DISCUSSION PAPER

Report No. UDD-67

MONITORING AND EVALUATING URBAN DEVELOPMENT PROGRAMS: A Handbook for Program Managers and Researchers

by

Michael Bamberger
and
Eleanor Hewitt

January, 1985

Water Supply and Urban Development
Operations Policy Staff
The World Bank

During the period in which this Handbook was prepared, Michael Bamberger was a Sociologist with the Water Supply and Urban Development Department of the World Bank. From January 1, 1985 he has been a Regional Training Coordinator with the Economic Development Institute. Eleanor Hewitt is a Research Assistant with the Water Supply and Urban Development Department.

Work on the Handbook began as part of a collaborative evaluation project between the International Development Research Center (IDRC) of Canada and the World Bank. Many of the cited studies were produced as part of this project. The authors would like to thank IDRC for their support throughout this cooperative venture.

Guidance and comments have been received from many people within the Bank. Special thanks are expressed to Douglas Keare who was Chief of the Urban and Regional Economics Division, where work on the Handbook began in 1980. Within the Water Supply and Urban Development Department helpful comments have been received from Michael Cohen, Andrew Hamer, Friedrich Kahnert, Robert Satin, Stephen Mayo, Kyu Sik Lee, Stephen Malpezzi and Ronald Parlato. Comments and suggestions have also been received at various points from colleagues in the Evaluation and Sociology Groups, in particular Dennis Casley, Barbara Searle, Michael Cernea, Teresa Ho and Ronald Ng. The authors are also grateful to their urban projects colleagues, Evan Rotner, Alberto Smith, Mario Rothschild, Edward Echeverria, Neil Boyle, Faye Johnson, Richard Westin, Aura Garcia, Thakoor Persaud and Carolyn Gochenour who commented on various drafts and helped with field testing. Lawrence Salmen, a consultant with the Bank, has provided helpful guidance on the role of qualitative evaluation.

Finally, the authors would like to thank the following local researchers and project staff who helped with field testing: Januario Flores and Antonio Jorge Araujo (Brazil), Jaime Medrano and Freddy Quiton (Bolivia), Juardo Velez (Colombia), Mila Reforma (the Philippines) and Gibson Maina, Angela Hamau, Davinder Lamba and Diana Lee Smith (Kenya).

ABSTRACT

"Monitoring and Evaluating Urban Development Programs: A Handbook for Managers and Researchers" is a comprehensive but easily understood handbook for urban policy makers, managers and evaluation practitioners in developing countries. It provides guidance on all stages of the design and implementation of a monitoring and evaluation system and presents the main options with respect to scope, key research issues and organization. Monitoring and evaluation systems are described which can be applied to both individual projects and to integrated multi-component urban development programs.

Urban development projects vary widely in scope and complexity, and in terms of the resources which are available for monitoring and evaluation. The Handbook is designed to help managers and policy makers decide on the types and complexity of the studies which are most appropriate for their project, and to select among the range of available research and analytical procedures. A distinction is made throughout between basic monitoring and evaluation techniques which are simple and economical to apply in any project, and more complex techniques which are only appropriate in certain circumstances.

All of the methods described in the Handbook have been field tested, most but not all of them as part of World Bank projects. The unique contribution of the Handbook is to show how approaches taken from the fields of sociology, economics, anthropology and accountancy can be combined in an integrated monitoring and evaluation strategy.

TABLE OF CONTENTS

(Continued)

(Table of Contents - Cont.)

(Continued)

(Table of Contents - Cont.)

(Continued)

(Table of Contents - Cont.)

(Continued)

(Table of Contents - Cont.)

(Continued)

(Table of Contents - Cont.)

(Continued)

(Table of Contents - Cont.)

(Continued)

(Table of Contents - Cont.)

TABLES AND CHARTS

(Continued)

(List of Tables - Cont.)

(Continued)

(List of Tables - Cont.)

(Continued)

(List of Tables - Cont.)

FIGURES

CASE STUDIES (Boxes)

(Continued)

(List of Case Studies - Cont.)

INTRODUCTION - A USERS GUIDE TO THE HANDBOOK

This Handbook is designed as an easily understandable guide to the monitoring and evaluation of urban development projects. It is written primarily for policy makers, managers and evaluation practitioners in developing countries, but it will also be useful to international development agencies and other readers involved in urban policy and research.

Monitoring and evaluation are practical tools which should form an essential part of good management practice. Monitoring is an internal project activity which assesses (a) whether project resources (money, materials, staff etc) are being delivered and used in accordance with the approved budget and timetable (b) whether the intended outputs (numbers of houses constructed, training courses given, patients treated etc) are being produced in a timely and cost-effective manner and (c) assesses the efficiency with which the project is being implemented.

The primary purposes of evaluation are (a) to assess the extent to which the intended impacts (increases in income, reduced incidence of certain infections, improved housing quality etc) have been produced and (b) to compare the cost-effectiveness of a project with possible alternatives.

The following examples, based on the experience of urban development projects during the past decade, illustrate the importance of effective monitoring and evaluation systems:

** Even the most carefully designed projects undergo substantial modifications during the process of implementation. Timely and appropriate decisions on project modification can only be made if rapid feedback is received throughout the implementation process.

** Many projects have social objectives (reaching certain economic or cultural groups, developing community institutions etc.). Monitoring these social objectives requires the regular presentation of socio-economic indicators which cannot be obtained from the administrative reports produced by most programs.

** As urban development strategies increase their scope and complexity, it becomes increasingly difficult for a central coordinating agency to monitor each project component and to have a means of evaluating overall progress. Consequently there is a demand for a system which can rapidly provide a set of indicators on the progress of each component and of the project in general.

** Most projects are part of an ongoing urban development strategy in which lessons from one project are used as inputs in the design of subsequent projects. Development planners require information on the contribution of particular projects to overall development goals, the impacts on particular target groups and a comparison of the cost-effectiveness of alternative approaches.

A well designed monitoring and evaluation system can contribute to all of these issues as well as providing most of the basic information required for implementing the project and for satisfying the reporting requirements of government and international agencies.

The methods described in this Handbook are based on 10 years of World Bank experience and a review of the extensive evaluation literature produced since the early Seventies. The World Bank urban evaluation experience began in 1975 with a cooperative venture with the International Development Research Centre (Ottawa) which supported a 5 year evaluation of 4 of the first World Bank financed urban shelter projects in El Salvador, Zambia, Senegal and the Philippines. Since then the Bank has provided assistance to governments in Asia, Africa and Latin America in the design of their own monitoring and evaluation systems; and new evaluation techniques designed to provide more rapid feedback or a better understanding of the point of view of intended beneficiaries, have been tested.

All of the methods described in the Handbook have already been field tested. The unique contributions of the Handbook are (a) to bring together, for the first time, and in an easily understandable form, approaches taken from the fields of sociology, economics, anthropology and accountancy (b) to combine the experiences of both developing and industrialized countries (c) to present examples from studies which have been conducted in all parts of the World and (d) to show how all of these techniques can be combined in an integrated monitoring and evaluation strategy.

One of the main problems in writing a comprehensive Handbook is that readers have different interests and levels of research experience. A number of methods have been used to help managers and other readers who require a general overview of the system, to locate the sections of interest and to avoid material which is too technical or detailed for their purposes:

(a) A brief Managers Guide has been prepared which outlines the essential elements of the proposed systems and which contains extensive cross-references to the appropriate sections of the Handbook.

(b) More technical and detailed material has been placed in Annexes. The chapters contain cross references to these annexes.

(c) An extensive bibliography is included to guide readers interested in a more detailed treatment of particular issues.

The Handbook has 6 chapters and 10 annexes containing more detailed technical material. The subject areas of each chapter can be summarized as follows:

Chapter 1: The Monitoring and Evaluation Framework

This chapter presents the framework within which the objectives, scope and organization of a monitoring and evaluation system are defined. Monitoring and evaluation are management tools and it is important that the project manager be actively involved in the design, review and application of the studies. The system should include performance monitoring (to control the use of inputs and the production of outputs), process monitoring (to evaluate the efficiency and effectiveness with which the project is implemented), impact evaluation to estimate the quantitative effects of the project on the social and economic conditions of the target population; and cost-effectiveness analysis to compare alternative projects in terms of the outputs produced for a given cost. The key decisions which management must take with respect to the planning and management of the studies are discussed and finally a number of common problems in the design and implementation of monitoring and evaluation are discussed, and some possible solutions are proposed.

Chapter 2: Performance Monitoring

This chapter describes the design and implementation of a system to provide periodic feedback on the progress of a project, the extent to which inputs are being used in accordance with the approved budget and timetable, and whether the intended outputs are being produced in a timely and cost-effective

manner. Two systems are described, a basic system which can be simply and economically applied to any project; and a more complex system, based on network analysis, which is more appropriate for larger and more complex projects.

Chapter 3: Process Monitoring: Monitoring the Project Delivery System

This chapter reviews key issues and designs for monitoring the efficiency and effectiveness of the project implementation process. Among the issues discussed are the trade-offs between different indicators of efficiency, and some of the organizational factors which have prevented many projects from effectively monitoring the implementation process. Process monitoring can either be used to provide regular information on the progress of implementation, or to provide rapid feedback when problems arise. The main data collection methods are presented, with the recommendation that a multi-method approach should always be used in which quantititative and qualitative methods are combined. Three study designs are described: continuous observation throughout a project, periodic studies, and studies conducted at only one point in time. Techniques are described for monitoring the efficiency of the implementation process. Finally recommendations are presented on how to define the appropriate strategy for process monitoring.

Chapter 4: Impact Evaluation and Cost-Effectiveness Analysis

This chapter presents alternative research designs for the evaluation of project impacts. The history of impact evaluation is reviewed and the debate between the advocates of qualitative and quantitative approaches is discussed. Key issues, including whether and when to conduct impact evaluations, are discussed and examples of different research designs are presented. A number of simple evaluation designs are presented as alternatives to the large-scale quantitative approaches in those situations where it is not necessary to obtain precise quantitative estimates of project impacts. With respect to quantitative evaluation, a distinction is made between approaches which estimate net project impacts (the quasi-experimental design) and those which estimate cost-effectiveness (cost-benefits analysis, cost-effectiveness analysis and cost-utility analysis). Finally, guidelines are presented on how to choose the appropriate impact evaluation strategy.

The chapter includes cross-references to technical material in the annexes on sampling, research design and statistical analysis.

Despite efforts to simplify the presentation, this is the most difficult chapter to follow for those readers with limited social science research experience. Some readers may wish to skip parts of this chapter and to leave the details to the researchers who will be directly responsible for the design of the impact evaluation studies.

Chapter 5: Managing the Evaluation

This Chapter discusses the main issues involved in defining the appropriate organizational structure for monitoring and evaluation at the level of the implementing agencies, the local coordinating agency, specialized sectoral agencies and national development and financial agencies. Some of the issues include: defining who should conduct the evaluation, the role of consultants, the appropriate organizational location of the monitoring and evaluation units, the role of a steering committee, and the distribution of evaluation responsibility between the national, sectoral and local agencies. There is no single best organizational structure and the location of the monitoring and/or evaluation units is determined in each case by the scope and complexity of the project and the relative size and research experience of the different agencies involved. Guidelines are provided for estimating the financial and human resource requirements for different types of monitoring and evaluation programs. The final section discusses common problems in the organization of an evaluation program and some of the possible solutions.

Chapter 6: Issues and approaches in the evaluation of new style urban development projects

This chapter reviews some of the new directions in urban development which have evolved in recent years and discusses the extent to which the monitoring and evaluation framework presented in earlier chapters is applicable to them. The 4 types of projects which are discussed are: income and employment generation, health, transport and urban and municipal development. It is concluded that the techniques of performance monitoring can be readily applied to all of these new types of projects, and that process monitoring can be easily applied to the first three and with some difficulty

to municipal and institutional development. The main difficulties occur in the evaluation of the impacts of the projects. Problems arise because the size and scope makes it difficult to identify a control group, because it is difficult to specify and measure impacts or because the project does not have a single set of outcomes and impacts which can be clearly defined and measured. Strategies are recommended for the application of each type of monitoring and evaluation study to each of the 4 project areas.

CHAPTER 1: THE MONITORING AND EVALUATION FRAMEWORK

This chapter presents the framework within which the objectives, scope and organization of a monitoring and evaluation system are defined. Monitoring and evaluation are management tools and it is important that the project manager is actively involved in the design, review and application of the studies. The system should include performance monitoring (to control the use of inputs and the production of outputs); process monitoring (to assess the efficiency and effectiveness with which the project is implemented); impact evaluation to estimate the quantitative effects of the project on the social and economic conditions of the target population; and cost-effectiveness analysis to compare alternative projects in terms of the output produced for a given cost. The key decisions which management must take with respect to the planning and management of the studies are discussed. Finally a number of common problems which arise in the design and implementation of monitoring and evaluation are discussed, and some possible solutions are proposed.

A. MONITORING AND EVALUATION: ESSENTIAL MANAGEMENT TOOLS

Monitoring and evaluation are essential management tools which help improve the efficiency of on-going projects and the selection and design of future projects. Monitoring is an internal project activity designed to provide constant feedback on the progress of a project, the problems it is facing, and the efficiency with which it is being implemented. Projects which do not have an effective monitoring system are more likely to suffer delays and cost overruns; to exclude or under-represent certain sectors of the target population; to have problems of quality control; or to take longer to detect antagonisms among the implementing agencies or between the agencies and beneficiaries.

Evaluation, on the other hand, is mainly used to help in the selection and design of future projects. Evaluation studies can assess the extent to which the project produced the intended impacts (increases in income, better housing quality etc) and the distribution of the benefits between different groups, and can evaluate the cost-effectiveness of the project as compared with other options. When a project does not have an effective evaluation system the danger is increased of deciding to continue a project which is not producing the intended benefits; money may be wasted by not selecting the most cost-effective option; and it may be more difficult to detect and correct some of the factors which are reducing project impact.

The following are examples of monitoring studies which were of direct operational utility to project managers. The first study (Box 1-1) examined reasons for the slow rates of house construction and occupancy in Dakar, Senegal. An "Action Plan", based on the findings of this study, was successful in getting families to build smaller houses which could be completed more economically and rapidly. The second study (Box 1-2) found that renters and poor households living on the periphery of an upgrading project in La Paz, Bolivia, were not sharing in most of the project benefits such as domestic water and paved roads. As a result of the study the project began to require landlords to give tenants access to water, and discussions began with families in some of the peripheral areas about ways in which they could be incorporated into the project. The third study (Box 1-3) examined reasons for the poor performance of an artesan credit and assistance project in Campina Grande, Brazil. Weaknesses in the operating methods of the credit

and training agencies were identified, and after some very lively discussions, many of the problems were corrected. Each of the above monitoring studies was conducted at the request of management and each of them provided operationally useful results economically and quickly (one of the studies was completed within 4 weeks and the other two within 3 months).

