them to compare the importance of the enterprise they are discussing to the total of these other sources.

It is important that the enumerators do not ask the respondents about amounts of income, only about the relative importance of the business to income. Amounts of money are never discussed in the baseline survey for three reasons: they make respondents uncomfortable and may even result in noncooperation, they are notoriously inaccurate, and they take too much time.

This question has traditionally been difficult for enumerators to ask correctly, and often requires extra attention during training.

* * * * *

Column 8**"LOCATION TYPE (b)"****Definition: TYPE OF LOCATION IN WHICH ENTERPRISE OPERATES**

This column records the location of the enterprise. Codes for this column are listed at the bottom of the questionnaire next to the letter (b). The code choices are:

- | | |
|-------------------------|--|
| (1) Home/homestead: | The business is located in the home or within the compound in which the proprietor lives. |
| (2) Traditional market: | The business is located in a marketplace, (often open-air or within a walled enclosure) where many people do business. |

- (3) **Commercial district:** The enterprise is located where formal businesses are concentrated. Although formal businesses may be the most obvious activities in commercial districts, many informal businesses locate there in hopes of capturing consumers who come to those areas to shop, bank, or work.
- (4) **Roadside:** The business is located along a road or footpath outside of a commercial area or traditional market.
- (5) **Mobile:** The business is not in a fixed location, rather it can be found at different places at different times.

This column can be filled in by inspection, either during or immediately following the interview.

* * * * *

Column 9
"STRUCTURE TYPE (c)"

Definition: TYPE OF STRUCTURE IN WHICH ENTERPRISE OPERATES

This column records the type of structure in which the enterprise operates. Codes for this question are at the bottom of the page next to the letter (c), and include the following:

- (1) **Permanent structure:** The enterprise operates from within a structure with walls and a roof, usually of cement, wood, brick, metal sheeting, or other permanent material.
- (2) **Temporary structure:** The enterprise operates from within a structure with sides and a cover, usually constructed of materials such as sticks, brush, cardboard, metal scraps, or polyurethane sheeting.
- (3) **Roof only:** The enterprise operates under a cover that gives some protection from rain and sun. The cover may be permanent (such as on the patio of a building) or temporary (such as under a lean-to roof).
- (4) **Open air:** The enterprise operates in the open air with no protection from rain or sun.

This question is not always included on the questionnaire. It may, however, be particularly useful in situations where the client is interested in enterprise working conditions or access to space.

* * * * *

Column 10

"# MONTHS PER YEAR"

Definition: NUMBER OF MONTHS PER YEAR THAT ENTERPRISE IS OPEN

This column records the number of months per year that the enterprise operates. It is designed to uncover seasonality of enterprises, which is particularly relevant when studying enterprises in the rural areas.

If the amount of time the enterprise operates, or the amount of time the proprietor spends in the enterprise, is of particular concern, an additional column can be added that captures the number of days per month that the enterprise operates or the hours per week that the proprietor spends in the business. These are considered optional questions.

* * * * *

NOTE: Columns 11-18 record information on the current work force of the enterprise. They count workers in four categories to come up with the total work force, then asks three additional questions about the total work force. Temporary or occasional workers are not counted. Any individual who works in the enterprise on a regular basis — whether part-time or full-time, paid or unpaid, adult or child — is counted.

* * * * *

Column 11

"# WORKING PROPRIETORS"

Definition: NUMBER OF PROPRIETORS WORKING IN THE ENTERPRISE

This column records the number of proprietors who work in the enterprise on a regular basis. This is not the same as the total number of proprietors, since some proprietors may not work in the business on a regular basis, and would therefore not be included in this column. The proprietor is defined as the person (or persons) who owns the enterprise. If the enterprise has a corporate structure, the proprietor is the person who makes major business decisions about the business (what will it produce or sell, how profits will be reinvested, and so forth).

Enumerators should be cautioned not to assume that the proprietor works in the business on a regular basis if he or she is there at the time of the interview.

* * * * *

Column 12

"# UNPAID FAMILY"

**Definition: NUMBER OF UNPAID FAMILY MEMBERS
WORKING IN THE ENTERPRISE**

This column records the number of family members who work in the business on a regular basis but are NOT fully compensated for their efforts. These family members may receive pocket money or no money for their efforts, but they do not draw a salary for the time they spend in the business.

* * * * *

Column 13

"# PAID WORKERS"

Definition: NUMBER OF PAID WORKERS IN THE ENTERPRISE

This column records the number of workers who are fully compensated for their efforts. It includes both family members who are paid and nonfamily members who are paid. It includes workers paid by the piece as well as those paid by the day or month, as long as the workers report to the business on a regular basis.

* * * * *

Column 14

"# APPRENTICES"

Definition: NUMBER OF APPRENTICES WORKING IN THE ENTERPRISE

This column records the number of apprentices working in the business on a regular basis. Apprentices differ from hired workers who receive on-the-job training in that apprentices do not receive monetary compensation for their efforts.

* * * * *

Column 15

"TOTAL # OF WORKERS"

Definition: TOTAL NUMBER OF WORKERS IN THE ENTERPRISE

This column records the total number of people working in the business on a regular basis. It includes a complete count of working proprietors, unpaid family workers, paid workers, and apprentices, and is therefore the sum of the numbers recorded in columns 11-14.

* * * * *

Column 16
"# FEMALE"

Definition: NUMBER OF FEMALE WORKERS OUT OF TOTAL

Of the total number of workers in the enterprise (recorded in column 15), how many are female?

Column 17
"# PART-TIME"

Definition: NUMBER OF PART-TIME WORKERS OUT OF TOTAL

Of the total number of workers in the enterprise (recorded in column 15), how many work on a part-time basis? A part-time worker is defined as anyone who works in the business less than 30 hours per week.

Column 18
"# CHILDREN"

Definition: NUMBER OF CHILDREN OUT OF TOTAL

Of the total number of workers in the enterprise (recorded in column 15), how many are under 15 years of age?

Columns 19-20
"PAY OF BEST PAID WORKER" and "PAY PERIOD (d)"
Definition: HOW MUCH IS THE BEST-PAID WORKER RECEIVING?

Columns 19 and 20 together record the amount that the best paid worker is receiving for his or her time. Two columns are required to capture the two pieces of information needed. The first, column 19, records the amount of money, while the second, column 20, records the pay period or increment. Codes for column 20 are listed at the bottom of the page next to the letter (d), and include:

- (1) Per hour
- (2) Per day
- (3) Per week
- (4) Per month

- (5) Per year
 (6) Per piece
 (7) Other

As an example, if the best-paid worker in an enterprise was paid 30 shillings per day, "30" would be recorded in column 19, and "2," the code for "per day," would be recorded in column 20.

Column 21

"YEAR BUSINESS WAS ACQUIRED"

Definition: THE YEAR THE PROPRIETOR ACQUIRED THE ENTERPRISE

This column records the year in which the proprietor began or took control of the business. To avoid crowding on the questionnaire, only the last two digits of the year are recorded in this column. For example, if the answer is 1985, the enumerator writes "85" in the space.

Column 22

"# WORKERS AT START"

Definition: TOTAL NUMBER OF WORKERS IN THE ENTERPRISE WHEN IT FIRST OPENED

This column records the total number of workers in the firm when it was first acquired by the proprietor. The proprietor should explicitly be reminded to count himself, as well as any family members, paid workers, and apprentices who worked in the business.

Column 23

"SELLS TO WHOM? (e)"

Definition: WHAT IS THE MOST IMPORTANT MARKET FOR THE ENTERPRISE?