BOX 1-1 HOW A RAPID MONITORING STUDY HELPED DEVELOP AN ACTION PLAN TO SPEED UP THE RATE OF HOUSE CONSTRUCTION AND OCCUPATION IN A SITES AND SERVICES PROJECT IN DAKAR, SENEGAL.

A monitoring study was conducted in Daker, Senegal at the request of project management to determine the reasons for the slow rate of plot occupancy in a early sites and services project. The study showed that most of the houses were much larger and more expensive than intended and that this was both slowing down occupancy and excluding lower income families. Some of the reasons included: (a) administrative difficulties in obtaining approval of plans for small houses; (b) families were not aware of the cost implications of designing a larger house; (c) families were not familiar with the concept of progressive development whereby a small core house could be built and occupied and later expanded; and (d) many higher income families were able to obtain a plot due to the lax screening procedures. As a result of the study, an "Action plan" was initiated which successfully encouraged people to build smaller houses and which imposed stricter screening procedures on applicants. Technical assistance was provided in recruiting and supervising subcontractors and the procedures for approving the plans for smaller houses were streamlined.

The study, which involved interviews with a sample of participants and discussions with project staff, was completed in under 3 months by a team of 4 interviewers working with a supervisor and research director.

Source: Bureau d'Avaliacion. "Study of purchasers of serviced plots who have not yet begun construction". Office d'Habitation de Loyer Moderee. Dakar. Senegal. November 1979.

Chapter 4 presents examples of how impact evaluation and cost-effectiveness analysis can help improve the selection and design of future projects. Box 4-1 (Chapter 4) describes how a rapid impact study demonstrated the need to define much more precisely and realistically the intended beneficiaries and impacts of a small business credit program in Francophone Africa. Box 4-2 describes a Cost-Benefit study which compared the social and

economic efficiency of current approaches to the provision of low-cost housing in El Salvador. It was shown that progressive development approaches such as sites and services and squatter upgrading scored much higher on the indicators of economic and social efficiency (internal economic rate of return and the ratio of net present value to costs) than did any of the conventional housing programs. Consequently, if the progressive development approach were implemented on a larger scale it could significantly improve the efficiency of future housing strategies. The final example is the evaluation of the US Experimental Housing Allowance Program (described in Section C-1 of Chapter 4). Although the evaluation cost around $ 50 million over a 10 year period, it has been estimated that the recommendations which it produced could potentially save the Federal Government up to $ 8 billion per year.

BOX 1-2 USING PARTICIPANT OBSERVATION TO IDENTIFY SECTORS OF THE POPULATION WHO WERE NOT BENEFITING FROM A SQUATTER UPGRADING PROJECT IN BOLIVIA.

One of the findings of a participant observer study in La Paz, Bolivia was that the community leadership in an upgrading project consisted almost exclusively of middle-income houseowners. It was found that the interests of poorer renters, and of low-income households on the periphery of the community, had not been adequately taken into account. As a result of the study, discussions were started with representatives of some of the peripheral areas and plans were made to extend infrastructure to include them.

The study consisted of a trained observer living in the community for 3 months (the participant observation approach is explained in Annex B)

Source: Lawrence Salmen "Participant Observer Evaluation of Urban Projects in La Paz, Bolivia and Guayaquil, Ecuador." Water Supply and Urban Development Department Discussion Paper No. 36. Nov. 1983.

These examples show that a well designed and implemented monitoring and evaluation program can be a cost-effective way to:

* Provide constant feedback on the extent to which the projects are achieving their goals.
* Identify potential problems at an early stage and propose possible solutions.
* Monitor the accessibility of the project to all sectors of the target population.
* Monitor the efficiency with which the different components of the project are being implemented and suggest improvements.

* Evaluate the extent to which the project is able to achieve its general objectives.
* Provide guidelines for the planning of future projects.

BOX 1-3 HOW A RAPID MONITORING STUDY HELPED IMPROVE THE PERFORMANCE OF AN ARTESAN CREDIT AND TECHNICAL ASSISTANCE PROGAM IN BRASIL

A rapid survey was conducted in Campina Grande, Brazil, to determine the reasons for the disappointing performance of an artesan credit and technical assistance program. It was found that most artesans considered that the amount of credit was too small to permit purchases of machinery or the hiring of more workers. Very few had participated in training courses offered by the project and the number of technical assistance visits from project staff had been quite small. In addition to the smaller than expected number of credits, it was found that about 25% of the beneficiaries were in occupational sectors the project was not intended to cover. As a result of the study a number of important changes were made in the operating procedures of both the credit and training agencies.

The study, which involved interviews with 57 artesans, an analysis of secondary data and conversations with project staff, was completed in 4 weeks.

Source: Unidade d'Avaliacao "Estudo rapido de acompanhamento do projeto de apoio as atividades produtivas." Prefeitura Municipal de Campina Grande. Jan 1984

Many project managers have been discouraged from starting monitoring and evaluation by the belief that these are highly technical fields which should be left to research specialists or which are too expensive and complex to be of practical utility. However, monitoring and evaluation need not be complicated or expensive, and the size and complexity of the studies can be adapted to suit the needs and resources of each project. For example, the monitoring studies described above were all simple, short and economical. In some projects monitoring and evaluation may be conducted by one part-time staff member whose main responsibility is to produce a quarterly report, whereas in other cases there may be a separate monitoring and evaluation unit with several full-time professionals producing a wide range of studies. The role of the manager is to define the topics which need to be studied, to make sure that researchers use the most cost-effective methods, and to arrange for reports to be reviewed, discussed and acted upon.

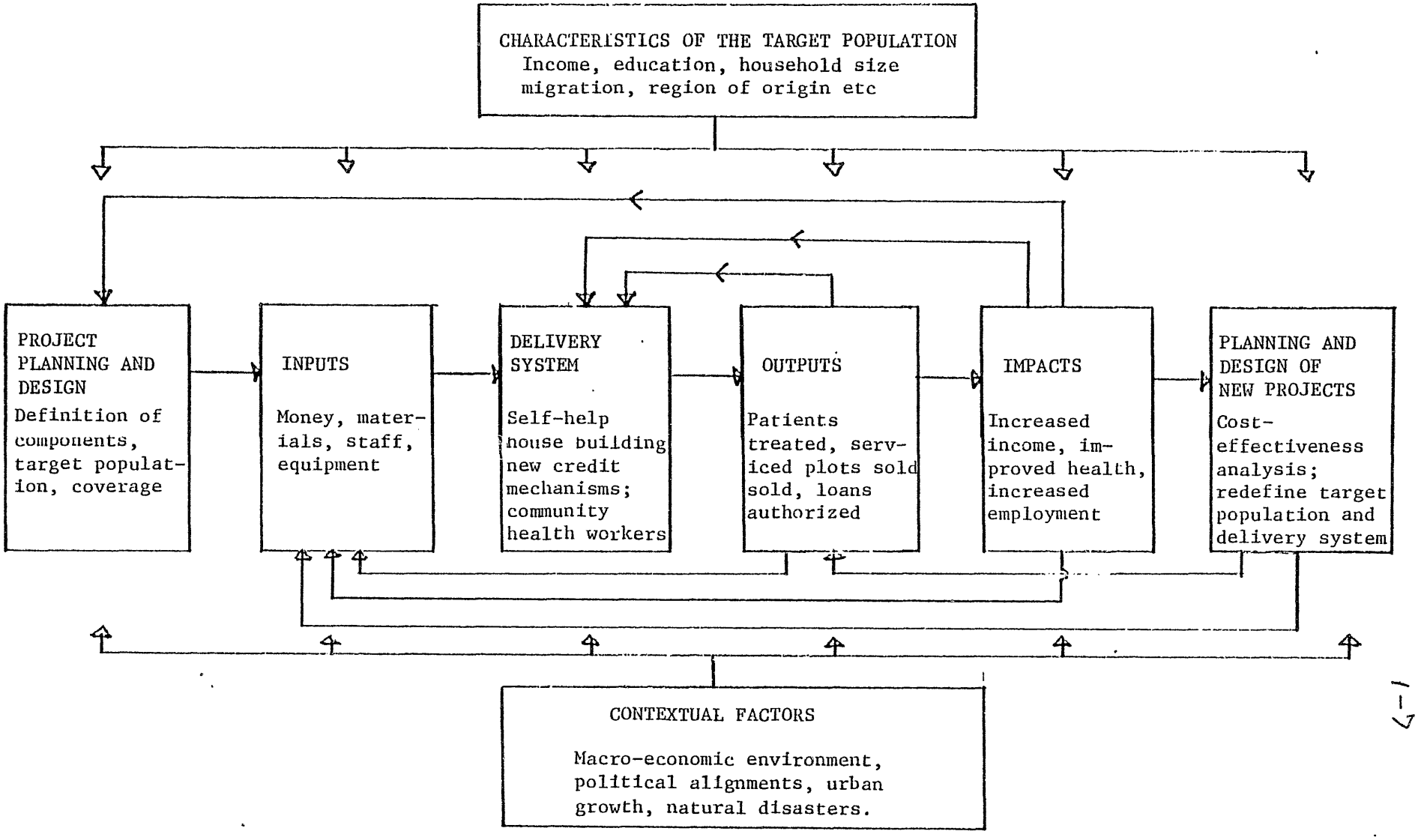
B. FRAMEWORK FOR MONITORING AND EVALUATION OF URBAN DEVELOPMENT PROJECTS

1. A model of the project implementation process

For the purposes of this Handbook, an urban project is defined as the set of urban development activities included in a government grant or authorization or in an agreement with an international development agency. A project will usually include a number of different components each of which may be the responsibility of a different executing agency. Each component may be further divided into sub-components. A program refers to a long term development strategy and will usually include several projects. For example, the First World Bank Urban Development Loan to Bolivia would be defined as a project. This comprised 6 main components: squatter upgrading, 2 separate sites and services components, maternal and child health care, promotion of small scale industries and artesan credit, which were managed by 5 separate agencies. Each component involved a number of sub-components. For example, the artesan credit component included: meetings and other forms of dissemination to inform artesans about the project, visits and selection procedures, lectures and training, technical assistance, approval and supervision of the credits. This project was intended to lay the groundwork for a long-term urban development program which would replicate successful components on a larger scale and which would initiate similar projects in other cities.

The components of an urban development project can range from the provision of basic core housing units to the improvement of municipal tax collection systems. Despite their diversity, any component can be represented by the simple model given in Fig 1.1. The figure shows that the implementation of all urban development components can be broken down into 5 or 6 sequential stages: planning and design, inputs, delivery system, outputs, impacts and in most cases planning and design of future projects. All components, whether solid waste disposal, municipal transport, or small business development begin with a planning and design stage in which objectives are defined and resource requirements are estimated. During the planning stage, assumptions are made about the needs of certain sectors of the population and about the relative effectiveness of alternative delivery systems. Assumptions are also made about how the project will be affected by

Figure 1-1 CONCEPTUAL FRAMEWORK FOR THE ANALYSIS OF THE PROJECT IMPLEMENTATION PROCESS



political and economic factors external to the project, such as upcoming elections, inflation and unemployment.

Once the design is approved, inputs (money, equipment, staff, land etc) are authorized and the intended outputs (number of houses to be built, number of small business loans approved etc) are defined. A series of delivery systems are selected through which the inputs will be used to produce the desired outputs. For example, a housing project must decide whether to organize self-help construction groups or to hire building contractors to construct the housing units (outputs). These are alternative delivery systems. Similarly artesan credit could be administered through normal commercial credit channels, or banks could be given incentives to create special artesan credit programs.

Although the immediate concern of managers is to ensure that the outputs stated in the project agreement are successfully completed, almost all components have a set of impacts they are intended to produce. For example, project planners may hope that the provision of economical core housing will: increase household income (through subletting, because they are nearer to places of employment or because the new house-owners will increase their demand for certain goods and services); reduce geographical mobility (house-owners may have more incentive to remain in the community as the value of their property is appreciating) and have beneficial effects on health (due to better water supply and sanitation). The definition of impacts involves assumptions about how people will respond to project outputs. It may be assumed, for example, that access to credit will induce artesans to hire more labor or purchase more machinery; or that access to drinking water will change hygene practices and the way in which food is prepared.

It is important to know how much importance will be given to the successful production of impacts in decisions about future projects. In some cases projects are judged mainly on their cost-effectiveness in producing certain outputs (number of houses built, number of patients treated), and relatively little importance is given to whether the project also produced the expected impacts on income, employment, health etc. Planners may argue, for example, that the provision of drinking water and sanitation is an essential pre-condition for long term improvements in public health and that these components are justified irrespective of whether they produce short-term health impacts. In a case like this, where decisions on future projects do

not depend on whether or not health impacts had been produced by the first project, there may be no justification for conducting a rigorous impact evaluation. In other cases where alternative approaches are being compared in order to select the most cost-effective option, the ability of projects to produce certain impacts may be a key factor in deciding on future investment strategies. Before a decision can be made on whether or not to conduct a quantitative impact evaluation, it is essential to understand the importance which policy makers attach to the production of impacts.

Many projects are intended to form part of an ongoing urban development program and frequently the results of the first project are intended to provide guidelines for the definition and design of future projects. If the first project is perceived to have been successful, future projects are more likely to adopt a similar approach; but, if it is perceived to have been unsuccessful, future projects in this field may be cut back or modified. We stress the importance of the perceived results, as most of the decisions on future project design are made without access to systematic information on the results of earlier projects.

Projects do not develop in a vacuum, and success is significantly affected by the characteristics of the target population, and by the political and economic context within which the project is implemented. For example, if identical core housing is provided in a fishing community and in an area where most household heads have permanent employment in factories, there may be significant differences in the outputs and impacts produced by the two projects and in the amount of private housing investment which takes place. The fate of projects is also influenced by the economic and political environment in which they operate. High inflation and unemployment affect affordability; and the election of a new mayor, or a conflict between the central and local government may suddenly increase or decrease resources and support for the project. All of these factors must be taken into account when trying to understand why a project has been more or less successful than expected.

2. The functions of monitoring and evaluation

Monitoring and evaluation are designed to provide project management, and national and international development agencies with timely and operationally useful information on how efficiently each stage of the

project is operating, the degree to which intended impacts are being achieved and the lessons for future projects.

Monitoring is an internal project activity concerned to assess whether project resources (inputs) are being administered and used as intended and whether they are producing the intended outputs. It is useful to distinguish between performance monitoring and process monitoring.

a. The purpose of Performance Monitoring is to assess the extent to which project inputs are being used in accordance with the approved budget and timetable and whether the intended outputs are being produced in a timely and cost-effective manner. It may also assess whether project benefits are reaching the intended population groups. Performance Monitoring is intended to improve project supervision, and it is essential that management receive constant feedback on key indicators of project performance so that problems can be detected and corrections made. The main approaches to Performance Monitoring are described in Chapter 2, and in Annexes G,H and I.

b. The purpose of Process monitoring (referred to by some authors as process evaluation is to provide feedback to management on the efficiency and effectiveness of the project delivery system. Two areas of particular importance are the analysis of how the project is perceived by, and actually operates at the level of, the intended beneficiaries; and the effectiveness of the communication and organizational linkages between the implementing agencies and beneficiaries. The speed and costs of the delivery system and the quality of the outputs are measured, and where possible a comparison is made with alternative approaches. The effects may also be assessed. The methods of process monitoring are described in Chapter 3.

Whereas monitoring is an internal project activity intended to improve the performance of an ongoing project, the purpose of evaluation is to help with the selection and design of future projects. Evaluation can be divided into impact evaluation and cost-effectiveness analysis.

Impact evaluation estimates the net impacts of a project on the target population by comparing the conditions of the affected groups after the project has taken place with what they would have been, had there not been a project. Depending on the types of information required for future project planning, studies can either be designed to provide general descriptive information on impacts, or to produce quantitative estimates of the magnitude of the impacts. The studies estimate the degree to which the observed changes

can be attributed to the effects of the project (have been "caused" by the project) and examine the factors which contribute to the degree and direction of the impacts.

The purpose of Cost-effectiveness analysis is to compare alternative projects in terms of the cost of producing a given output. In the housing allowance program described in Chapter 4, output was defined in terms of the amount invested by households in upgrading their dwelling unit. It was shown that providing loans which families could use to make their own improvements was more cost-effective than alternative programs in which federal agencies were directly responsible for making housing improvements. Both impact evaluation and cost-effectiveness analysis are described in Chapter 4.

Although each of these 4 types of study has different purposes, they complement each other and should be used together in an integrated monitoring and evaluation program. Let us assume, for example, that the evaluation of a water supply and sanitation project finds that the expected improvements in the height and weight of young children have not taken place. The impact evaluation on its own cannot explain whether this finding means that the underlying assumptions of the project are invalid, or whether the problem is due to the way in which the project was implemented. In this case process monitoring could identify any problems which occurred during implementation and could assess the extent to which these problems have contributed to the lack of expected impacts. If serious problems occurred during implementation, it would clearly be difficult to interpret the negative impact findings as evidence that improved water and sanitation do not affect child health. In this case the monitoring studies could greatly increase the operational utility of the impact study. In other cases the impact study can be used to complement the findings of monitoring studies by assessing whether a project which is being implemented very efficiently, is producing the intended impacts. For example, if an efficiently run job training program does not have any impact on income or employment there may be no justification for its being continued.

C. PLANNING THE MONITORING AND EVALUATION SYSTEM

This section decribes the main decisions and actions which a manager must take in selecting and planning the appropriate types of monitoring and evaluation studies for a particular project. The issues are discussed in more detail in Annex C.

1. Defining the issues to be studied

Every study involves costs of money and staff time, and it is therefore essential for management to select carefully the issues which should be covered by the monitoring and evaluation. There are usually a number of different ways in which any topic can be studied, and as the approaches vary considerably in terms of detail, complexity, time and cost it is important for the manager to define carefully exactly what type of information is needed. There are 4 main sets of issues on which the studies can focus and managers must decide the relative importance of each and how resources will be allocated between them:

a. Monitoring the use of project inputs and the production of outputs according to cost and time schedules. Although Performance Monitoring is an essential supervision and control procedure which must be included in all projects, the manager must decide whether to use a relatively simple approach or the develop one of the network based (and possibly computerized) monitoring systems described in Annexes H and I. The basic operational procedures for performance monitoring are described in Chapter 2.

b. Monitoring the process of project implementation. It is strongly recommended that resources are made available to monitor the implementation process as this can contribute to project efficiency and ensure that benefits are reaching the target population. Management must decide whether to only conduct studies when problems have been identified during implementation, or to develop an ongoing program of process monitoring which also studies projects which are going well. Chapter 3 describes the various approaches which can be used.

c. Evaluating project impacts on the target population. Whereas all projects require relatively standardized monitoring procedures, managers have more flexibility in defining the scope and intensity of the impact evaluation. The importance of precise estimates of impacts as a factor in the

selection of future projects will usually determine whether the evaluation is designed to produce general, non-quanttive estimates of project impacts (described in Chapter 3 and Annex B), or whether carefully controlled experimental studies will be used to produce quantitative estimates of project impacts. For example, project impact on household income could be studied either through qualitative interviews and observation of a small number of subjects, or through a longitudinal impact study which would require interviews with a large sample of project participants and a control group at two or more points in time. The first option would be much cheaper and faster, but much less precise. The manager, not the researchers, must decide which approach will provide the information needed for future policy decisions. Chapter 4 describes the techniques for conducting each type of study. The importance is stressed of combining quantitative and qualitative evaluation techniques and of integrating impact evaluation with process monitoring so as to understand the way in which project implementation affects outputs and impacts.

d. Using cost-effectiveness analysis to compare alternative projects in terms of their costs for producing certain outputs. These techniques are described in Chapter 4 section F.

2. Organizing the monitoring and evaluation

Decisions have to be made about whether monitoring and evaluation will be conducted within the executing agencies or whether some or all of the studies will be subcontracted to other organizations. Where monitoring and evaluation are conducted internally, decisions have to be made on the location of the monitoring and/or evaluation units within the organizational structure. The decision must also be made as to how and when to use consultants. All of these issues are discussed in Chapter 5.

3. The scope and intensity of the studies

For large projects with many components, and possibly covering many different geographical areas, research priorities must be established. It would be excessively expensive and time-consuming to intensively study every component in every area and consequently management must decide which components or areas are the most important to study. As discussed previously,

it is also important to define whether precise quantitative estimates of impact are required, or whether descriptive studies will suffice.

4. Ensuring that the evaluation is "user oriented" and not "technique oriented"

Many monitoring and evaluation researchers have become experts in the use of particular research techniques, and some researchers are more concerned with the selection and refinement of techniques than with ensuring that the studies satisfy the information needs of the project manager. All research methods have their strengths and their weaknesses and a good evaluation design should always include a number of different techniques so as to obtain a wider understanding of the problem and to avoid the types of bias which innevitably occur when only one method is used.

When talking to evaluators, the project manager will discover that there is a controversy between the advocates of quantitative ("objective") methods, which seek to provide rigorous statistical estimates of project impacts; and the advocates of qualitative ("subjective") methods, which seek to understand the value systems of the affected populations and the meanings they attach to the project. These issues are discussed in Chapters 3 and 4. It is argued that both quantitative and qualitative methods have their strengths and weaknesses, and that the evaluation design should include both.

5. Defining resource requirements

Monitoring and Evaluation techniques vary widely in terms of their scope, complexity and cost. One of the functions of the manager it to select the types of studies and the organizational structure which best respond to the needs and resources of the project. Resources should be allocated to high priority issues and care must be taken to avoid investing resources in studies with limited operational utility. Chapter 5 provides guidelines for estimating the likely resource requirements for different types of monitoring and evaluation programs.

When addressing the issue of affordability of an evaluation it is important to keep in mind the concept of cost-effectiveness. In a complex project there may be a justification for investing large sums of money in order to conduct a rigorous evaluation of alternative investment strategies,

as this may permit very large cost savings in the design of future projects (see the example in Chapter 4 Section C.1 and table 4.1.

D. MANAGING THE MONITORING AND EVALUATION

1. The role of the project manager in monitoring and evaluation

Monitoring and evaluation are management tools, and in order for them to be operationally useful the project manager must be actively involved in all stages of the design and review of the studies and in the implementation of the recommendations. Although the technical aspects of the studies are the responsibility of the research team or consultants, the project manager must control the selection of studies and ensure the important operational questions are being addressed. When the responsibility for the design and execution of monitoring and evaluation is left to researchers, the studies may become too academic and their operational utility may be reduced.

2. Defining the main users of monitoring and evaluation

In order to ensure that the monitoring and evaluation reports will be operationally useful, it is first necessary to define the potential users of the studies and their information needs. There are usually four main groups of users:

a. Project managers: who use the information to control the implementation of the projects for which they are responsible; to ensure that objectives are being reached and that the benefits are reaching the target population; and to help in the design of future projects or the extension of the project to new areas.

b. Urban development planners: who use the information to evaluate the project impacts on the overall development of the city; to compare costs and performance with alternative investments and to plan new projects.

c. National finance and planning ministries: who use the information to speed the receipt of future disbursements by international lending agencies; to control internal cash-flow and to evaluate revenue generating impacts of the projects.

d. International lending agencies: who use the data as the basis for their regular supervision and planning activities and to assist in planning future disbursements.

In addition to these groups of primary users, the information is of utility to a wide range of researchers and academic institutions.

3. The outputs of monitoring and evaluation and their practical utility.

It is important to ensure that the monitoring and evaluation studies will produce reports which appear in a timely fashion, which are simple to read and which address the key operational issues. In order to ensure this, the studies and their outputs should be closely coordinated with the main stages of the project development cycle (see Table 1-1). The following are the main types of operationally useful studies and reports which can be produced at each stage of the project cycle.

a. Studies produced during the project planning and design stage

(i) The monitoring and evaluation system should be developed and appraised during the project design stage in the same way as other project components. A detailed research design, stating objectives, methods and organization should be prepared and agreed to at this stage.

(ii) The preliminary findings of evaluation studies can be very helpful in the planning of a new project. For example, participant observation or other in-depth studies can estimate the likely reactions of a community to a proposed project and can identify potential problems which may arise in certain communities.

(iii) Although they are potentially very useful, most evaluation reports are not produced until towards the end of a project, by which time the second project may already have been designed. In order to ensure that full use is made of the experience of the first project in the design of subsequent projects, special studies may need to be conducted.

b. Studies produced during the process of project implementation

(i) Regular monitoring reports are produced to assess the use of resources and the production of outputs and to identify key issues requiring more intensive study. The monitoring reports will often be published in the form of a Quarterly Progress Report (see Chapter 2). The purpose of this report, which is frequently required by the national

TABLE 1-1: EVALUATION OUTPUTS AND THE PROJECT CYCLE

Stage of Project Cycle	Evaluation Output	Comments
Regular project supervision	Quarterly Progress Report	This will complement and be prepared at the same time as the Quarterly Supervision Report sent to the World Bank. Whereas the Supervision Report is prepared for use in Washington, the Progress Report is designed for project management and for the project officer.
	Rapid feedback studies	More detailed analysis of issues identified in the quarterly progress report.
	Intensive studies	Occasionally these will be requested to assist with project supervision.
Mid-term project review	Mid-term project review	This will normally be a synthesis of existing studies although additional data collection may be required.
	Intensive or rapid feedback studies	These may be conducted to produce information required for the mid-term review.
Project completion and audit	Final report	This will complement the project completion report submitted to the Bank, or in some cases the two reports may be merged.
	Intensive or rapid feedback studies	May be conducted to produce information required for the final report.
Appraisal of second project	Mid-term project review or final report.	Depending on the timing one or both of these reports will provide inputs into the appraisal and design of new projects.

government or the lending agency, is to provide a summary of key indicators on the progress of each project component, together with an indication of key problems and issues requiring attention. The report should provide feedback on community attitudes and interactions with other organizations, as well as summarizing information collected within the agency itself.

(ii) Continuous panel studies (or periodic studies) may be conducted on certain projects to provide continuous feedback on implementation and to identify potential problems.

(iii) One-time studies are conducted whenever management requires help in the evaluation of problems or requires information to help in making an important decision. These may either be rapid studies designed to produced feedback within a few weeks or they may be more intensive studies when the problem to be studied in more complex.

c. Studies produced when the project reaches mid-term or is completed

(i) A mid-term evaluation review is frequently required by the central government or the lending agency. The report reviews progress, identifies key issues and makes recommendations about possible design changes which should be considered. This report is important as it is prepared when the project has been underway for a sufficient amount of time for there to be a solid basis on which it can be evaluated; whilst at the same time there is still sufficient time and resources for it to be possible to make significant corrections if they are required.

(ii) A project completion report is normally required at the completion of a project. This will review in detail the extent to which project goals have been achieved, the efficiency with which the project was organized and recommendations for the design of future projects.

4. The importance of regular reviews of monitoring and evaluation outputs

The design of a useful monitoring and evaluation system is an iterative process in which the quality of the study designs and the reports will gradually improve on the basis of experience. The quality of the first studies and reports will often be somewhat limited, as both researchers and managers are inexperienced in this field. In order to ensure that the quality of studies improves, regular review procedures should be instituted by which

all reports are discussed and evaluated and by which the key issues for future studies are defined.

The review can be organized by inviting the evaluation team to make a presentation on their report in one of the regular management meetings and by assigning sufficient time for the report to be discussed and for agreement to be reached on the actions to be taken. Reports should be circulated before the meetings and divisional managers should be responsible for obtaining comments from their staff.

5. POTENTIAL PROBLEMS IN THE DESIGN AND IMPLEMENTATION OF A MONITORING AND EVALUATION SYSTEM AND SOME POSSIBLE SOLUTIONS

This section is intended to alert the project manager to some of the most common problems and criticisms which are made about evaluations, and to propose approaches to their solution.

Common Criticisms and Potential Problems

* Evaluations can become very expensive:
  In many countries evaluation exercises have proved to be very expensive. In addition to expensive expatriate consultants, large numbers of expensive local staff and sophisticated computing systems have been used. The evaluation may also be competing for scarce professional staff who could have been used in other activities.

* Limited involvement of program management in the design of the evaluation:
  Many evaluations are designed by consultants or researchers with very limited input from management, either in the definition of the objectives or in the review of the system once it is operational. As a result the findings may be of limited operational utility.

* Too much emphasis on long-term studies which do not produce immediate results for management:
  Many evaluation manuals place emphasis on longitudinal impact studies designed to measure overall project impact. These long-term studies often become one of the major components of the evaluation and, as a result, too little attention is given to short and medium run studies.

* Division of the research between separate and uncoordinated units:
  Monitoring and Evaluation are often set up as separate units, with different staff reporting to different parts of the organization. Often there is very little coordination between the Monitoring and Evaluation teams.

* Inadequate methods for the dissemination of findings:
  A high proportion of the research is never read by the people for whom it was intended. Decision-makers are very busy and they do not have the time to read long reports.

* Very little feedback or control from the organization on the way in which the research develops:
  Researchers complain that they receive very little feedback or guidance from the organization, so they do not know whether their studies are useful or how they should be changed. Management, for its part, often complains about the large amount of time which must be devoted to the research. As a result, the evaluation program often develops in a vacuum.