This column was introduced to explore patterns of forward linkages by identifying the largest buyer of the small enterprise's goods or services. The codes, listed at the bottom of the page next to the letter (e), give the following options:

- | | | |
|-----|-------------------------------|--|
| (1) | Individuals: | The business sells its goods or services primarily to individuals. |
| (2) | Urban commercial enterprises: | The business produces goods for resale by urban businesses. |

- | | | |
|-----|----------------------------------|---|
| (3) | Urban manufacturing enterprises: | The business produces inputs (goods or services) for further productive activities in the urban area. |
| (4) | Rural commercial enterprises: | The business produces goods for resale by rural businesses. |
| (5) | Rural manufacturing enterprises: | The business produces inputs (goods or services) for further productive activities in the rural area. |
| (6) | Export: | The business exports the majority of its products. |
| (7) | Other: | All answers that cannot be categorized in (1)-(6) fall under this code. |

During training, enumerators should be warned that although some respondents will give multiple answers to this question, the survey is looking for the business's most important type of buyer, not a list of all types of buyers.

* * * * *

Column 24

"NATURE OF INPUTS (f)"

Definition: AT WHAT LEVEL OF PROCESSING DOES THE PROPRIETOR BUY THE ENTERPRISE'S MOST IMPORTANT INPUT?

This column complements Column 23 in that it searches for patterns of backward linkages to the enterprise by examining the source of the enterprise's inputs or stock. Because enterprises use multiple inputs, the question focuses only on the most important input into the business (as defined by the proprietor).

Codes for this question are listed at the bottom of the page next to the letter (f), and include the following five categories:

- | | | |
|-----|----------------------------|--|
| (1) | Makes/gathers own inputs: | The enterprise makes or gathers its most important input. Beer-brewers may, for example, make their own hops. |
| (2) | Buys unprocessed inputs: | The enterprise buys unprocessed inputs, then does the processing as a stage of production within the enterprise. |
| (3) | Buys semiprocessed inputs: | The enterprise buys inputs that have already been processed to some extent. |

- | | | |
|-----|----------------------------------|--|
| (4) | Buys finished products for sale: | The enterprise does not buy inputs for production, but rather buys finished goods for resale. This category is usually applicable for retail concerns. |
| (5) | Other: | All answers that do not fall under (1)-(4) above are coded as (5). |

In addition to providing information on forward and backward linkages, columns 23 and 24 provide the material to develop an index of the enterprise's level of vertical integration. In such an index, for example, the most integrated business would process its own inputs and market directly to final consumers. A less integrated business would buy semi-processed inputs and/or sell to other businesses that service the final consumer.

* * * * *

Column 25

“CREDIT SOURCE (g)”

Definition: WHAT HAS BEEN THE ENTERPRISE'S MAJOR SOURCE OF CREDIT?

This column asks for the enterprise's most important source of credit, including during the start-up period. Codes are listed at the bottom of the page next to the letter (g), and include the following:

- (1) Never received loans
- (2) Loans from family/friends
- (3) Loans from moneylenders
- (4) Loans from banks
- (5) Loans from elsewhere

* * * * *

Column 26

“PROPRIETOR GENDER (h)”

Definition: WHAT IS THE PROPRIETOR'S GENDER?

This column records the gender of the proprietor. Codes for this question are listed at the bottom of the questionnaire next to the letter (h). The code choices are:

- | | | |
|-----|-------------------------|--|
| (1) | Female, one proprietor: | The business has only one proprietor, who is female. |
| (2) | Male, one proprietor: | The business has only one proprietor, who is male. |

- (3) Female, more than one proprietor: The enterprise has more than one proprietor, and all of them are female.
- (4) Male, more than one proprietor: The enterprise has more than one proprietor, and all of them are male.
- (5) Mixed-gender joint proprietorship: The enterprise has more than one proprietor, and includes at least one male and one female proprietor.

* * * * *

Column 27

“PROPRIETOR CATEGORY (i)”

Definition: WHAT IS THE PROPRIETOR’S ETHNIC AFFILIATION?

This column, while not always included in the questionnaire, may be useful for identifying proprietors by ethnic or national group. In Zimbabwe, the color of the entrepreneur was considered the relevant distinction; in Lesotho, tribal identity was used. Since categories are country-specific, codes need to be added at the bottom of the questionnaire next to the letter (i) if the question is used.

* * * * *

Column 28

“TOTAL # OF HH ENTERPRISES”

Definition: TOTAL NUMBER OF ENTERPRISES IN THE HOUSEHOLD

This column records the total number of enterprises undertaken by the household as a unit. Although the household is not the unit of analysis used in the survey, it is still of concern to policy makers, who may want to know how many enterprises the average household runs. Household has traditionally been defined as including the individuals who eat out of the same pot.

* * * * *

Column 29

“# IN HOUSEHOLD IN PAID EMPL.”

Definition: NUMBER OF INDIVIDUALS IN HOUSEHOLD IN PAID EMPLOYMENT

This column records the number of individuals in the household who have paid jobs outside of the family’s enterprises. When combined with the information in Column 28, it provides a profile on how the household supports itself.

* * * * *

Column 30

**“PROPRIETOR’S NAME, NICKNAME, ENTERPRISE NAME,
PHYSICAL ADDRESS, AND CONTACT’S NAME”**

This space is used to record five types of information to help the researcher contact the respondent again. Because the space is so small, enumerators will have to write neatly to fit all the information in the space provided.

Proprietor’s name: Enumerators record the proprietor’s first and last name, asking for assistance in spelling if necessary.

Nickname: If the proprietor has a nickname that is used by friends, neighbors, or customers, it is written down in quotation marks after his/her formal name.

Business name: Does the business have either a formal or informal name? A formal name will probably be displayed on the building, while an informal name may be how the business is referred to in conversation, such as “Mama’s Cak^{es}.”

Physical address: Where is the business situated with respect to streets, intersections, and landmarks? The information should be recorded with as much detail as possible, and include any helpful details. An example is: “Mud house with green door on Tsamaya Street, two blocks down from Mveledzo Primary School.”

It is important to remind enumerators that mailing addresses are not acceptable. The enumerators should write down whatever information they would need to return to that site a month later.

Contact’s name: What is the name of someone who will be able to help find the proprietor if he or she cannot be located? This person may be a neighbor, another business person, or a friend. Although this heading is not explicitly written on the questionnaire, the information can be very important in some situations. In particular, contacts’ names should be requested for mobile vendors or those with businesses in squatter settlements.

* * * * *

Column 31

“SUPPLEMENTARY QUEST.? (1=YES, 0=NO)”

Definition: DID THE ENUMERATOR ADMINISTER THE SUPPLEMENTARY QUESTIONNAIRE FOR THIS ENTERPRISE?”

This column is only used if a supplementary questionnaire is included in the baseline survey. The column is needed to later help the coder and researcher link the basic questionnaire interviews to the corresponding supplementary questionnaire interviews.

Codes for this column are given in the heading, where “Yes” is “1” and “No” is “0”. If the enumerator administers the supplementary questionnaire for a given interview, he marks a “1” in the column for that row, signifying that there is a supplementary questionnaire pertaining to this business. If he does not administer the supplementary questionnaire in a given interview, he marks a “0” in the column for that row, signifying that there is no supplementary questionnaire on that business.

Column 32

“ANY PAST ENTERPRISES? (1=YES, 0=NO)”

Definition: DID THE ENUMERATOR ADMINISTER THE CLOSED ENTERPRISE QUESTIONNAIRE FOR THIS PROPRIETOR?”

This column is only used if the closed enterprise questionnaire is included in the survey. It is needed to help the coder and researcher link the basic questionnaire interviews to the corresponding closed enterprise questionnaire interviews.