It is important to emphasize that, in most cases, these problems have arisen not from lack of interest. Usually, great time and effort have been devoted by managers and researchers to trying to make the evaluation useful. Difficulties seem to arise not from disinterest, but from the way in which the evaluation activities are perceived and how they are organized.

Possible solutions

The above problems have no simple solutions as can be seen from the difficulties which efforts to improve evaluation performance have often encountered. A number of basic guidelines are proposed to ensure a more efficient use of an organization's research activities:

** The evaluation system must be designed and controlled by senior program management:
  This is perhaps the most important and the most difficult requirement. Evaluation is intended as a management tool which, to be useful to management must be designed and controlled by them. The system must be sufficiently simple for management to maintain control without having to expend large amounts of time.

** An integrated system must be developed for defining information needs and for generating the information:
Program management must be responsible for defining the information needs of the organization in a unified way. Once these needs have been determined, an integrated monitoring and evaluation system must be established for collecting and presenting the required information.

** Rigid distinctions between Monitoring and Evaluationmust be avoided:
The attempt to distinguish between monitoring and evaluation has often produced rigid segmenting of research efforts. In contrast, this Handbook recommends an integrated monitoring and evaluation system which can satisfy all management information needs. Although there are several distinct users, each with different information needs, there is considerable complementarity between the research methods used to generate the information required by each group.

** The research system must start by identifying information users and their requirements:
The key to a successful evaluation system is to begin with a clear definition of the users of the information and their information needs. The system can then be designed to provide this information.

** Development of a successful evaluation system is an iterative process:
The system must be constantly reviewed as it develops. There should be built-in procedures for making modifications as they are required. Information or types of study which are not proving useful must be identified and rapidly eliminated.

** The system must be kept sufficiently simple not to exceed the ability of the organization to absorb the information:
The temptation must be avoided of setting up a comprehensive system which can provide every type of information which might ever be asked for. A system like this would be excessively expensive and time consuming, produce much unwanted information, and not have the flexibility to respond to management's changing needs. The aim is to reduce the amount of information being collected to a minimum. It is much better to start simple and to expand later if the need arises.

CHAPTER 2: PERFORMANCE MONITORING

This chapter describes the design and implementation of a system to provide periodic feedback on the progress of a project, the extent to which inputs are being used in accordance with the approved budget and timetable and whether the intended outputs are being produced in a timely and cost-effective manner. Two systems are described, a basic monitoring system which can be simply and economically applied to any project; and a more complex system, based on network analysis, which is more appropriate for larger and more complex projects.

A. PERFORMANCE MONITORING

Successful accomplishment of a project's objectives requires a plan that identifies specific tasks to be done, the costs, the sequence in which they are to be done, the linkages of these activities and effective management of this plan.

The process of managing such a plan toward project completion embraces the concept of monitoring. As explained in Chapter 1, monitoring is an internal project activity designed to provide periodic feedback on the progress of a project, the problems it is facing, and the efficiency with which it is being implemented. Monitoring can be broken down into closely related but distinct sets of activities: performance monitoring which assesses the extent to which inputs are being used in accordance with the approved budget and timetable and whether the intended outputs are being produced in a timely and cost-effective manner; and process monitoring which assesses the efficiency of the project implementation process. This chapter will discuss performance monitoring and the following chapter will discuss process monitoring.

Monitoring is an essential management tool. Projects which do not have an effective monitoring system are more likely to suffer delays and cost overruns; to exclude or under-represent certain sectors of the target population; to have problems of quality control; or to take longer to detect antagonisms among the implementing agencies or between the agencies and project beneficiaries.

Performance monitoring provides useful information to management to assess progress of implementation and take timely decisions for efficient and effective project completion. Complex data collection systems are usually not required to achieve this. Neither does formal monitoring necessarily have to result in voluminous reporting. The overall goal of a monitoring system is to make the data collection as relevant as possible, and to ensure that the means exist for fast collection and summary of data so that it can be presented as useful information to the decision-makers.

A monitoring system established early in the project cycle can serve as an effective project management tool by providing continuous, relevant, action-oriented information in order to make timely modifications as

necessary. A typical urban project lasts around seven years. During this period there are likely to be significant changes in the economic, institutional and political environment which are likely to affect the project's performance.

Developments take place that were not foreseen at the time of project design: cost overruns occur, demand could be low, counterpart funding becomes scarce which could affect disbursements, time delays in construction occur due to inclement weather conditions, consulting engineers could perform unsatisfactorily, all of which require adjustments in the scope and composition of the project to suit changing conditions. It is indeed unrealistic to expect a project to be completed exactly as planned. Effective monitoring of a project's development based on a thorough understanding of its inputs, sequence of activities and their interrelationships would enable management to detect early any unexpected situation so that corrective or adaptive measures could be taken.

All urban projects undergo some form of monitoring whether by one line manager or through a small unit designed explicitly for this purpose. To the extent that a project is managed it is monitored. Monitoring is an aid to project management. All implementing agencies submit quarterly or monthly monitoring reports on a regular basis, in some cases to their own government, in others to an international lending agency. Sometimes financial status reports on project accounts are submitted separately from reports on physical progress. These two kinds of reports are usually prepared by different units of the implementing agency or by different agencies altogether with little coordination in compiling these documents, with the effect that their usefulness as project management or project control tools is minimized. The system outlined below attempts to address this weakness. It integrates project budgets with actual project timing, costs and performance of a project's components.

B. DESIGNING THE PERFORMANCE MONITORING SYSTEM

The exercise of thinking through the design of a monitoring system for a particular project helps to internalize into the management system the project's objectives, the underlying logic - how these objectives will be accomplished, performance indicators for measuring progress, and available

data sources from which the measures can be drawn, as well as defining how the results should be disseminated to the coordinating agencies. It is thus preferable that the design be done by the project management unit or the project manager so that management information needs are adequately reflected.

Projects vary greatly in terms of their complexity and in terms of the resources available to monitor them. Consequently two alternative performance monitoring systems are presented. Annex G describes a relatively simple system which can be implemented in any project, and which requires very little additional financial and human resources. This basic performance monitoring system can provide management with all of the essential information required for project control. Annexes H and I present a somewhat more complex network based monitoring system which provides a much more rigorous control of physical and financial implementation, and is recommended for larger and more complex projects. Reference is made to both systems throughout this chapter. Annex G provides a set of pro forma monitoring tables which can be used as guidelines and adapted to suit the specific requirements of each project.

A project's schedule and budget will normally predict when certain stages of physical progress should be met, when expenditures should be disbursed and when "software" project elements should begin. The chart given in Table 2 of Annex G illustrates how this approach could be applied in planning a hypothetical project to renovate the existing housing stock in a large city. The chart, by showing the logical sequence of activities, provides a simple way to identify those components which are falling behind schedule, and where some action must be taken. Adhering to the suggested sequence is essential to ensuring that a project's objectives are met, or if deviations do occur that they are identified early and that corrective or adaptive measures are taken.

Steps which would be taken in developing a monitoring system are listed below:

First Step

Subdivide the project appropriately into its major physical, financial and social features. For example:

(a) Infrastructure
(b) Equipment and vehicles
(c) Community facilities--markets, workshop, dispensaries, etc.
(d) Project management--salaries, office furniture and equipment
(e) Technical assistance--consultants, contractors, studies, etc.
(f) Construction loan program

Second Step

Using network scheduling (CPM) (Annex H, Charts 1 through 4) or a simple Gant bar (Annex G, Table 1) set up the projected sequence of planned operations in detail indicating the logical relationships between the various activities. Set up a project financial schedule, including also parallel disbursements for principal physical features, showing the expected payment schedules for each major item (Annex G, Tables 2 and 6 and Annex I, Charts 1 through 3).

Third Step

Determine required staffing for the monitoring unit according to the project composition, scope and the required tasks to be performed. Usually an accountant (or assistant) and an engineer (or assistant) will be required to monitor costs and the quality of physical progress. Sufficient authority and facilities must be given to the project monitoring team early in the project cycle to allow supervision of project records to be maintained and the collection and correlation of information to be provided regularly for the effective monitoring of physical progress and project funds.

Fourth Step

Design bar charts for physical progress and forms for budget control. As mentioned above (Second Step) a Gant Chart (3-bar type) can be used and is usually adequate for most urban projects because of the linear sequence of most activities. For projects which involve many simultaneous activities, the network-based system described in Annex H may be more appropriate.

On the Gant Chart show (a) original estimated starting and completion dates for each principal item of physical work; (b) on second bar

actual starting date for the item and the estimated percentage of completion as of the date of the report; and (c) on the third bar re-estimates of completion dates if necessary.

Budget information forms should include at least the following columns:

Column 1 Official estimates with a total equal to the budget estimates listed in Step Two above

Column 2 Authorized changes in each item

Column 3 Sum of the above to arrive at official control budget

Column 4 Contract orders placed to date

Column 5 Accountable expenditures to date

Column 6 A new estimate should be made at the end of each report period, showing amounts required to complete each item - not made merely by subtraction of the pertinent columns, but actually taking the quantities of work remaining to be done, re-estimating them and entering this expected cost. (Usually this is done by the field engineering contractor on a monthly basis). This then constitutes a new official estimate of expenditures still to be made.

Column 7 Estimated percentage of completion of project items

The above constitute the relevant data which must be submitted to management; they are comprehensive and correlate key information (in a concise way) on physical progress and project expenditure. The pertinent forms are provided in Annex G, Tables 1 to 7.

Performance monitoring reports prepared in this way and submitted according to an agreed schedule provide management with an effective tool for appraising progress and determining which, if any, steps need to be taken to make corrections and changes. Since a major objective of monitoring is to guide project performance, timeliness is of the essence. To some extent there is a trade-off between speed and completeness of the information. For monitoring, as opposed to evaluation, purposes, less precise information which is available in time to help with decision making, is far preferable to more complete information which arrives after the decision has been made.

C. MONITORING SOFTWARE COMPONENTS

Project monitoring systems were originally designed to control the implementation of "hardware" project elements such as construction of roads and houses. However, most urban projects include one or more "software" components, either as a final goal or as a necessary step in the implementation process. "Software" components include: organization of communities to assist the planning and implementation of a project; use of social promotors to ensure that the community is informed about and in agreement with the goals and methods of a project and health education programs. All of these components are difficult to monitor using conventional methods because there are no easily quantifiable indicators of performance.

A frequent, and usually unsatisfactory approach, is to try and reduce all of the social activities to numbers, even though the numbers do not capture the essence of the activities involved. For example, social workers may be required to report on the number of houses visited, the number of meetings held, the number of people who attended the meetings, etc. If these numbers are used as a basis for evaluating the performance of the social workers, then staff will have an incentive to either falsify the numbers (not an uncommon occurrence) or to substitute quantity for quality in their work. Instead of spending several hours with a family which is in need of help, the social worker will have an incentive to make rapid calls on a large number of families. Several approaches are possible to obtain more meaningful monitoring information on these software components:

1. Rapid sample surveys of the actual and intended beneficiaries to determine their attitudes and knowledge about the project. Table 8 in Annex G provides an example of how such a survey can be conducted and the results presented. Beneficiary awareness is very important as many projects suffer lengthy delays due to inadequate communication channels. It can be very useful to management to have a simple table showing the proportion of families who have received, and understood, the basic information about the project. Feedback on opinions is equally important. Again projects should have adequate feedback mechanisms to keep management informed of beneficiaries' points of view as the project progresses.

2. Service Delivery Assessment (Hendricks, 1981) can be used to observe the effectiveness with which services are actually delivered to the target population. The method, which has been pioneered by the Department of Health and Human Services in the US, involves visits by agency staff (not consultants) to project sites to meet with clients, front-line service providers, local administrators, program officials at the lcoal and state level, and other people who are knowledgeable about how the project actually operates. In this way, rapid, qualitative feedback can be obtained on the "soft" services which are actually delivered.

D. METHODS FOR COLLECTING AND ANALYZING DATA FOR PERFORMANCE MONITORING

In order for the monitoring system to provide timely and economical feedback it must use data which is either already available or which can easily be obtained. Some of the main sources of data include:

1. Financial information (expenditures, disbursements received, cost variation, etc) which already exists within the organization. *In most cases the preparation of these special tables does not involve much extra work, as long as the table format is agreed to at the start of the project. However, it is usually extremely difficult to restructure the formats once the project is underway.

2. Credit information on loan disbursements and repayments and information on payment for service charges can readily be obtained from project records or from the records of a cooperative or similar organization supported by the project. A common problem is a long delay before the information is available.

3. Information on the physical implementation of project components. This information is usually readily available from the division responsible for construction or supervision of construction.

4. Observation guide to monitor house consolidation by residents and the construction and maintenance of community infrastructure. Often this information will not be available from project records, as the families

themselves are responsible for completing their houses. A simple observation guide can be developed which includes questions on (a) whether building materials are on the plot, (b) whether construction has started, (c) how many rooms have been completed, (d) what kinds of materials are used, etc. The information can usually be obtained by direct observation without having to interview families, which means that it can be collected very rapidly.

5. Short survey of community residents to obtain their opinions on different aspects of the project and to estimate the proportion of households who have certain types of information about the project (see Section C).

6. Informal conversations with project staff, beneficiaries and others to obtain opinions and suggestions about the project and the way it is being implemented.

Depending on the complexities of the project, and the resources of the organization, the data can be processed in one of the following ways:

a. Hand-tabulation. This method is quite effective and rapid when the numbers and types of tables to be produced are relatively simple.

b. Automatic generation of tables through the project's main computer system. This is in many ways the ideal system, however, it is subject to a number of potential dangers. First, the monitoring reports may have to use the same tables, categories and formats as are being used by the organization for other purposes, which may mean that certain types of desired information will not be available. Second, many management information systems are subject to considerable delays, so that if the monitoring reports rely mainly on this source, they too will be subject to delays.

c. Use of a micro-computer. This alternative permits much greater flexibility and speed.

E. REPORTS

The monitoring system will usually produce all, or most, of the following reports. The amount of detail, and the frequency of the reports will be largely determined by the level of resources assigned to the monitoring exercise.

- Quarterly or Monthly Progress Report which can be used as a regular project control tool by managers.
- Interim Report which is to be prepared at critical points in the project development. This report would provide in-depth technical analysis of major issues and problems which could result in project restructuring, scaling back, or early completion.
- Final Report which presents an ex post analysis of project performance shortly after its completion comparing the costs, outputs and expected benefits with appraisal estimates and thereby contributing to an assessment of the effectiveness of the project and the executing agency.

The data base provided by the accumulated quarterly reports should contain most of the information needed for preparation of the interim and final reports.