Codes for this column are given in the heading, where “Yes” is “1” and “No” is “0.” If the enumerator administers the closed enterprise questionnaire for the proprietor of a given interview, he marks a “1” in the column for that proprietor’s rows, signifying that there is at least one closed enterprise questionnaire pertaining to that proprietor. If he does not administer the closed enterprise questionnaire, he marks a “0” in the column for that row, signifying that there is no closed enterprise questionnaire for that proprietor.

INFORMATION AT THE TOP OF THE QUESTIONNAIRE

REGION, STRATUM, ENUMERATION AREA

Enumerators are given a list of enumeration areas, categorized by stratum and by region. The list assigns a code to each enumeration area, stratum, and region. At the top left-hand corner of the questionnaire, the enumerator fills in the code for each of these categories to match where the interviews are being carried out. He also writes out the name of the enumeration area, stratum, and region, to ensure that the questionnaires are properly classified.

DATE COMPLETED

The enumerator fills in the date he completed the interviews and turned the paper in to his supervisor at date completed.

DATE PROOFED, DATE ENTERED, DATE VERIFIED

The enumerator leaves these spaces blank. The coder fills in the date he completes coding on the sheet at date proofed, the data enterer fills in the date she completes entering the data from the questionnaire at date entered, and also fills in the date she completes verifying the data from the questionnaire at date verified.

PAGE #

The enumerator keeps track of the pages he has filled in for that enumeration area. The first complete page is "1," the next "2," and so on. Numbering continues from day to day until the enumeration area is complete. As soon as the team starts a new enumeration area, the enumerator begins with a fresh page (with the new enumeration area's name and number recorded at the top) that is again numbered "1."

ENUMERATOR/SUPERVISOR

In these spaces, the enumerator enters his name and the name of his supervisor.

SITES WITH NO ACTIVITY

This is the space where a running tally is kept of all sites visited where no activity is taking place. For each such site, the enumerator places a stroke in this space. After the enumeration area has been covered, the coder or researcher can add up the number of strokes made by each enumerator in that enumeration area to arrive at the number of sites with no activity in that area.

SITES CLOSED

This is the space where a running tally is kept of all sites that are either closed or where people refused to grant an interview. For each such site, the enumerator places a stroke in this space. After the enumeration area has been covered, the coder or researcher can add up the number of strokes made by each enumerator in that enumeration area to arrive at the number of sites closed in that area, which is then used to calculate the percentage of sites that were included in the survey.

APPENDIX 6

COLUMN EXPLANATIONS FOR CLOSED ENTERPRISE QUESTIONNAIRE

Column 2

"ENTERPRISE SEQUENCE #"

Definition: ENTERPRISE SEQUENCE NUMBER

This column is filled in by the enumerator. It is designed to keep track of the number of now-closed enterprises a given proprietor has had. For the first closed business a proprietor mentions, a "1" is written in the column. For the second closed business the proprietor mentions, a "2" is assigned, and so on.

Because the enterprise sequence numbers do not provide a unique identifier for a given enterprise, they are not entered into the computer. They are only useful for the coder who is responsible for tying together the business history of a single proprietor.

Column 3

"PROPRIETOR ID #"

Definition: UNIQUE PROPRIETOR IDENTIFICATION NUMBER

This column is filled in by the coder, not by the enumerator, and is entered into the computer by the data enterer. It contains a unique identification number for each proprietor who was interviewed in the baseline survey.

The coder uses the information in Column 1 to fill in this column. The coder begins with "1" and numbers the individuals sequentially through to the end of the survey. For example, if 8,567 proprietors are interviewed over the course of the survey, numbers in this column will run from 1 through 8,567. The coder must keep track of the last number assigned on any given day, and begin the next day with the subsequent number. By the end of the survey, each proprietor will have a unique number connected with his or her interviews in the computer.

The coder assigns a given proprietor the same unique ID number each time he or she appears on any questionnaire. A person will appear more than once when he or she has more than one closed business (therefore has more than one row on the closed enterprise questionnaire) or when the person is interviewed using more than one questionnaire.

Previous Page Blank

Column 4

“CLOSED ENTERPRISE ID #”

Definition: UNIQUE CLOSED ENTERPRISE IDENTIFICATION NUMBER

This column is filled in by the coder, not by the enumerator, and is entered into the computer by the data enterer. It contains a unique identification number for each closed enterprise that is listed in the baseline survey.

The coder begins with “1” and numbers the closed enterprises sequentially through to the end of the survey. In effect, each case, or row of the questionnaire, has its own number that identifies it as a unique closed enterprise in the computer. If the 8,567 proprietors discussed above have 2,000 closed businesses among them, then the unique closed enterprise identification numbers will run from 1 through 2,000. The coder must keep track of the last number assigned on any given day, and begin the next day with the subsequent number.

This column must be given a different variable name than the corresponding column on the basic questionnaire because it is also numbered from “1” onward, and tracks a different set of business activities. For example, the unique identification number for current enterprises might be called “ENT_ID”, while the unique identification number for closed enterprises might be called “C_ENT_ID”. (This differs from the proprietor identification number variable, which is given the same variable name in both data files so that individuals can be linked between data files.)

* * * * *

Column 5

“CLOSED ENTERPRISE TYPE”

Definition: COMPLETE DESCRIPTION OF ACTIVITY UNDERTAKEN

This column records in words the exact nature of the closed enterprise. Enumerators are encouraged to be as specific as possible when writing down the type of activity. This column is not entered into the computer. It is used by the coder to check the accuracy of Column 6, which is entered into the computer.

* * * * *

Column 6

“CLOSED ENTERPRISE CODE”

Definition: NUMERIC CODE CORRESPONDING TO ENTERPRISE TYPE

This column is for the numeric code that corresponds to the type of activity listed in column 6. The enumerator uses the same list of codes given for the basic questionnaire. This column is usually left blank during the interview, in order to minimize paper shuffling and delays while the

enumerator looks for the correct code. It is filled in by the enumerator before the questionnaire is turned in to the supervisor.

* * * * *

Column 7
“LOCATION TYPE (a)”

Definition: TYPE OF LOCATION IN WHICH ENTERPRISE OPERATES

This column records the location of the closed enterprise. Codes for this column are listed at the bottom of the questionnaire next to the letter (a). The code choices are:

- (1) Home/homestead: The business was located in the home or within the compound in which the proprietor lived at the time.
- (2) Traditional market: The business was located in a marketplace, where many people did business.
- (3) Commercial district: The enterprise was located where formal businesses were concentrated.
- (4) Roadside: The business was located along a road or footpath outside of a commercial area or traditional market.
- (5) Mobile: The business was not in a fixed location, rather it moved to different places at different times.

* * * * *

Column 8
“YEAR ENTERPRISE STARTED”

Definition: THE YEAR THE PROPRIETOR ACQUIRED THE ENTERPRISE

This column records the year in which the proprietor began or took control of the business. To avoid crowding on the questionnaire, only the last two digits of the year are recorded in this column. For example, if the answer is 1985, the enumerator writes “85” in the space.

* * * * *

Column 9

"# WORKERS AT START"

**Definition: TOTAL NUMBER OF WORKERS IN THE ENTERPRISE
WHEN IT FIRST OPENED**

This column records the total number of workers in the enterprise when it was first acquired by the proprietor. The proprietor should explicitly be reminded to count himself, as well as any family members, paid workers, and apprentices who worked in the business.

* * * * *

Column 10

"YEAR ENTERPRISE CLOSED"

Definition: THE YEAR THE PROPRIETOR CLOSED THE ENTERPRISE

This column records the year in which the proprietor closed the business. To avoid crowding on the questionnaire, only the last two digits of the year are recorded in this column. For example, if the business closed in 1990, the enumerator writes "90" in the space.

* * * * *

Column 11

"# WORKERS AT CLOSE"

**Definition: TOTAL NUMBER OF WORKERS IN THE ENTERPRISE
WHEN IT CLOSED**

This column records the total number of workers in the enterprise just before it closed. The proprietor should explicitly be reminded to count himself, as well as any family members, paid workers, and apprentices who worked in the business.

* * * * *

Column 12

"HIGHEST NUMBER OF WORKERS"

**Definition: THE LARGEST SIZE THE ENTERPRISE ATTAINED IN TERMS OF
NUMBER OF WORKERS**

This column records the largest size the enterprise attained in terms of total number of workers. When combined with the information from columns 9 and 11, the data provide a sketch of how much businesses grew and when. The proprietor should explicitly be reminded to count himself, as well as any family members, paid workers, and apprentices who worked in the business at any given time.

* * * * *

Column 13

“SOLD PRODUCT TO WHOM? (b)”

Definition: WHAT WAS THE MOST IMPORTANT MARKET FOR THE ENTERPRISE’S OUTPUT?

This column was introduced to explore patterns of forward linkages by identifying the largest buyer of the small enterprise’s goods or services. The codes, listed at the bottom of the page next to the letter (e), give the following options:

- (1) **Individuals:** The business sold its goods or services primarily to individuals.
- (2) **Urban commercial enterprises:** The business produced goods for resale by urban businesses.
- (3) **Urban manufacturing enterprises:** The business produced inputs (goods or services) for further productive activities in the urban area.
- (4) **Rural commercial enterprises:** The business produced goods for resale by rural businesses.
- (5) **Rural manufacturing enterprises:** The business produced inputs (goods or services) for further productive activities in the rural area.
- (6) **Export:** The business exported the majority of its products.
- (7) **Other:** All answers that cannot be categorized in (1)-(6) fall under this code.

During training, enumerators should be warned that although some respondents will give multiple answers to this question, the survey is looking for the business’s most important buyer, not a list of all types of buyers.

Column 14

“NATURE OF INPUTS (c)”

Definition: AT WHAT LEVEL OF PROCESSING DID THE PROPRIETOR BUY THE ENTERPRISE’S MOST IMPORTANT INPUT?

This column complements Column 13 in that it searches for patterns of backward linkages to the enterprise by examining the source of the enterprise’s inputs or stock. Although most enterprises use multiple inputs, the question focuses only on the most important input into the business (as defined by the proprietor).

Codes for this question are listed at the bottom of the page next to the letter (f), and include the following five categories:

- (1) **Makes/gathers own inputs:** The enterprise made or gathered its most important input.
- (2) **Buys unprocessed inputs:** The enterprise bought unprocessed inputs then did the processing as a stage of production within the enterprise.
- (3) **Buys semi-processed inputs:** The enterprise bought inputs that had already been processed to some extent.
- (4) **Buys finished products for sale:** The enterprise did not buy inputs for production, but rather bought finished goods for resale. This category is usually applicable for retail concerns.
- (5) **Other:** All answers that do not fall under (1)-(4) above are coded as (5).

In addition to providing information on forward and backward linkages, columns 13 and 14 provide the material to develop an index of the enterprise's level of vertical integration.

* * * * *

Column 15

"REASON ENTERPRISE CLOSED"

Definition: FOR WHAT REASON WAS THE ENTERPRISE CLOSED?

In this column the enumerator writes out in words the reason that the enterprise was closed as explained by the respondent.

* * * * *

Column 16

"CLOSURE CODE (d)"

Definition: NUMERIC CODE CORRESPONDING TO REASON FOR CLOSURE

This column contains the numeric code for the reason for closure given in column 15. Respondents may give a wide variety of responses, but the enumerators must categorize each answer into one of the following responses, which are recorded at the bottom of the page next to the letter (d):

- | | | |
|-----|---|---|
| (1) | Market problems: | This category includes all demand-related problems (for example, not enough customers), as well as problems with competition (for example, too many producers, or other producers undercut prices). |
| (2) | Personal reasons: | This category includes all nonbusiness reasons for closure, such as proprietor death, sickness, or retirement, family responsibilities (such as child care or marriage), a move by the household, or any other such reason. |
| (3) | Shortage or expense of operating funds: | The proprietor could not gain access to or could not afford working capital. |
| (4) | Shortage or expense of | The proprietor could not gain |
| | | stock or raw materials:access to or could not afford stock or raw materials. |
| (5) | Got a job: | The proprietor closed the business to work for someone else. |
| (6) | Started another MSE: | The proprietor closed the business in order to open another business activity. |
| (7) | Legal troubles/
government interference: | The business was closed due to legal troubles or government interference. |
| (8) | Other | All answers that cannot be categorized as (1)-(7). |

Column 17

“WHAT ARE YOU DOING NOW?”

**Definition: WHAT INCOME-EARNING ACTIVITIES
IS PROPRIETOR DOING NOW?**

In this column the enumerator writes out in words what the respondent is now doing to earn his or her income.

Column 20

“PROPRIETOR CATEGORY (g)”

Definition: WHAT IS THE PROPRIETOR’S ETHNIC AFFILIATION?

This column, although not always included in the questionnaire, may be useful for identifying proprietors by ethnic or national group. Because categories are country-specific, codes need to be added at the bottom of the questionnaire next to the letter (g) if the question is used.

Column 21

“PROPRIETOR’S NAME, NICKNAME, ADDRESS, AND CONTACT”

This space is used to record three types of information to help the researcher contact the respondent again in the future.

Proprietor’s name: Enumerators record the proprietor’s first and last name, asking for assistance in spelling if necessary.

Nickname: If the proprietor has a nickname that is used by friends, neighbors, or customers, it is written down in quotation marks after his/her formal name.

Address: Where does the proprietor live or work? The information should be recorded with as much detail as possible. It is important to remind enumerators that mailing addresses are not acceptable. The enumerators should write down whatever information they would need to contact that individual again a month later.

Contact’s name: What is the name of someone who will be able to help find the respondent if he or she cannot be located? This person may be a neighbor, another business person, or a friend.

Column 22

“ENUMERATOR: DID YOU FILL IN BQ? (1= YES, 0=NO)”

Definition: DOES INDIVIDUAL HAVE CURRENT ENTERPRISES RECORDED ON THE BASIC QUESTIONNAIRE?

The enumerator puts a “1” in this column if the respondent has a current business as well as a closed business, and therefore appears on the basic questionnaire as well. If the respondent has no current businesses, the enumerator puts a “0” in the column, signifying that he did not fill in the basic questionnaire for this individual.

TOP OF THE QUESTIONNAIRE:

For instructions on filling in the top of the questionnaire, see the end of Appendix 5, which describes the top of the basic questionnaire form.