F. POTENTIAL PROBLEMS AND POSSIBLE SOLUTIONS IN THE DESIGN AND IMPLEMENTATION OF A MONITORING SYSTEM

The following are some of the typical problems which arise during the design and implementation of a Monitoring System. Possible solutions are proposed for each problem.

Problem No. 1: There are considerable delays in setting up the system so that the project has been underway for some time before monitoring reports begin to appear.

This is a common problem and is due to a number of causes. First, monitoring often has a low priority so that no thought is given to it until major operational problems appear. Second, a related problem is that it is much more difficult to obtain budget and authorization for hiring additional staff once the project is underway as resources have already been committed for other purposes. Third, management will often reassign monitoring staff to other temporary duties, again delaying the production of monitoring reports.

The following guidelines can help overcome these problems:

1. Decisions on organization, budget and staff for monitoring should be made at the time of project appraisal. At this time the budgetary line item for monitoring should be approved as well as the terms of reference and grades of all staff and consultants who must be hired or reassigned.
2. It is the duty of management to ensure that monitoring is given a high priority.
3. Management must enforce a timetable for the production of reports. They must ensure that the reports are sufficiently straightforward for the deadline to be feasible.

Problem No 2: Monitoring reports are too long and tend to be published too late:

This problem is due in part to management's lack of involvement in the design of the monitoring system. The following approaches can be used to address this problem:

1. Management must thoroughly review what types of information are really needed, making an effort to reject unnecessary material.
2. Management must insist on reports being produced on time.

Problem No. 3: Monitoring may only cover the "hardware" components which are easy to quantify and may ignore important "software" components

This is a common problem, particularly when project managers have previously worked in civil construction of similar programs. Section C proposes methods which can be used for monitoring the software components. An important preliminary step is to increase the organization's awareness of the importance of the "software" components. This can be done through a combinaton of (a) training and seminars, (b) qualitative studies which describe the importance and impacts of components such as communication, community particiaption in planning and decision making, (c) rapid studies of beneficiary opinions and information about the project.

CHAPTER 3: PROCESS MONITORING - MONITORING THE PROJECT DELIVERY SYSTEM

This chapter reviews key issues and designs for monitoring the efficiency and effectiveness of the project implementation process. Among the key issues discussed are the trade-offs between different indicators of efficiency, and some of the organizational factors which have prevented many programs from effectively monitoring the implementation process. Process monitoring can either be conducted to provide regular information on the progress of implementation, or to provide rapid feedback when problems arise. The main data collection methods are presented, with the recommendation that a multi-method approach should always be used in which quantititative and qualitative methods are combined. Three types of studies are explained: continuous observation throughout a project, periodic studies, and studies conducted at only one point in time. A number of techniques are then presented for monitoring the efficiency of the implementation process. Finally recommendations are presented on how to define the appropriate strategy for process monitoring.

A. KEY ISSUES

Although the final outcomes of a project are obviously affected by the initial design and the availability of the necessary resources, in practice an equally important determinant is the way in which the project was actually implemented. Many projects are developing new approaches and often they are much less successful than expected simply because the delivery system did not work as planned (see Box 3-1). In some cases there are difficulties in hiring or retaining well qualified staff, while in others the administrative

> BOX 3-1 THE EFFECTS OF THE DELIVERY SYSTEM ON PROJECT OUTCOMES
>
> The following are typical examples of ways in which the delivery system can affect project outcomes:
>
> ** Project planners in Zambia had assumed that following land nationalization, acquisition of land for squatter upgrading would not present major problems. In fact lack of surveys, inexperience in the implementation of the new legislation and problems of interagency coordination served to delay the project by almost 3 years. This caused substantial increases in project costs and the consequent elimination of a number of schools and community centers from the project.
>
> ** In order to use house construction as a means to develop community organizations, participants in a sites and services project in El Salvador were required to participate in mutual help construction groups every weekend. Although this was successful in developing grass-roots organizations, it may have discouraged people such as small shop-keepers and self-employed artesans from participating as the opportunity cost of their labor during the weekend was much higher than for wage labororers.

procedures were more difficult to implement than expected. Communication between implementing agencies and intended beneficiaries has proved to be particularly problematic (See Box 3-2). In some cases the lack of adequate communication linkages prevented a project from starting; in other cases leaders misrepresented the interests and willingness to pay of certain sectors of the community, or only represented the views of the wealthier beneficiaries. Due to all of these factors it is clear that the evaluation of the project delivery system is an essential component to understanding why a project was or was not able to achieve its desired objectives.

Despite the importance of understanding the way in which the delivery system actually works, many monitoring and evaluation programs do not include process monitoring. One of the reasons for this is that monitoring and evaluation are frequently divided between a small administrative unit which uses secondary data to prepare performance monitoring; and a separate unit which only conducts impact evaluations. Within this scenario there is frequently no institutional capacity to study the delivery system.

BOX 3-2 SOME EFFECTS OF INADEQUATE COMMUNICATIONS ON PROJECT IMPLEMENTATION

The following are examples of how poor communications can affect the process of project implementation:

** It was several years before the executing agency in Guayaquil, Ecuador realized that community leaders, who were from one of the opposition parties, were deliberately misinforming the community about the services being offered. Due to this the community showed no interest in the project and it was never started.

** In a squatter relocation project in Recife, Brazil the implementing agency believed that the community had been informed by their leaders as to the nature and costs of the project and that most families were in agreement. A rapid study revealed that most families had not in fact been informed and due to this were becoming hostile to the project.

** In Usulutan, El Salvador it was found that many low-income and illiterate households were not applying for houses as they had not been reached by the mass media communication techniques which had relied on publicity through the cinema and radio or through written communications.

Table 3-1 illustrates some of the key issues which arise when monitoring the efficiency of implementation of a typical shelter project. In addition to monitoring each indicator individually (for example the speed and cost of the process of selecting beneficiaries, it is also important to assess the trade-offs between the different indicators. For example, household income is frequently measured in order to use "capacity to pay" as a selection criterion for a sites and services project. It is often possible to speed up the selection process by only taking into account formal labor market earnings which are easy to verify. However, this approach will tend to exclude people who work in the informal sector and whose earnings are harder to verify. Thus

there may be a trade-off between speed (and perhaps costs) and equity. Table 3-2 indicates some of the key issues in monitoring the efficiency of implementation of health, employment, transportation and municipal development projects.

The issue of trade-offs between different efficiency indicators is particularly important when assessing the performance of agencies who traditionally work with higher income groups in the formal sector (for example financial and housing institutions) and who are required to adapt their procedures to the characteristics of low-income households who live and work in the informal sector.

The method and effectiveness of the delivery system can also have important consequences for final outcomes. For example, the poorest households often live in the least accessible parts of the community and may be the last to receive water, roads, or sewage connections. Consequently delays in implementation, or reductions in scale, may have the severest impact on the poor or may in fact mean that the projects never reach them. Implementation delays, particularly during periods of high inflation, can also increase costs, so that poorer households may no longer be able to afford the project.

The choice of delivery system can also affect the distribution of benefits. Box 3-1 showed how the requirement to participate in mutual help construction may discourage certain groups from applying to the project. Similarly requirements about formal proof of earnings or certain types of financial documentation may eliminate many types of small businesses from access to credit. Legal requirements relating to property ownership may also make it difficult for women to acquire houses.

The success of most types of projects is very much affected by the economic and political environment within which they operate. For example, low-cost housing projects tend to be politically controversial due to issues relating to land invasion and property titles, or because politicians are concerned that the standard of housing is too low or that the projects will become slums. The issues of subsidies and cost recovery often create conflicts between implementing agencies who wish to use the threat of eviction and service cut-offs to improve cost recovery, and local politicians who seek to protect the rights of their constituents and who oppose these measures. The

Table 3-1 ISSUES IN EVALUATING THE EFFICIENCY OF PROJECT DESIGN AND IMPLEMENTATION

Criteria for Evaluation of Project Efficiency

Project Element	Speed of Implementation	Cost	Quality	Accessibility	Replicability
Planning and Design	Self-help benefits offset if progress is slow; this raises costs and delays benefits.	i. High design standards raise costs. ii. Use of local vs. imported materials is an important cost issue.	The quality of construction will affect capital value, rental value and owner satisfaction.	i. High standards, leading to high costs can result in poorer families being excluded.	i. Complex designs are more costly to replicate. ii. Projects materials and equipment requirements may limit large scale applications.
Selection Processes	Some selection procedures are very time consuming.	Selection procedures are not costless; costs, furthermore can vary considerably.		i. Richer families may falsify income to be selected, forcing out poorer families. ii. Selection processes may be biased against certain groups.	Replication may require use of computer selection to speed up processing of of larger numbers of applicants. Danger of eliminating certain groups through mechanical selection.
Construction Methods	Mutual help construction is often much slower than self-help, which may	Major cost differences between construction methods; importance of timing.	How much supervision is required to ensure adequate quality.	Are some groups excluded by the use of certain construction methods?	Can the method be repeated on a larger scale? (This is closely related to supervision requirements).
Materials and Credit	Substantial delays may occur in the distribution of materials.	i. Materials from project stores are not always cheaper. ii. Material stores usually offer a smaller selection.	Quality of material from project stores are always better.	i. Are loans large enough? ii. Would project be more accessible if people could buy materials outside.	Replicability is again essentially a question of complexity, cost and administrative capacity.
Plot Occupancy	Any implementation problem will slow occupation.	Slowness of occupation means higher costs: i. Rising real costs during inflation. ii. Expenses of two residences (S/S).	Pace of occupation may be affected by perceptions of quality.	i. Management may offer plots to higher income groups to accelerate occupation.	The slower the process, the less can be done in any period.

Table 3-1 (cont'd)

Project Element	Speed of Implementation	Cost	Quality	Accessibility	Replicability
Maintenance	i. Problem of ensuring speedy maintenance through local government agencies. ii. Should project have own maintenance arrangements?	Are benefits of lower charges lost through higher maintenance costs?	Is quality related to the level of community participation?	The maintenance policies limit access of certain groups to services.	Is possibility of replication threatened by inability to maintain project services?
Cost Recovery	Project delays, especially in plot occupation, can make cost recovery more difficult.	Cost recovery problems substantially increases project costs.		Is low cost-recovery a disguised subsidy which may make project more accessible to lower income groups.	i. This will have a crucial direct impact on replicability. ii. In secondary cities, where finances are weakest and subsidies are not possible, the revenue generating potential of the project may be a key issue.
Community Participation	This may slow project implementation through long discussions or speeding up implementation by more community support.	This can both reduce costs by having them absorbed by the community, or increase through need set up an additional administrative structure.	This affects maintenance and cost recovery.	May be important in terms of cost reduction and ensuring the services are those desired by the community.	Community organization is both expensive and complex to organize on large scale and may discourage government from replicating.

Table 3-2 Some Key Issues in Process Evaluation for Non-Shelter Components

Income and Employment Generation

1. Do procedures for ensuring speed and economy in the selection and operating procedures, negatively affect the access of certain groups (such as the smallest artesans, female entrepreneurs) to credit and other services?
2. How well are credit and training institutions able to adapt to operating in the informal sector of the economy?
3. Are the training programs adapted to the needs of small, informal sector businesses?
4. How well are the businesses able to compete for raw materials and markets?
5. What type of intermediary organizations are developed and how effective are they as communication links and for representing the views of beneficiaries?

Health

1. How do traditional cultural beliefs and practices affect the success of child-care, health and sanitation programs?
2. How effective are the community promotion and out-reach procedures in ensuring the project is understood by, and accessible to all sectors of the target population?
3. Are there any cultural or economic groups who do not have access to the program?
4. How cost-effective are the programs in comparison with conventional health and child-care?

Transport

1. Cost-effectiveness comparison of small-scale transport programs with conventional public transport or large-scale private companies.
2. How effective are the projects in reaching all sectors of the target community, and what are the main barriers to a wider coverage?
3. Effectiveness of community involvement in decisions on the transport routes within the communities and on the policies for relocation and compensation of households who must be moved.
4. Effectiveness of community participation in construction and maintenance.

Municipal Development

1. Effectiveness of communication and coordination linkages between the implementing agencies and intended beneficiaries.
2. Effectiveness of communication linkages between implementing agencies.
3. Impacts of the economic and political environment on the effectiveness of project implementation.
4. Effectiveness of internal monitoring and management information systems.
5. Satisfaction of intended beneficiaries with interagency coordination.
6. Effectiveness of interagency meetings and problem solving mechanisms.
7. Indicators of internal organizational efficiency and assessment of how they are affected by the project.
8. Project impact on financial administration and control.
9. Implications of the organizational structures which are set up for project implementation on the long-term institutional development at the municipal level.

interplay between political and administrative pressures is a crucial factor in monitoring project peformance.

Many impact evaluations are designed only to measure the extent to which objectives are achieved, and consequently tend to overlook unexpected outcomes. It is essential to understand the causes and consequences of these unanticipated outcomes and to assess their implications for the design of future projects. For example: many community organizations become involved in projects which were not considered in the original project design; low-income shelter projects may either stimulate the interest of other agencies in similar low-cost approaches, or may create hostility and negative reactions; small business credit programs may result in credit being withdrawn from other less glamorous projects so as to enjoy the publicity generated by an international project.

B. WHEN AND WHY TO MONITOR THE IMPLEMENTATION PROCESS

Process monitoring can be used to provide regular information on the efficiency and effectiveness of the delivery system, or as a means to provide feedback to management when problems arise. The studies can be conducted on a continuous basis throughout a project's development, at several points in time or at one particular moment (see Table 3-3).

An example of a continuous study is the selection of a panel of households who are observed, through questionnaires and direct observation, throughout the process of house construction and consolidation. (FSDVM,1977). Information can be obtained on the sequence of construction (do they build extra bedrooms before upgrading the quality of the existing living areas), about sources of building materials (where were they bought and were they new or used), types of labor (hired or provided by the family or neighbors) and sources of finance. This permits a much deeper understanding of the process of house construction than could be obtained by simply asking questions at the end of the project. Continous observation can also be used to study project impact on income and expenditure patterns (RAD,1978) or on the way in which a small business develops.