APPENDIX 7

ILLUSTRATIVE BUSINESS CODE SHEET

A. PRODUCTION/MANUFACTURING

Garments/
Textiles

1. Tailoring
2. Dressmaking
3. Knitting
4. Weaving
- 5.
6. Other Textiles

Forest-
Based
Products

7. Sawmilling
8. Furniture making
9. Carpentry
10. Wood carving
11. Grass/Cane works
12. Coal/Wood production
- 13.
14. Other Woodworking

Metal
Work

15. Welding
16. Tinmithing
17. Blacksmithing
- 18.
19. Other metalworking

Repair
Work

20. Auto repair
21. Bike repair
22. Electrical repair
23. Clock/watch/jewelry repair
24. Radio/TV repair
- 25.
26. Other repairs

Leather/
Rubber/
Plastics

27. Shoework and repairs
28. Other leatherwork
29. Rubber work
30. Plastic work

Food
Process-
ing

31. Flour/maize milling
32. Butchery
33. Baking
- 34.
35. Other food processing

Beverages

36. Beer brewing/shebeen
- 37.
38. Other beverage making

Masonry

39. Brick/block making
40. Pottery
- 41.
42. Other masonry

Other

43. Chemical production

Manufac-
turing

44. Photo studio
45. Glass work
46. Printing work
47. Jewelry work
48. Art/artifact production
- 49.
50. All other manufacturing

B. COMMERCE/TRADE

Vending

55. Vending foods
56. Vending drinks
57. Vending farm products
58. Vending forest-based products
59. Vending hardware
60. Vending garments
61. Vending art/artifacts
62. Food catering
- 63.
- 64.
65. Vending other

Retail

66. Grocery
67. Retail food
68. Bottle store
69. Retail livestock
70. Retail farm products
71. Retail hardware
72. Retail forest-based products
73. Retail garments
74. Retail leather/shoes
75. Stationers/bookstore
76. Filling station
77. General trader/dealer
78. Pharmacy
79. Restaurant
80. Bar/Pub
- 81.
- 82.
83. Other Retailing

Wholesale

84. Liquor distributor
85. Wholesale

C. SERVICES

86. Hotel
87. Laundry
88. Dry cleaning
89. Hair salon/barber
90. Construction
91. Traditional healer
92. Funeral services
93. Bus/taxi services
94. Goods transport
95. Renting flats/rooms
- 96.
- 97.
- 98.
99. Other services

APPENDIX 8

INTERNATIONAL STANDARD INDUSTRIAL CLASSIFICATION CODES

- 3 MANUFACTURING**
- 31 Food, Beverage, and Tobacco Processing**
 - 3111 Butchery
 - 3116 Flour/Maize Milling
 - 3117 Baking
- 32 Textile and Wearing Apparel Production**
 - 3221 Dressmaking
 - 3222 Tailoring
 - 3223 Knitting
 - 3224 Other Textiles
 - 3225 Weaving
 - 3226 Crocheting
- 33 Wood Production and Processing**
 - 3311 Sawmilling
 - 3312 Grass/Cane/Bamboo Work
 - 3313 Coal/Wood Processing
 - 3319 Wood Carving
 - 3320 Carpentry
 - 3321 Furniture Making
 - 3322 Other Woodworking
- 34 Paper, Printing, and Publishing**
 - 3420 Printing Work
- 35 Chemical and Plastics Production**
 - 3520 Chemical Production
 - 3525 Plastic Production

Previous Page Blank

36 Non-Metallic Mineral Processing

- 3610 Pottery
- 3620 Glass Work
- 3690 Brick/Block Making
- 3699 Other Masonry

38 Metal Fabrication

- 3811 Blacksmithing
- 3814 Tinsmithing
- 3818 Other Metalworking
- 3819 Welding

39 Miscellaneous Manufacturing

- 3901 Jewelry Production
- 3904 Art/Artifact Production
- 3909 All Other Manufacturing

5 CONSTRUCTION

50 Construction

6 TRADE

61 Wholesale Trade

- 6100 Liquor Distributing
- 6110 Wholesaling

62 Retail Trade

- 6201 Vending foods
- 6202 Vending Drinks
- 6203 Vending Farm Products
- 6204 Vending Garments
- 6205 Vending Forest-Based Products
- 6206 Vending Hardware
- 6207 Vending Art/Artifacts
- 6208 Other Vending
- 6209 Food Catering
- 6213 Grocery
- 6214 Retail Food

- 6215 Bottle Store
- 6216 Retail Livestock
- 6217 Retail Farm Products
- 6220 Retail Garments
- 6221 Retail Leather/Shoes
- 6230 Retail Forest-Based Products
- 6240 Stationers/Bookstore
- 6250 Filling Station
- 6251 Pharmacy
- 6280 Retail Hardware
- 6290 General Trader/Dealer
- 6291 Other Retailing

63 Hotels, Restaurants, and Bars

- 6309 Hotel
- 6310 Restaurant
- 6311 Bar/Pub
- 6312 Shebeen

7 TRANSPORTATION

71 Transportation

- 7113 Bus/Taxi
- 7114 Goods Transport

8 FINANCE, REAL ESTATE, AND BUSINESS SERVICES

83 Real Estate

- 8310 Renting Flats/Rooms

9 SERVICES

95 Services

- 9520 Laundry
- 9521 Dry Cleaning
- 9531 Traditional Healing
- 9591 Hair Salon/Barber
- 9592 Photo Studio
- 9597 Funeral Services
- 9599 Other Services

APPENDIX 9

INFORMATION ABOUT HOUSEHOLDS

The survey described in this manual is based on the enterprise, and not the household, as the unit of analysis. As discussed in Chapter Four, the necessity of interviewing on-site (to avoid double-counting) prevents this particular survey from providing comprehensive data at the household level. Still, some insights into the workings of the household can be gleaned from the MSE survey. Two of the columns in the Illustrative Basic Questionnaire, for example, gather information about the household. These are discussed in Appendix 5.

Information generated by the survey can also answer another question. Sometimes clients or other interested parties want to know what proportion of households in the country or region under study are engaged in MSE activity. The tallies in the upper right corner of the questionnaire identify the numbers of sites in each cluster that either have no MSE activity or are closed. The number of interviews tells how many sites were found with activity. Unfortunately, because sites are composed of not only households but also shops and mobile enterprises, the proportion of households with activity is not directly available. With a small amount of preparation, however, this question can be addressed.

First, note that sites with no activity are almost exclusively households, since by definition shops and mobile enterprises have activity. Second, the number of households within the cluster engaged in MSE activity must be ascertained. This is not difficult, for the survey gathers information on enterprise location. In other words, the number of home-based enterprises are tallied, using the location identifier in the basic questionnaire. This number is then added to the number of sites (households) without activity, to come up with a number of open households visited. The number of households with activity is then divided by this number to calculate the proportion of households engaged in MSE activity. If the researcher is interested in the countrywide proportion, she can follow the weighting procedures described in Chapter Five.

Previous Page Blank

APPENDIX 10

FORMULAE FOR SAMPLE SIZE

How many MSEs ought to be enumerated in a stratified sampling procedure? Let us first consider a stratified technique without cluster sampling (i.e., enterprises are randomly selected within each stratum). In a design with h strata, where the total number of enterprises in the country, N , is given by

(1)

$$\sum_h N_h = N,$$

the sample size necessary for a given level of precision is:

(2)

$$n = \frac{n_0}{1 + \frac{n_0}{N}}$$

where

$$n_0 = \frac{1}{V} \sum_h W_h s_h^2$$

$$s_h^2 = \frac{1}{n_h - 1} \sum_{i=1}^{n_h} (y_{hi} - \bar{y}_h)^2$$

$$W_h = \frac{N_h}{N}$$

and

$$y_{hi}, \bar{y}_h$$

Previous Page Blank

are the values of some continuous variable of interest and the mean of this variable, respectively.

The variance, V , can be calculated in two different ways. Either it can be estimated from the data, or if a particular desired margin of error is specified, it can be approximated as:

(3)

$$V = \left(\frac{d}{t} \right)^2$$

where d is the desired margin of error, and t is the confidence level.