Periodic studies can be conducted to observe a project during its main stages. For example, households can be observed at the point when a new sanitary system is being installed, after perhaps one year and again after

Table 3-3: Timing and uses of process evaluation (monitoring the project delivery system)

Type of Study	STAGES OF THE PROJECT IMPLEMENTATION: Planning	Implmenetation	Completion	Examples
PERIODIC	☐	☐	☐	1. Periodic visits to panel of families moving to new project to observe process of construction, sources of financing, etc. 2. Selection of panel of households who keep diaries on all sources of income and types of expenditure during an upgrading project. 1. Interviews with households at the start, during implementation and at some point after completion of a maternal and child health care program. Information collected on health practices, illnesses, health expenditures, attitude to the project, etc. 2. Participant observer lives in a fishing community for several weeks or months at different stages in the development of a fishing cooperative.
ONE TIME At start of project	☐			1. Rapid survey to determine which sectors of the target population are informed about, and interested in, a cooperative. 2. Participant observation to understand community organization and likely response to a shelter project.
During implementation		☐		1. Rapid study to determine why an artesan credit and training program has started more slowly than planned. 2. Intensive study, lasting several months, in which participant observation, qualitative techniques and short surveys are combined to understand why conflicts have arisen around an artesan cooperative program.
At end of project			☐	1. Rapid survey to assess effectiveness of project in reaching target group and to receive feedback on how to improve design of future project. 2. Intensive study, combining quantitative and qualitative techniques, to assess why an upgrading project was less successful than expected in developing grass-roots organizations.

several years. In this way it is possible to assess whether the services continue to be used and maintained when the implementing agency is no longer closely supervising the project. Maintenance is a major concern with many types of projects (roads, community facilities etc) and consequently it is important to be able to assess what happens after a number of years. As many evaluations are funded as part of the project loan, they often end when the project is administratively completed which usually occurs when the physical works are completed or the final loan disbursement has been made. Consequently it is often not possible to assess the longer term impacts or the efficiency of maintenance and cost recovery.

One-time studies can either be conducted as a form of quality control in which projects are selected at random, or to provide feedback when a problem has arisen. The one-time study can be used at the start of a project to understand the community organization and likely responses to the project. This can help select project sites or types of small businesses or provide information on areas or types of enterprises which have already been selected. The studies may also be conducted during the implementation of the project, for example when a certain number of business credits have been approved. Finally the studies can be used at the end of a project to provide an overview of the implementation process.

In practice there is often an operational distinction between rapid studies which are intended to provide feedback as quickly as possible; and more intensive studies which are used when the situation is more difficult to understand. A rapid study may be used to evaluate the management style and efficiency of a fishing cooperative (Box 3-3); to review the progress of a small business credit program; or to assess the efficiency with which a community construction project is being implemented. On the other hand more intensive studies, lasting several months may be used to understand the role of a local junta and its relationships with the community and the implementing agencies; to evaluate the overall development and problems of a cooperative program (see Section C.3 of Chapter 4); or to understand why a community was not willing to accept a project (Chapter 1 Box 1-1).

C. DATA COLLECTION METHODS

1. The importance of a multi-method approach

Many researchers have a preference for a particular research method which they try to apply in all their studies with the result that many monitoring or evaluation studies are based upon a single method. All research methods have their strengths and weaknesses and a study which relies on only one will give a limited and biased view (see Annex B Sections 1 and 2). For example, when a structured questionnaire is used to estimate income and employment, it is difficult to obtain valid estimates of income from people

BOX 3-3 EXAMPLE OF A RAPID PROCESS MONITORING STUDY

Reviewing the progress of a fishing cooperative in Natal, Brazil

The evaluation unit was requested to conduct a rapid monitoring study to identify some of the reasons why the fishing cooperative was progressing much more slowly than expected. Interviews were conducted with 10 fishermen, including both members and non-members of the cooperative, with project staff and with fishing experts. Fishing communities were visited and joint meetings were held with fishermen and project staff. The main findings were as follows:

a. At the time of the study the cooperative was so small and new that it was not able to offer many services and hence was not able to attract new members.

b. It was difficult to compete with the middlemen who could offer short-term credit (albeit at very high rates) and could provide supplies. Fishermen expected the cooperative to be able to protect them from the middlemen and to help them get out of debt.

c. Most fishermen had little awareness of the functions of a cooperative.

d. The management of the cooperative was suffering from poor book-keeping and administrative procedures, which was further complicated by a duplication of functions between the cooperative and the government fishing development program.

e. Many of the fishermen mainly fished for lobster, and as the cooperative did not purchase lobster they had little incentive to join.

Partly as a result of this study, the management systems were reorganized and their performance has improved.

Source: Nucleo de Acompanhamento e Avaliacao "Desenvolvimento da Pesca Artesanal". Nov 1983. Natal. Brazil

who work in the informal sector as their income varies constantly. They may also not wish to reveal their true income, particularly when a project has income eligibility criteria. Structured questionnaires are also widely used to obtain information on attitudes and behaviour, and again there is extensive evidence of biases in the information. For example, in El Salvador, many people did not wish to admit that they participated in community development or similar organizations for fear of possible reprisals. Similarly many people do not like to express negative views about organizations or people and so tend to respond very cautiously to survey questions.

Intensive ethnographic techniques such as participant observation and the preparation of in-depth case studies on individuals or families, have the advantage of providing much more insight and reliable information on particular individuals. However, these techniques are subject to a number of problems. First, the number of cases tends to be very limited so it is difficult to generalize. Second, it is difficult to know whether there is any bias in the selection of cases. Frequently certain types of people are willing, or even anxious to be interviewed, whereas others are very reluctant. It is likely therefore that the cases selected will not be representative. Third there is a problem of evaluating the extent to which the results reflect a bias on the part of the researcher. As the study is based on personal relationships between the observer and the community it is difficult to evaluate the results in the same objective way as can be done with quantitative methods.

Many studies rely heavily on secondary data such as records of community organizations, credit information from cooperatives, reports of community organizers, records of the construction department etc. Most of these sources have an incomplete coverage of the community or include certain biases.

For all of the above reasons it is important to ensure that a number of independent research techniques are combined so as to provide consistency checks and independent interpretations of the data. This approach is called <u>triangulation</u> (analogy from topography) in the sense that two or more indicators are used to get a more accurate estimate of the true value. A simple example of triangulation is the use of in-depth case studies with a small number of households to check the information obtained in a questionnaire. Another example is the use of direct observation (for example

of the number of "to rent" signs in a street) to check on questionnaire information about sub-letting. The use of a multi-method approach in which at least two different methods are used to provide a consistency check, is strongly recommended as a standard operating procedure in all the evaluations. Box 3-4 illustrates how verbatim reporting of statements in a meeting can provide a deeper understanding of the results of a sample survey.

Unfortunately there are a number of reasons why a multi-method approach is not more frequently used. It is, of course, more expensive and time consuming to use two (or three) methods than to use only one. Also, many researchers have a preference for a particular method and are reluctant to use others with which they are less familar. The most important reason, unfortunately, is that most researchers have not been willing to criticize their research methods and to admit the need to complement them with other techniques.

BOX 3-4 USING QUALITATIVE DATA TO COMPLEMENT SURVEY FINDINGS

Survey of a fishing cooperative in Natal, Brazil

A rapid evaluation of a fishing cooperative, in addition to presenting statistical analysis, also included the following verbatim comments of fishermen to help the reader understand the meaning of the statistical findings.

"The majority of the fishermen here don't own their boats, but are hired by a boat owner. He gives us everything we need, and when we return everything is for him."

"When the cooperative arrived they said that they were going to help us by providing materials and buying our fish.. but this hasn't happened. We go on suffering and have to buy smaller quantities of everything elsewhere at higher prices and we still have to sell our fish at lower prices."

"The cooperative has to offer us something more attractive if we are going to believe in it".

"This cooperative decides things without consulting with us fishermen. They even close the doors of the room where the fish are washed and we are not allowed to enter".

Source: Previously cited study from Natal (Nov 1983).

2. Modelling the process of project implementation

A helpful way to begin the design of an evaluation is to prepare a simple diagram illustrating how the project is expected to work (see Annex A) By illustrating the inputs, the main processes to be used, the expected outputs and impacts, it is easier to identify the points where evaluation studies are likely to be needed. It also becomes easier to understand the ways in which project implementation might be affected by the characteristics of participants or by the external environment within which the project operates.

3. Quantitative surveys (Annex B Section 7)

One of the most common methods for obtaining information on how a project is operating is to design a questionnaire and to apply it to a sample of project participants (and possibly to a control group not affected by the project). The questionnaire may contain the following types of information (among many others):

a. Information on the socio-economic characteristics of participants.

b. Information on how they participate in the project (frequency of attending meetings, amount of loans received, type of participation etc)

c. Knowledge about the project and its objectives.

d. Opinions on the project, its organization, the people and organizations involved etc

e. Changes which the project has produced (in income, health, access to services, operation of the business etc).

The surveys can be applied to different groups or at different points in time so as to compare groups or to measure changes over time.

4. Direct observation (Annex B Section 5)

Many aspects of a project can be directly observed without the need to ask questions. For example:

a. Progress of house construction or upgrading.

b. Numbers of people participating in community work groups and the way in which the work is organized.

c. Attendance at meetings, the decision making process and the level of group participation.

d. Indicators of changing economic conditions of the community (for example, quality of housing, number of houses with cars or bicycles, products on sale in community stores, quality of clothing etc)

Annex B discusses some of the problems which can arise in the interpretation of observational data.

5. Secondary data

Most projects produce large amounts of written and statistical documentation which can be of great assistance in the evaluation. Some common examples include:

a. Information on the socio-economic characteristics of successful and unscucessful applicants for shelter and credit programs.

b. Financial information on approval and repayment of loans.

c. Records of health centers.

d. Records of cooperatives and other community or business organizations.

Care must be taken in the use of this secondary data sources as the information may be inaccurate, incomplete or contain certain biases.

6. Ethnographic methods (Annex B Section 3)

Many ethnographers have criticized quantitative methods claiming they are inappropriate for community studies and evaluation. It is argued that questionnaires cannot be used to study organizational processes, to study how groups operate or to measure attitudes. The ethnographic methods proposed to overcome these problems involve the in-depth study of individuals, groups or the whole community. The methods seek to understand the way in which the community operates and to understand the meanings which people place on their world and on the project interventions.

A common technique is participant observation in which the researcher lives in, or spends extended periods of time, in the community or group. The purpose is to observe the natural behaviour of group members and to understand how they interact without asking them to explain or verablize their feelings or behaviour. Another method is to intensively study a particular person or family and to present a detailed descriptive monograph (Oscar Lewis' "Children of Sanchez" is the classic).

Ethnographic techniques are extremely useful in the evaluation of urban development, as there are many barriers between low-income participants and the implementing agencies which make it difficult to understand the feelings and behaviour of the population by using formal survey methods.

D. METHODS FOR MONITORING THE EFFICIENCY OF THE IMPLEMENTATION PROCESS

1. Monitoring the overall efficiency of project implementation

Many projects seek to develop institutions which can plan and implement future projects, and consequently it is important to assess the effects of the project on developing the capacity of institutions at the community, implementing and planning levels. It is important to understand the factors which affect institutional performance as well as to measure outcomes. To achieve this it is necessary to combine qualitative and descriptive techniques with the more usual quantitative indicators.

The evaluation will normally include a descriptive analysis of factors such as the following:

a. Achievement of program goals: Usually evaluated by using the management by objectives approaches explained in Chapter 2.

b. Satisfaction of participating agencies with interagency coordination:

This is achieved through participant observation,informal interviews and the use of formal questionnaires

c. Community satisfaction with the coordination with the implementing agencies:

Many organizational problems are caused by poor communication between the implementing agencies and the intended beneficiaries, and it is, therefore, important to include interviews with community leaders and residents. Useful techniques include participant observation, direct observation and meetings with community organizations. Structured questionnaires are useful for studying information flows and for comparing knowledge and opinions about the project in different sectors of the community. However, questionnaires are less effective for studying attitudes and conflicts due to people's reluctance to openly criticize the implementing agencies or the community leaders.

The amount and accuracy of information flowing between implementing agencies and the community should be measured as there are often

serious communication blockages of both downward and upward information flows. Typical examples include:

i. agencies not trying to explain "technical" information in the belief that the community will not understand;

ii. agencies only communicating with a small number of community leaders who then either do not pass on the information or who unintentionally distort it;

iii. deliberate efforts by certain groups within the community to distort information;

iv. different agencies using different and contradictory communication channels;

v. agencies being unresponsive to community complaints or questions; and

vi. social workers not communicating up through their organizations, either due to their low status or to their belief in a lack of interest or responsiveness of their organization.

d. Effectiveness of inter-agency meetings as communication and problem-solving mechanisms

Inter-agency meetings often provide a good indicator of how well coordination is taking place. Some of the indicators from the meetings (which can be studied by reviewing minutes or through direct observation) are:

i. how frequently meetings are held;

ii. which organizations participate and what level of staff do they send;

iii. are the major issues (defined in project progress and similar reports) discussed;

iv. are decisions made and carried out;

v. do these meetings provide an efficient information flow; and

vi. what types of issues are not discussed or resolved and what appear to be the reasons for this.

e. Effectiveness and efficiency of monitoring and evaluation systems

The speed of collection and comprehensiveness of coverage of the monitoring and evaluation systems are an indicator of managerial efficiency. Some of the specific indicators include:

i. Comparison of planned and actual production of reports;

ii. Quality of reports;

iii. Regularity and effectiveness of meetings to review and take action on the evaluation reports; and

iv. Opinions on the usefulness of the reports.

f. Internal organizational indicators

The following organizational indicators can be used:

i. clarity of organizational chart, definition of functions and forms of coordination;

ii. extent to which the actual system conforms to the organizational chart;

iii. main organizational bottlenecks and breakdowns;

iv. numbers and qualifications of staff at different levels;

v. number of unfilled positions at different levels (and reasons); and

vi. staff turnover at different levels (and reasons).

g. Financial administration and control

Some of the indicators of financial administration at the project and program level are the following:

i. quality and comprehensiveness of information on financial status;

ii. speed with which financial information is obtained and reports prepared;

iii. time taken to prepare and process disbursement requests and main bottlenecks;

iv. time between initiation of disbursement request and receipt of funds; and

v. conformity to financial goals.

2. Developing summary indicators

Although the descriptive analysis outlined above is useful for understanding the dynamics of a particular institution, it is also necessary to develop summary indicators which can be used to compare the performance of different institutions or to describe the evolution of one institution over the life of the project.

The following are some of the key indicators which can be used in this comparative analysis:

a. Achievement of project goals
b. General efficiency of organizational procedures and inter-agency interactions (including interactions with project beneficiaries and affected communities)
c. Speed of project implementation
d. Cost (as compared with target)
e. Quality of the project outputs
f. Accessibility/affordability to the target population
g. Replicability
h. Flexibility and adaptability

Each of these indicators can be reviewed separately, and then they can be combined to produce an overall index of project efficiency. Often there will be trade-offs between different indicators. For example, it may be possible to implement the project more quickly and cheaply if less attention is paid to ensuring that selected families fall within the targeted income ranges. Similar trade-offs may exist between replicability and speed/cost of implementation. Another common issue is that it may be possible to complete the project more rapidly by setting up a special unit not subject to normal administrative controls (and delays). However, once the first project is completed, the special unit will often be disbanded and no base will have been created for project replicability. In each project, management must decide the relative weighting to be given to each criterion in the overall evaluation.

The indicators can be used in various ways. The first is to present a separate evaluation of project performance with respect to each indicator, without necessarily trying to integrate them all into an overall index. A second possibility is to develop a system of ranking. For example, each indicator could be ranked from 1 (poor) to 10 (very good). Independent rankings can also be obtained from a number of different people and the

average computed. If there is a high level of agreement between judges, more reliance can be placed on the results.

A further refinement of the ranking system is to apply ranks to each project component (using a list similar to that outlined in the following section). An average is then calculated for each component. Table 3-4 presents an hypothetical example of this system. Eight project components are ranked using these indicators and an average rank is computed for each component and for each indicator. Participant selection, the installation of infrastructure, house construction, and project maintenance all received the relatively high overall mean ranking of 7, suggesting they performed well on all indicators. On the other hand, cost recovery, land acquisition, and the administration of the material loan programs received lower mean ranking of 4 or 5. Not all indicators were applied to all components as some were not appropriate. Care must be taken in the interpretation of these averages, and they should only be considered as providing a very general comparison between components.

Overall rankings can also be obtained on each indicator. Thus quality, goal achievement, and overall organizational efficiency were ranked relatively highly (7 or 8), whereas speed, cost, accessibility to the target groups, and flexibility of implementation were ranked lower.

3. Indicators of the efficiency of individual project components

It is also important for management to obtain information on the efficiency of each project component. Special criteria can be applied to each component but it is also useful to develop a set of standard indicators which can be used for comparative purposes. One possible set of such indicators is the following:

a. Speed of implementation
b. Cost of implementation as compared to the original estimates
c. Quality of the final product or service
d. Accessibility to the target population
e. Replicability of the procedures and design

A similar ranking procedure can be used to that described above. Table 3-5 identifies some of the factors which should be taken into account in the application of these indicators to 8 typical project components.

Table 3-4 Example of the application of ranks in the monitoring of overall project efficiency

INDICATORS OF EFFICIENCY

COMPONENTS	a	b	c	d	e	f	g	h	Mean	
Land acquisition 2	3	8	3	7	–	–	–	5		
Participant selection	8	7	8	7	8	8	–	4	7	
Material loans program	5	3	8	4	7	7	–	4	5	
Infrastructure install.	5	5	8	7	7	7	–	–	7	
House construction	6	4	9	6	9	8	–	9	7	
Cost recovery		4	3	–	–	4	4	–	3	4
Maintenance	8	7	8	–	7	7	–	7	7	
Community participation	6	4	7	7	7	6	–	7	6	
OVERALL RATING	6	5	8	6	7	6	7	6		

Use of ranks: 1 = Poor. 5 = Average. 10 = Excellent.

(i.e. if costs were higher than expected the ranking would be low).

In computing average ranks the average is rounded up to the nearest whole number.

Key:
- a. Achievement of goals
- b. Efficiency of procedures
- c. Speed of implementation
- d. Cost
- e. Quality
- f. Accessibility to target population
- g. Replicability
- h. Flexibility

Table 3-3 INDICATORS OF EFFICIENCY FOR TYPICAL PROJECT COMPONENTS

Project Element	Indicators of Project Efficiency: Speed of Implementation	Cost	Quality	Accessibility	Replicability
Planning and Design	i. Speed of completion by contrac- ii. Speed of completion by families. iii. Are delays caused by standards and inspection procedures.	i. Total cost per unit. ii. Effect on cost of each main component. iii. Effect of material quality on cost.	i. Family satisfaction. ii. Evaluation by architects or engineers. iii. Rental value.	i. Effects of design standards on costs. ii. Effects of standards on ease with which families can build.	i. Is design easy to follow and implement on large scale. ii. Are required materials available for large scale replicability.
Selection Procedures	i. Time between application and decision. ii. Total time between start and finish of selection.	i. Total cost including staff time.	i. Is the system accurate and fair.	i. Are there any groups sytematically under-represented. For example: women, self-employed.	i. Can the procedures be replicated in larger projects or in other areas which have more limited access to staff or computers.
Construction Process	i. Time taken for families to complete house. ii. Main causes of dealy.	i. Total cost to family and comparison with expected costs. ii. Main reason for higher costs.	i. Quality of contracting. ii. Quality of family construction. iii. Main difficulties.	i. Which type of family has problem with construction. ii. Are some families eliminated by self-help requirement.	i. Can the method be replicated on larger scale and in other areas.
Material Loans and Material Stores	i. Time taken to obtain materials. ii. Comparison with processing time in free market.	i. Total cost of running store, including staff time. ii. Comparison of costs of materials with free market. iii. Effects on costs of only providing certain types of materials.	i. Comparison of quality with materials from free ii. Comparing quality of construction with control group.	i. Do loans make it easier for certain groups to participate in project.	i. Could/should stores/loans be replicated in new projects.

Table 1 (Cont'd)

Project Element	Speed of Implementation	Cost	Quality	Accessibility	Replicability
Maintenance	i. Speed with which repairs and other services performed.	i. Effects of project design on cost of maintenance. ii. Costs to family/government of maintenance.	i. How well are services maintained. ii. How well are houses and private areas maintained.	i. Do maintenance policies limit access to certain groups of services.	i. Should maintenance policies/costs be replicated or modified.
Cost Recovery	i. Speed of collection. ii. Default rates.	i. Collection costs. ii. Costs of defaults.	i. General efficiency of collection		i. Could collection procedures be replicated.
Community Participation	i. Effect on speed of project implementation.	i. Effect on project costs.	i. Effects on quality of construction and maintenance. ii. Effect on consolidation of community and obtaining other facilities and services.	i. Are there groups who are helped to enter project through community action. ii. Are there groups who are eliminated or discouraged.	i. Could the community organization methods be replicated in other areas areas and in larger projects.
Project Occupation	i. Time from start of project to occu- ii. Main reason for delay.	i. Effects of costs on occupancy rates.		i. Which groups are not occupying and why. ii. Who is dropping out.	

4. Studying community level organizations

One of the objectives of many projects is to develop or strengthen community organizations, and consequently it is important to evaluate how these organizations are affected by the project. Some of the methods which can be used include:

a. Studying communication linkages between the organization and the community (how many people are informed about the organization and its activities, are people consulted etc).

b. Studying participation of different sectors of the comunity in the organization.

c. Studying the internal efficiency of the organization.

d. Observing decision making processes and the form of community participation.

e. Studying the interactions between the community organization, the executing agency and other outside groups.

E. THE DESIGN OF CONTINUOUS, PERIODIC AND ONE-TIME PROCESS MONITORING

1. Continuous panel studies

a. The panel methodology

The panel methodology consists of selecting a sample of people, households, small businesses etc and maintaining contact with them throughout the period of the study so as to record their response to the project and to assess the factors which have affected this response. The panel sample may only cover potential project beneficiaries, or it may also include a control group made up of households (businesses etc) with similar characteristivcs but who will not be affected by the project. The information can be collected through the application of a formal questionnaire, informal interviews, direct observation or by asking respondents to record the information on their own behaviour. In some cases, such as a cooperative, some of the information may be obtainable from secondary sources.

b. Initial Diagnosis

Before designing a study, it is advisable to conduct an initial diagnosis the purpose of which is to assess the issues which need to be studied, to test alternative research methods and to identify potential problems and issues which are likely to arise during the course of the study. The following are some of the steps commonly followed in the design and implementation of the diagnosis.

i. Define all of the groups likely to be affected by the project and contact them all (or use other methods to determine their likely reactions).

ii. Contact all major organizations outside the community who are inovlved in the project and determine their reactions.

iii. Conduct unstructured interviews with a small sample of beneficiaries to understand their perceptions and to test ways in which the study might be conducted.

iv. Review existing documentation on the project.

v. Spend time visiting the project areas and talking informally to residents so as to obtain a "feel" for the situation.

c. Defining the universe and selecting cases

The objective of a panel study is normally to make generalizations to a wider population, such as all project beneficiaries or all households in a certain category living in the city. In order to make valid generalizations it is essential to ensure that the cases are selected to be representative of the total population. With respect to the project population a list can usually be obtained of all participants and a random sample can easily be selected. Sometimes this can be made more difficult when records are not kept up to date.

The universe is more difficult to define for the control group. For example, if the study covers small businesses, should the control sample cover all businesses of a certain size in the city, only those in areas close to the project, or only those with similar economic activities. In practice the choice is also constrained by the cost and difficulty of selecting and studying a completely random sample.

d. Some common measurement problems

A common problem is that of households who drop-out of the sample, either by moving or because they refuse to continue. In most areas of the city there are high population turnover rates so it is quite common to find that 20% or more of a population group will move every year. This presents various problems for the study. The first is that the original sample becomes progressively smaller unless replacements are chosen. Even more serious are the differences between households who move and those who remain in the same place. As the study will only obtain information on the families who have not moved, this is likely to provide a biased picture of the total population. It may be for example, that the businessmen who moved were

those who were least successful. Consequently, a sample of only those who have remained will have a bias towards the most successful businesses. It is important to note that increasing the sample size by replacing drop-outs will not solve this problem of bias.

Respondent fatigue is another common problem. Even though respondents may be willing to cooperate at the start of the study, they are likely to become less willing to continue cooperating if the interviewer returns every month (or three months) to ask them the same questions. This may either lead to drop-outs or to respondents providing less information so as to end the interview more quickly.

The interview process itself may lead to distortion. The fact that a small businessman is constantly being asked to provide information on stocks, sales, profits etc may lead him to rationalize his accounting and stock control systems. If this happens, some of the improvements in his management system may be due to the interaction with the interviewer rather than to the effects of the program which is being evaluated. A similar phenomenon has been observed in health studies, where child-care and health practices may improve as a result of the observation.

e. Data collection

Although the panel study will innevitably rely primarily on one or two methods of data collection, it is essential to use the multi-method approach to check on the reliability of the data and to help with its interpretation. Some of the main data collection methods include:

i. Structured questionnaire in which information in obtained on income, employment, housing investment etc. Questions may also be included on the reasons for certain actions (why did the respondent decide to build another room etc) and on sources of finance and assistance.

ii. Observation guide: for example, observing the changes which have been made in the house. The method can also be used for observing changes in the general economic or physical conditions of the community. Photographs can provide a useful complement to the reported information.

iii. Case studies in which a small sample of individuals, households, businesses etc is studied intensively. An unstructured interview guide may be combined with observation and general conversation.

iv. Self-reporting by respondents. Keeping a record of income and expenditure is one common example.

v. Use of secondary data is sometimes possible (for example, credit or production/sales information from a cooperative).

f. Analysis

Statistical analysis can range from the presentation of simple graphs showing, for example, average investment in house construction, business sales, proportion of children with parasites; to more sophisticated types of multivariate analysis.

2. Periodic studies

Periodic studies follow a similar logic to the panel study in that they observe a process of change over time, but a smaller number of observation periods are used. Often a larger sample, and more extensive data collection methods are used in order to obtain a systematic comparison between certain points in time. The most common periodic study is where observations are made at the start of a project, at some point during the implementation process and again when the project is completed or nearing completion. For example:

a. A study to evaluate the effectiveness of the mutual help house construction process. Observations were made at the beginning of the process when the groups were being formed and the initial orientation given. They were repeated when the process was well underway. A final observation was made when the process had been completed.

The same observations about choice and selection of data collection techniques apply as for panel studies.

3. One-time studies

a. The approach

One-time studies may either be conducted as part of an overall and pre-planned evaluation design, or more commonly, to provide assistance when a problem has arisen or a decision has to be taken. A distinction can be made between a rapid study which is specifically designed to provide results within a given period of time (say a month), and a more intensive study where there is less time pressure and where the emphasis is on obtaining a more in-depth understanding of the processes involved. The choice between these two approaches is usually determined in part by the nature of the problem being studied and in part by the information needs of the organization.

In general a rapid study will be used when:

i. The information is needed by a certain date.

ii. It is conducted as part of an ongoing evaluation program where the goal is to produce a certain number of studies per year (for example, in some projects the goal is to produce an average of one study per month).

iii. Where the issues to be studied are relatively specific and are not perceived to be excessively complex.

iv. Where the option exists of conducting a more complete follow-up study.

In contrast, a more intensive study will be preferred when:

- It is known that the issues being studied are complex or delicate (for example, when a project has been rejected by a community; when relations between the community and the implementing agency are known to be bad)

- When it is desired to obtain a more in-depth understanding of the processes of change and of how the beliefs and attitudes of the community affect the project.

- When the study is used for quality control and to provide feedback on how the program is developing and what are some of the key policy issues to be addressed.

F. DEFINING THE STRATEGY FOR PROCESS MONITORING

Process monitoring should be given a much more central place in the overall evaluation strategy than it has conventionally enjoyed in the past. Although strategies must adapt to available resources, the following are some general guidelines:

1. Resources should be reserved to permit the conducting of one-time studies at fairly regularly intervals as they are a very valuable aid to project management. Mechanisms should be developed for selecting and reviewing these studies so that they became a regular part of project planning and review.

2. Given resource constraints, and the need for feedback to management, most one-time studies should probably be rapid.

3. At least one or two longitudinal panel studies should be conducted as they help understand the complexities of the development process and identify key policy issues requiring further study. These studies should always use a multi-method approach in which quantitative and qualitative methods are combined.

CHAPTER 4: IMPACT EVALUATION

This chapter presents alternative research designs for the evaluation of project impacts. The history of impact evaluation is reviewed and the debate between the advocates of qualitative and quantitative approaches is discussed. Key issues, including whether and when to conduct impact evaluations, are discussed and examples of different research designs are presented. A number of simple alternatives to large-scale quantitative evaluations are discussed before describing the design and analysis of quasi-experimental designs. A distinction is made between approaches which estimate net project impacts (the quasi-experimental design) and those which estimate cost-effectiveness (cost-benefits analysis, cost-effectiveness analysis and cost-utility analysis). Finally, guidelines are presented on how to choose the appropriate impact evaluation strategy. Technical material on sampling, research design and statistical analysis, is presented in annexes at the end of the document.

A. APPROACHES TO IMPACT EVALUATION

At the time when urban development programs began to expand rapidly in LDC's in the early seventies, many of the researchers responsible for the evaluation of these programs looked for guidance to the United States. Following the massive social investments of the War on Poverty, evaluation research had become a growth industry and a large research literature was being published. Two main approaches were recommended for the quantitative estimation of project impacts. The first sought to estimate the net impact of a project on the target population, and the second to compare the "effectiveness" of two or more alternative strategies.