As noted in Chapter Three, the use of cluster sampling decreases the precision relative to randomly sampling units within the stratum. The degree factor of adjustment, as reported in Casley and Lury (1987) is given by:

$$(4) \quad 1 + \delta(M - 1),$$

where δ is the intraclass correlation and M is the size of the cluster. More specifically,

(5)

$$\delta = \frac{S_b^2 - S^2}{(M-1) S^2}$$

where S_b^2 is the between-cluster variance, S^2 is the variance among all MSEs, and M is the average number of MSEs per cluster. δ is calculated for each stratum, and the resulting factor of adjustment for each stratum is multiplied by the sample size estimates calculated for each stratum using equation (2). As the average cluster size, M , increases, this imprecision increases. Cochran (1977) notes that the intraclass correlation, δ , is also a function of cluster size (specifically, S_b^2 and S^2 are decreasing functions of M), and that increases in the average cluster size decrease δ . However, this effect is not enough to offset the increase in imprecision described in equation (4). The especially diligent reader is referred to a sampling text, such as Cochran (1977) for further detail.

Given that the variances in the equations above are unknown, these formulae are not useful for calculating the appropriate sample sizes in the survey design phase. Still, these equations have uses. If estimate of variances from previous surveys are calculated, these formulae can provide approximations for sample sizes in future surveys. Also, it is useful, though sometimes painful, to ascertain after the survey how appropriate the sample size for a given survey was so that the inaccuracies it may contain are known.

GEMINI PUBLICATION SERIES

GEMINI Working Papers:

1. "Growth and Equity through Microenterprise Investments and Institutions Project (GEMINI): Overview of the Project and Implementation Plan, October 1, 1989-September 30, 1990." GEMINI Working Paper No. 1. December 1989. [not for general circulation]
- *2. "The Dynamics of Small-Scale Industry in Africa and the Role of Policy." Carl Liedholm. GEMINI Working Paper No. 2. January 1990. \$5.50.
3. "Prospects for Enhancing the Performance of Micro- and Small-Scale Nonfarm Enterprises in Niger." Donald C. Mead, Thomas Dichter, Yacob Fisseha, and Steven Haggblade. GEMINI Working Paper No. 3. February 1990. \$6.00
4. "Agenda Paper: Seminar on the Private Sector in the Sahel, Abidjan, July 1990." William Grant. GEMINI Working Paper No. 4. August 1990. \$3.00
- *5. "Gender and the Growth and Dynamics of Microenterprises." Jeanne Downing. GEMINI Working Paper No. 5. October 1990. \$10.50.
6. "Banking on the Rural Poor in Malaysia: Project Ikhtiar." David Lucock. GEMINI Working Paper No. 6. October 1990. \$3.30.
7. "Options for Updating AskARIES." Larry Reed. GEMINI Working Paper No. 7. October 1990. \$3.50.
- *8. "Technology — The Key to Increasing the Productivity of Microenterprises." Andy Jeans, Eric Hyman, and Mike O'Donnell. GEMINI Working Paper No. 8. November 1990. \$3.60.
9. "Lesotho Small and Microenterprise Strategy — Phase II: Subsector Analysis." Bill Grant. GEMINI Working Paper No. 9. November 1990. \$15.50.
- *10. "A Subsector Approach to Small Enterprise Promotion and Research." James J. Boomgard, Stephen P. Davies, Steven J. Haggblade, and Donald C. Mead. GEMINI Working Paper No. 10. January 1991. \$3.10.
11. "Data Collection Strategies for Small-Scale Industry Surveys." Carl Liedholm. GEMINI Working Paper No. 11. January 1991. \$1.30.
12. "Dynamics of Microenterprises: Research Issues and Approaches." Carl Liedholm and Donald C. Mead. GEMINI Working Paper No. 12. January 1991. \$6.50.
13. "Dynamics of Microenterprises: Research Priorities and Research Plan." Carl Liedholm and Donald C. Mead. GEMINI Working Paper No. 13. August 1990. [not for general circulation]

*Publications of general interest

127

14. "Review of Year One Activities (October 1, 1989 to September 30, 1990) and Year Two Work Plan (October 1 to November 30, 1990)." GEMINI Working Paper No. 14. January 1991. [not for general circulation]
- *15. "The Process of Institutional Development: Assisting Small Enterprise Institutions to Become More Effective." Elaine Edgcomb and James Cawley. GEMINI Working Paper No. 15. February 1991. \$9.70.
16. "Baseline Surveys of Micro and Small Enterprises: An Overview." Donald C. Mead, Yacob Fisseha, and Michael McPherson. GEMINI Working Paper No. 16. March 1991. \$2.60.
17. "Kenya: Kibera's Small Enterprise Sector — Baseline Survey Report." Joan Parker and C. Aleke Dondo. GEMINI Working Paper No. 17. April 1991. \$6.40.
- *18. "A Financial Systems Approach to Microenterprises." Elisabeth Rhyne and Maria Otero. GEMINI Working Paper No. 18. April 1991. \$3.00
- *19. "Agriculture, Rural Labor Markets, and the Evolution of the Rural Non-farm Economy." Steve Haggblade and Carl Liedholm. GEMINI Working Paper No. 19. May 1991. \$2.50.
- *20. "The Microenterprise Finance Institutions of Indonesia and Their Implications for Donors." Elisabeth Rhyne. GEMINI Working Paper No. 20. June 1991. \$3.40.
21. "Microenterprise Growth Dynamics in the Dominican Republic: The ADEMI Case." Frank F. Rubio. GEMINI Working Paper No. 21. June 1991. \$3.10.
- *22. "Credit Unions: A Formal Sector Alternative for Financing Microenterprise Development." John H. Magill. GEMINI Working Paper No. 22. September 1991. \$3.80.
23. "A Proposed Subsector-Based Monitoring and Evaluation System for CARE/Thailand's Silk Promotion Efforts." Steven Haggblade. GEMINI Working Paper No. 23. September 1991. \$3.60.
24. "Steps to the Creation of a Viable Financial Institution for Microenterprise Development in the Philippines: Notes on a Process for the Staff and Board of Tulay sa Pag-Unlad, Inc." Doug Salloum and Nan Borton. GEMINI Working Paper No. 24. November 1991. \$2.00
- *25. "Village Banking: A Cross-Country Study of a Community-Based Lending Methodology." Sharon L. Holt. GEMINI Working Paper No. 25. December 1991. \$12.60.
26. "Dynamics of Small- and Micro-scale Enterprises and the Evolving Role of Finance." Carl Liedholm. GEMINI Working Paper No. 26. December 1991. \$3.00
- *27. "Opportunities for Intervention in Thailand's Silk Subsector." Steven Haggblade and Nick Ritchie. GEMINI Working Paper No. 27. January 1992. \$3.20.
- *28. "Apex Study of the Asociación de Grupos Solidarios de Colombia." Arelis Gomez Alfonso, with Nan Borton and Carlos Castello. GEMINI Working Paper No. 28. April 1992. \$4.60. [See Technical Reports No. 36 and No. 39 for apex studies in Senegal and Thailand.]
- 178

29. "The Subsector Methodology, A Field Orientation for CARE/Egypt, January 20-February 7, 1992." William Grant. GEMINI Working Paper No. 29. April 1992. \$9.50.
30. "'Poverty Lending' and Microenterprise Development: A Clarification of the Issues." Mohini Malhotra. GEMINI Working Paper No. 30. May 1992. \$3.60.
31. "The Solidarity Group Experience." Shari Berenbach and Diego Guzman. GEMINI Working Paper No. 31. June 1992. \$5.80.
32. "A New View of Finance Program Evaluation." Elisabeth Rhyne. GEMINI Working Paper No. 32. November 1992. \$1.50.
33. "The Role of Savings in Local Financial Markets: The Indonesian Experience." Marguerite S. Robinson. GEMINI Working Paper No. 33. November 1992. \$3.50.
34. "Small and Medium Enterprise Development in Poland." Adam Saffer, Miroslaw Zielinski, Jerzy Zielinski, Tadeusz Marek, and Matthew Gamser. GEMINI Working Paper No. 34. February 1993. \$5.20.

GEMINI Technical Reports:

1. "Jamaica Microenterprise Development Project: Technical, Administrative, Economic, and Financial Analyses." Paul Guenette, Surendra K. Gupta, Katherine Stearns, and James Boomgard. GEMINI Technical Report No. 1. June 1990. [not for general circulation]
2. "Bangladesh Women's Enterprise Development Project: PID Excerpts and Background Papers." Shari Berenbach, Katherine Stearns, and Syed M. Hashemi. GEMINI Technical Report No. 2. October 1990. \$13.00
3. "Maroc: Conception d'une Enquête pour une Etude du Secteur Informel." Eric R. Nelson and Housni El Ghazi. GEMINI Technical Report No. 3. November 1990. \$12.50.
4. "Small Enterprise Assistance Project II in the Eastern Caribbean: Project Paper." James Cotter, Bruce Tippet, and Danielle Heinen. GEMINI Technical Report No. 4. October 1990. [not for general circulation]
5. "Technical Assessment: Rural Small-Scale Enterprise Pilot: Credit Activity in Egypt." John W. Gardner and Jack E. Proctor. GEMINI Technical Report No. 5. October 1990. \$4.00
- *6. "Developing Financial Services for Microenterprises: An Evaluation of USAID Assistance to the BRI Unit Desa System in Indonesia." James J. Boomgard and Kenneth J. Angell. GEMINI Technical Report No. 6. October 1990. \$9.00
7. "A Review of the Indigenous Small Scale Enterprises Sector in Swaziland." David A. Schrier. GEMINI Technical Report No. 7. October 1990. [not for general circulation]
8. "Ecuador Micro-Enterprise Sector Assessment: Summary Report." John H. Magill and Donald A. Swanson. GEMINI Technical Report No. 8. April 1991. \$10.20.

9. "Ecuador Micro-Enterprise Sector Assessment: Financial Markets and the Micro- and Small-scale Enterprise Sector." Richard Meyer, John Porges, Martha Rose, and Jean Gilson. GEMINI Technical Report No. 9. March 1991. \$16.00
10. "Ecuador Micro-Enterprise Sector Assessment: Policy Framework." Bruce H. Herrick, Gustavo A. Marquez, and Joseph F. Burke. GEMINI Technical Report No. 10. March 1991. \$11.30.
11. "Ecuador Micro-Enterprise Sector Assessment: Institutional Analysis." Peter H. Fraser, Arelis Gomez Alfonso, Miguel A. Rivarola, Donald A. Swanson, and Fernando Cruz-Villalba. GEMINI Technical Report No. 11. March 1991. \$25.00
12. "Ecuador Micro-Enterprise Sector Assessment: Key Characteristics of the Micro-Enterprise Sector." John H. Magill, Robert Blaney, Joseph F. Burke, Rae Blumberg, and Jennifer Santer. GEMINI Technical Report No. 12. March 1991. \$19.60.
13. "A Monitoring and Evaluation System for Peace Corps' Small Business Development Program." David M. Callihan. GEMINI Technical Report No. 13. [not available for general circulation]
14. "Small-Scale Enterprises in Lesotho: Summary of a Country-Wide Survey." Yacob Fisseha. GEMINI Technical Report No. 14. February 1991. \$6.40.
- *15. "An Evaluation of the Institutional Aspects of Financial Institutions Development Project, Phase I in Indonesia." John F. Gadway, Tantri M. H. Gadway, and Jacob Sardi. GEMINI Technical Report No. 15. March 1991. \$8.80.
- *16. "Small-Scale Enterprises in Mamelodi and Kwazakhele Townships, South Africa: Survey Findings." Carl Liedholm and Michael A. McPherson. GEMINI Technical Report No. 16. March 1991. \$4.60.
17. "Growth and Change in Malawi's Small and Medium Enterprise Sector." Michael A. McPherson. GEMINI Technical Report No. 17. June 1991. \$2.20.
18. "Burkina Faso Microenterprise Sector Assessment and Strategy." William Grant, Matthew Gamser, Jim Herne, Karen McKay, Abdoulaye Sow, and Sibry Jean-Marie Tapsoba. GEMINI Technical Report No. 18. August 1991. Volume One, Main Report, \$7.60; Volume Two, Annexes, \$14.20.
- *19. "Women in the BPD and Unit Desa Financial Services Programs: Lessons from Two Impact Studies in Indonesia." Sharon L. Holt. GEMINI Technical Report No. 19. September 1991. \$3.80.
20. "Mali Microenterprise Sector Assessment and Strategy." William Grant, Kim Aldridge, James Bell, Ann Duval, Maria Keita, and Steve Haggblade. GEMINI Technical Report No. 20. October 1991. Volume One, Main Report, \$6.70; Volume Two, Annexes, \$13.00
21. "A Microenterprise Sector Assessment and Development Strategy for A.I.D. in Zambia." Eric L. Hyman, Robert Strauss, and Richard Crayne. GEMINI Technical Report No. 21. November 1991. \$10.00
22. "Bangladesh: Women's Enterprise Development Project Paper." GEMINI Technical Report No. 22. August 1991. [not for general circulation]

23. "Peru: Small Business and Employment Expansion Project Paper." GEMINI Technical Report No. 23. November 1991. [not for general circulation]
24. "A Country-wide Study of Small-Scale Enterprises in Swaziland." Yacob Fisseha and Michael A. McPherson. GEMINI Technical Report No. 24. December 1991. \$5.40.
- *25. "Micro and Small-Scale Enterprises in Zimbabwe: Results of a Country-wide Survey." Michael A. McPherson. GEMINI Technical Report No. 25. December 1991. \$5.00
26. "The Development Impact of Financing the Smallest Enterprises in Indonesia." GEMINI Technical Report No. 26. January 1992. [not for general circulation]
27. "Midterm Evaluation of the ASEPADE Component of the Small Business II Project, Honduras." Arelis Gomez Alfonso, Wesley Boles, and Donald L. Richardson. GEMINI Technical Report No. 27. February 1992. \$5.80. Also available in Spanish.
28. "Midterm Evaluation of the ANDI/PYME Component of the Small Business II Project, Honduras." Arelis Gomez Alfonso, Wesley Boles, and Donald L. Richardson. GEMINI Technical Report No. 28. February 1992. \$6.60. Also available in Spanish.
29. "The Role of Financial Institutions in the Promotion of Micro and Small Enterprises in Burkina Faso." John McKenzie. GEMINI Technical Report No. 29. February 1992. \$10.40.
30. "Small and Micro Enterprise Development Project No. 262-0212, Egypt. Midterm Evaluation." Katherine Stearns. GEMINI Technical Report No. 30. March 1992. \$7.60.
31. "A Review of the Prospects for Rural Financial Development in Bolivia." James J. Boomgard, James Kern, Calvin Miller, and Richard H. Pattern. GEMINI Technical Report No. 31. March 1992. \$4.60.
32. "The Role of Private Sector Advocacy Groups in the Sahel." William Grant. GEMINI Technical Report No. 32. March 1992. \$2.40.
- *33. "Access to Credit for Poor Women: A Scale-up Study of Projects Carried Out by Freedom from Hunger in Mali and Ghana." Jeffrey Ashe, Madeline Hirschland, Jill Burnett, Kathleen Stack, Marcy Eiland, and Mark Gizzi. GEMINI Technical Report No. 33. March 1992. \$11.80.
- *34. "Egyptian Women and Microenterprise: the Invisible Entrepreneurs." C. Jean Weidemann. GEMINI Technical Report No. 34. March 1992. \$11.20.
- *35. "A Pre-Project Identification Document Analysis of the Lesotho Agricultural Enterprise Initiatives Project." Mike Bess, Don Henry, Donald Mead, and Eugene Miller. GEMINI Technical Report No. 35. April 1992. \$20.00
36. "Apex Study of the Small Enterprise Development Program of Catholic Relief Services, Senegal." Arelis Gomez Alfonso. GEMINI Technical Report No. 36. May 1992. \$3.00
37. "The Private Operators' Perspective on an Agenda for Action," Dakar, Senegal, November 22-25, 1991. A Seminar on the Private Sector in West Africa. Organized by the Senegalese National

131

Employers' Union (CNP), the Club du Sahel, CILSS and USAID. GEMINI Technical Report No. 37. May 1992. \$7.00

38. "Background Documents to the Seminar on the Private Sector in West Africa," Dakar, Senegal. November 22-25, 1991. Technical Report No. 38. May 1992. \$5.00

39. "Apex Study of the Small Enterprise Development Program of Catholic Relief Services, Thailand." Arelis Gomez Alfonso. GEMINI Technical Report No. 39. May 1992. \$3.20.

40. "Study of Informal Cross-border Trade, Poland." SMG-KRC/Poland. GEMINI Technical Report No. 40. May 1992. \$3.20.

41. "Study of the Informal Commercial Sector, Poland." SMG/KRC Poland. GEMINI Technical Report No. 41. May 1992. \$4.20

42. "Evaluation of the Micro and Small Enterprise Development Project (MSED) in Bolivia." William Fisher, Jeffrey Poyo, and Ann Beasley. GEMINI Technical Report No. 42. June 1992. \$10.60. Also available in Spanish.

43. "Analysis of Funding Mechanisms for the Small and Micro Enterprise Development Project, Egypt." Kenneth J. Angell and John M. Porges. GEMINI Technical Report No. 43. June 1992. \$3.80.

44. "Get Ahead Foundation Credit Programs in South Africa: The Effects of Loans on Client Enterprises." Jennefer Sebstad. GEMINI Technical Report No. 44. June 1992. \$3.00

45. "Get Ahead Foundation in South Africa: Final Evaluation." Robert Christen, Elisabeth Rhyne, Doug Salloum, and Jennefer Sebstad. GEMINI Technical Report No. 45. June 1992. \$11.00

46. "Micro- and Small-Scale Enterprises in Botswana: Results of a Nationwide Survey." Lisa Daniels and Yacob Fisseha. GEMINI Technical Report No. 46. August 1992. \$9.40.

*47. "The Growth and Dynamics of Women Entrepreneurs in Southern Africa." Jeanne Downing and Lisa Daniels. GEMINI Technical Report No. 47. August 1992. \$3.10.

48. "Small Business Development Programming Trip: Peace Corps/Albania and the Office of Training and Program Support, Small Business Development Sector." Lauren Spurrier and Wesley Weidemann. GEMINI Technical Report No. 48. October 1992. \$6.00

49a. "Small Enterprise Development in the Russian Far East." Martha Blaxall, Yasuo Konishi, Virginia Lambert, Jennifer Santer, and Timothy Smith. GEMINI Technical Report No. 49a. October 1992. \$12.00 [not for general circulation]

49b. "Supporting Private Enterprises in Uzbekistan: Challenges and Opportunities." Nan Borton, John Magill, Neal Nathanson, and Jim Packard Winkler. GEMINI Technical Report No. 49b. November 1992. \$5.60. [not for general circulation]

49c. "Assessing the Prospects for Small Enterprise Development in Kazakhstan." Kenneth Angell, James J. Boomgard, Mohini Malhotra, and Robert A. Rodriguez. GEMINI Technical Report No. 49c. December 1992. \$3.90. [not for general circulation]

49d. "Small Enterprise Development in Ukraine." Dennis De Santis, Jean Gilson, Max Goldensohn, Jennifer Santer, and Timothy Smith. GEMINI Technical Report No. 49d. December 1992. \$8.10. [not for general circulation]

*50. "Skins and Hides in Four Countries in Africa: The Potential Role for Micro- and Small-Scale Enterprise Development." William Grant. GEMINI Technical Report No. 50. November 1992. \$3.00. Also available in French.

51a. "Morocco: Assessment of Programming Options for Microenterprise Development." Housni El Ghazi, Sheila Reines, Steve Silcox, Katherine Stearns, and Matthew Gamser. GEMINI Technical Report No. 51A. November 1992. [not for general circulation]

51b. "USAID/Morocco: Assessment of Programming Options for Microenterprise Development. Report on Workshop and Field Investigations." Matt Gamser, Housni El Ghazi, Sheila Reines, Steve Silcox, and Katherine Stearns. GEMINI Technical Report No. 51b. December 1992. Also available in French. [not for general circulation]

52. "Small Enterprise Development in Armenia: Programming Recommendations for Peace Corps Volunteers." Timothy J. Smith. GEMINI Technical Report No. 52. July 1992. \$2.20.

53. "Results of a Nationwide Survey on Micro, Small, and Medium Enterprises in Malawi." Lisa Daniels and Austin Ngwira. GEMINI Technical Report No. 53. January 1993. \$11.80.

*54. "A Review of Donor-Funded Projects in Support of Micro- and Small-Scale Enterprises in West Africa." William Grant. GEMINI Technical Report No. 54. February 1993. \$18.80.

Technical Notes:

Financial Assistance to Microenterprise Section:

*1. Series Notebook: Tools for Microenterprise Programs (a three-ring binder, 1 1/2 inches in diameter, for organizing technical notes and training materials) and "Methods for Managing Delinquency" by Katherine Stearns. \$7.50. Also available in Spanish and in French.

*2. "Interest Rates and Self-Sufficiency." Katherine Stearns. \$6.50. Also available in Spanish and in French.

*3. "Financial Services for Women." C. Jean Weidemann. March 1992. \$5.00

Nonfinancial Assistance to Microenterprise Section:

*1. "A Field Manual for Subsector Practitioners." Steven J. Haggblade and Matthew Gamser. \$4.65. Also available in Spanish and in French.

*2. "Facilitator's Guide for Training in Subsector Analysis." Marshall A. Bear, Cathy Gibbons, Steven J. Haggblade, and Nick Ritchie. December 1992. \$20.00

Field Research Section:

*1. "A Manual for Conducting Baseline Surveys of Micro- and Small-scale Enterprises." Michael A. McPherson and Joan C. Parker. February 1993. \$13.60.

Special Publications:

*1. *Training Resources for Small Enterprise Development.* Small Enterprise Education and Promotion Network. Special Publication No. 1. 1992. \$11.00

*2. *Financial Management of Micro-Credit Programs: A Guidebook for NGOs.* Robert Peck Christen. ACCION International. Special Publication No. 2. 1990. \$19.00

*3. *The ADEMI Approach to Microenterprise Credit.* A. Christopher Lewin. Special Publication No. 3. 1991. \$15.00

*4. *Microempresas y Pequeñas Empresas en la República Dominicana. Resultados de una Encuesta Nacional.* Miguel Cabal. Michigan State University and FondoMicro. Special Publication No. 4. 1992. \$9.00

Copies of publications available for circulation can be obtained by sending a check or a draft drawn on a U.S. bank to the DAI/GEMINI Publications Series, Development Alternatives, Inc., 7250 Woodmont Avenue, Bethesda, MD 20814, U.S.A.