The first approach estimates the net impacts of a project by comparing the conditions of beneficiaries after the project with what they would have been if the project had not taken place. Most textbooks at this time recommended the randomized experimental design in which subjects are randomly assigned to receive "treatments" (project services) or to form part of a "control group" which would not receive the treatments. Evaluations of this kind were implemented on several of the major War On Poverty programs in the Sixties, including the "negative income tax" programs; housing allowance programs; and comprehensive employment training programs. Many of these evaluations were mandated by the U.S Congress to assess whether the programs offered cost effective ways to achieve certain quantifiable social and economic goals (such as improving housing quality for low income families). The reviews of these evaluations (see for example Rossi and Wright,1984; Abt,1976; and the Evaluation Studies Annual Reviews for 1979 and 1981) demonstrate that these types of experimental designs are technically feasible and can provide operationally useful recommendations - some of which are able to save billions of dollars (see below).

It was soon realized, however, that the randomized experimental approach is expensive and in many cases not technically or politically feasible. For these reasons a number of quasi-experimental designs were developed in which statistical rigour could be applied when random assignment and other conditions of the experimental design could not be met (Campbell and

Stanley, 1966 and Cook and Campbell, 1979). Designs covered situations in which participants were selected on a non-random basis and where a "patched-up" control group had to be selected; or where it was not possible to measure conditions before the program began.

The second approach compares the effects of alternative strategies to determine which produces the greatest benefits for a given investment. In those cases where the benefits of each alternative can be given a monetary value (monetized), the most widely used technique is cost-benefit analysis. One of the earliest large scale applications was the evaluation of Federal Manpower Training Programs (Glennan,1972) where benefits were measured in terms of increased income of trainees. Where benefits cannot be monetized (and this has proved to be the case with many social programs) then cost-effectiveness analysis is used to compare the costs of producing a given output (increasing reading skills by a certain number of points, providing health care for a certain number of patients).

In 1975 the World Bank and the International Development Research Center of Canada initiated a 5 year evaluation of four of the first Bank urban shelter projects in El Salvador, Zambia, Senegal and the Philippines. All of these evaluations attempted to implement a quasi-experimental design with before and after measurements being conducted on a sample of participants and a control group. The initial objective was to measure net impacts on variables such as income, employment, health, housing investment, family size, and migration.

The usefulness of this approach soon began to be questioned by the executing agencies and World Bank project staff. A first problem was that a longitudinal impact evaluation did not help managers to improve the performance of their projects, and consequently additional ad hoc studies had to be carried out to provide more rapid feedback on short and medium range problems related to project implementation. Second, delays and changes in the project meant that by the time evaluation results were available, many of the findings were either already known or no longer relevant. Finally questions arose as to the validity of information obtained through formal quantitative survey approaches.

Partly for these reasons interest was shown in qualitative, ethnographic techniques such as participant observation in which the

researcher lived in the community and sought to understand the ways in which intended beneficiaries perceived and responded to the project. Advocates of these approaches argued that they were much more flexible, produced results more rapidly, were able to investigate delicate or conflictive issues, produced more reliable data and were able to study processes as well as outcomes. Importance was also attached to evaluating the projects from the point of view of the intended beneficiaries as many development projects were perceived to have been imposed on beneficiaries with very little attempt being made to understand their point of view.

For somewhat similar reasons, a number of rapid impact or rapid reconnaissance methods were developed which could provide rapid and economical indicators of project progress and impacts. AID, for example, has developed a rapid impact evaluation strategy which uses a multimethod approach to provide a quite comprehensive evaluation of a project's impacts with only about 4 weeks of fieldwork (see for example "Assisting small businesses in Francophone Africa: the Entente Fund African Enterprises Program", USAID,1982). Although less formalized than the AID approach rapid evaluation techniques have been successfully applied in a number of World Bank urban projects (see Chapter 3)

Initial experience with both ethnographic and rapid feedback techniques revealed some of their limitations as well as their strengths. The first problem is that in order to evaluate impacts and to compare different groups it is necessary to have some basis for comparision and quantification. This cannot be achieved with most qualitative approaches, so that it is difficult to generalize from the particular cases studied. It also became apparent that the claim of economy and speed is much less true when larger projects have to be evaluated. It is becoming clear that an evaluation strategy must use a multi-method approach in which qualitative and quantitative techniques are combined. Each method has its strengths and weaknesses, and an evaluation which relies on only one technique will innevitably give a limited and probably biased perspective. The combination of techniques for the purpose of consistency is known as triangulation and will be discussed later in this chapter.

B. SOME KEY ISSUES IN THE DESIGN OF URBAN IMPACT EVALUATIONS

The above review highlights a number of key issues which must be addressed by management and researchers in the definition of an impact evaluation strategy:

1. Is it necessary to evaluate project impacts?

While all project managers accept the practical utility of performance and process monitoring, many would question the necessity of conducting impact evaluations. It is argued that impact evaluations are too slow and expensive and that managers are concerned with implementation issues such as efficiency and the achievement of specific objectives, rather than with understanding whether their project has produced general changes in the conditions of the affected population. The success or failure of a project depends, it is argued, on a set of very specific circumstances and little useful information can be gained from the generalized statements produced by an impact evaluation.

In reply, defenders of quantitative impact evaluations point out that projects are intended to improve the social and economic conditions of the target population, and it is essential to know whether they are able to do this. Development agencies seek to maximize the utility of scarce resources, and if investment decisions are based on wrong assumptions then these scarce resources may be used for projects which have a much lower rate of return than one of the alternative strategies which could have been used. If studies are not conducted on likely outcomes and impacts, decisions on future projects costing millions of dollars would be made without systematic information on their cost-effectiveness or ability to produce the desired benefits. Decisions on future development strategies should make use of all useful sources of information which can be obtained in a cost-effective manner.

Part of the confusion on the utility of impact evaluations has arisen from a failure to clearly identify who are the audiences. Although many types of cost-effectiveness and impact analysis may not be of interest to the project manager, they are important for finance ministries and development planning agencies. Impact evaluations can be useful in at least 3 ways:

a. An impact report can assess the extent to which the project objectives have been achieved and the factors which have affected the outcomes. The reports can be of direct assistance in the design of future

projects by making goals more realistic or by defining more clearly the target population. (See Box 4-1)

b. An impact evaluation can compare different projects, for example in terms of their benefit/cost ratios, and provide guidance as to which project or projects are likely to be most cost-effective. (See Box 4-2)

BOX 4-1 IMPACT EVALUATIONS CAN HELP DEFINE REALISTIC PROJECT OBJECTIVES

Project designs will often be overly ambitious in the way in which the objectives of a new project are specified. The following examples are typical of ways in which impact evaluations can help make project objectives more realistic:

** An evaluation of small business development programs in francophone Africa (AID 1982) showed that many of the original goals were unrealistic in terms of expected impacts and that a clearer definition of the target population was needed.

** An evaluation of a cooperative program in El Salvador (see Section 3 of this chapter) showed that the cooperatives were unlikely to produce any major impact on employment of the low-income population, even though they significantly improved the economic conditions of the small number of families directly involved.

c. Many projects, such as squatter upgrading, involve a large number of components (water, roads, sanitation, housing credit, community centers, technical assistance etc) which different families receive in different combinations. A well designed impact evaluation can estimate the overall impact of the project and also assess the relative contribution of each component.

2. Is it possible to measure impacts and assess causation?

It is often argued that given the complexities of the urban environment, it is not possible to identify the effects produced by a particular set of project inputs. The urban environment is so different from animal psychology and agricultural research, that the randomized experimental designs developed in these fields are much less appropriate for urban research.

In reply, a number of examples can be cited from the U.S in which randomized experimental designs have been successfully conducted and where reasonably precise causal relationships have been established. Section C

gives the example of the Experimental Housing Allowance Program. Another example is the evaluation of the effects of U.S Negative Income Tax Experiment on school performance (Maynard and Murname, 1981). In addition, the careful use of quasi-experimental designs in combination with qualitative techniques can provide operationally useful estimates of the degree of association between the project interventions and changes in the conditions of the population for many projects where randomization is not possible.

BOX 4-2 USING IMPACT EVALUATIONS TO COMPARE THE BENEFIT/COST RATIOS OF ALTERNATIVE URBAN PROJECTS

A comparison was made in El Salvador of 8 low-income housing options in terms of their costs and benefits to both the nation and the participating households. It was shown that progressive development approaches (sites and services and squatter upgrading) had higher rates of return for both the nation and the participating households, than any of the conventional housing programs. The large scale implementation of progressive development would be very cost effective and would make projects more accessible to low-income households.

Source: Fernandez-Palacios and Bamberger "Economic Analysis of Low-Cost Housing Options in El Salvador". 1984

3: Do cost-effective methods exist for impact evaluation?

One of the most frequent criticisms of quantitative impact evaluations is that they are expensive in terms of both financial and human resources. However, it is meaningless to claim that an evaluation is "expensive" without considering its cost-effectiveness with respect to the benefits produced. For example, the evaluation of the Experimental Housing Allowance Program described in Section C cost about $50 million which would seem to be quite expensive. However, the evaluation demonstrated that housing allowances were a much more cost-effective than conventional construction programs in improving the housing conditions of low-income families. It has been estimated that the findings of the evaluation could save the federal government between $ 7 billion and $ 8 billion per year. Assuming these estimates are correct, the evaluation was clearly very cost-effective.

Although it is obviously not possible to predict the outcomes of an evaluation it is often possible to estimate the potential savings which the

evaluation could produce. Table 4.1 presents a hypothetical example of a pilot small business program costing around $10 million, whose purpose is to test the efficacy of technical assistance and innovative types of credit as means for generating new employment and increasing income of poor urban households. If the project is successful it is intended that a second project, costing about $ 100 million would be implemented in 10 cities. It is believed that the project will offer a more cost-effective way to generate employment and income than conventional commercial credit. A group of local consultants were asked to submit a proposal for an impact evaluation of the pilot project. They estimated that the study, which would interview a sample of beneficiaries and a control group several times throughout the life of the project, would cost around $ 500,000. How could the cost-effectiveness of the proposed evaluation be assessed? It is known that it costs about $ 5,000 to create a job through commercial credit and that on average each job will increase income by $4,500 over a period of 5 years. A group of experts indicated that under the best possible scenario the new approach could reduce the cost of employment generation to around $1,000 per job and could generate $6,000 of additional income per job over the 5 year period. The experts also indicated, however, that many programs of this kind have had very limited success and under the worst scenario it could cost as much as $20,000 to create a new job. Part A of the table shows that under the best scenario, the second project could produce up to 100,000 jobs and $600 million of additional income. Under the worst scenario only 5,000 jobs would be created and $25,000 of income. If, instead of approving a special project, the funds were made available to commercial banks to expand small business credit, it is estimated that 20,000 jobs could be created. Part B shows the costs of making a wrong decision about whether to approve a second project. If a second project is approved under the worst scenario it will cost $75 million more to create the jobs than it would have cost through commercial credit. On the other hand if under the best scenario the decision was made not to approve the second project, it would cost $80 million more to produce the 100,000 jobs through commercial credit. Under these circumstances it would seem very cost-effective to invest $500,000 in a careful evaluation if it could significantly increase the likelihood of making the correct decision and saving up to $75-80 million. Although information on impacts can sometimes be obtained cheaply and easily for small projects, large and complex projects usually require more

Table 4.1 Hypothetical example of the potential savings from an impact evaluation of a pilot employment project

A. PROJECTED OUTCOMES FOR $ 100 million PROJECT BASED ON BEST AND WORST ASSUMPTIONS OF EXPERTS

	Assumptions by experts about possible range		Projected outcomes for $ 100 million project	
	Cost per job created	Income generated per new job over 5 years	Jobs created	Income created over 5 years
BEST SCENARIO	$ 1,000	$ 6,000	100,000	$ 600 million
WORST SCENARIO	$20,000	$ 5,000	5,000	$ 25 million
CONVENTIONAL COMMERCIAL CREDIT	$ 5,000	$ 4,500	20,000	$ 90 million

B. PROJECTED COSTS OF WRONG DECISION ABOUT WHETHER TO APPROVE PROJECT OR TO USE FUNDS THROUGH COMMERCIAL BANKS

Decision taken	Cost to produce the same number of jobs if alternative decision had been taken	Household income lost over 5 year period through not making alternative decision
Approving project under worst scenario	Loss of $ 75 millions	Loss of $ 65 millions
Not approving project under best scenario	Loss of $ 80 million	Loss of $ 510 million

A WELL DESIGNED IMPACT EVALUATION COULD SAVE BETWEEN $ 75-80 MILLION BY AVOIDING A WRONG DECISION ON WHETHER TO APPROVE A SPECIAL PROJECT OR PROVIDE CREDIT TO COMMERCIAL BANKS

rigorous analysis in order to assess the likely outcomes of new types of programs and to compare the outcomes with other investment options.

The purpose of this example is to show that the decision on whether or not to conduct a rigorous impact evaluation should be based not simply on its cost, but on its cost-effectiveness. The cost, complexity and duration of an impact evaluation depends on the types of information which are required and the needed level of precision. A number of rapid impact evaluation methods exist which can provide operationally useful, but more general, information in very short periods of time and at a low cost. Management must decide what information is really required so that the appropriate design can be developed.

4. Quantitative versus qualitative methods

There is a continuing discussion among evaluation practitioners as to whether quantitative or qualitative methods are better. The debate is often conducted in quite heated terms as it involves philosophical and ethical issues as well as methodology. Also, many researchers have specialized in a particular technique and hence become very defensive when it is criticized. The approach adopted in the present document is that all techniques have their strengths and weaknesses and that a multi-method approach should always be used in which quantitative and qualitative methods are combined. Box 4-3 illustrates the utility of combining ethnographic methods and econometric analysis in the evaluation of interhousehold transfer behaviour in Colombia.

C. EXAMPLES OF IMPACT EVALUATIONS

In this section examples are presented of 3 different approaches to impact evaluation. The first example, the US Experimental Housing Allowance Program, is one of the few urban evaluations which comes close to using a randomized experimental design. The second example, which evaluates the impacts of a housing program in El Salvador on income and employment, is a typical example of a quasi-experimental design in which research methods are adapted to the realities of the urban context in which the project is developed. The third, an evaluation of the social and economic impacts of a cooperative program in El Salvador, is an example of a rapid impact evaluation in which qualitative and simple quantitative methods are combined.

1. A large-scale randomized experimental design: the experimental Housing Allowance Program

The Experimental Housing Allowance Program (EHAB) sought to test the efficacy of providing poor households with allowances they could use to improve their present housing or to find better rental accommodation. It was believed that stimulating the private housing market would be more effective than the government directly providing housing.

BOX 4-3 AN EXAMPLE OF THE BENEFITS OF COMBINING ETHNOGRAPHIC CASE STUDIES AND ECONOMETRIC METHODS

Evaluation of the impact of an upgrading project on interhousehold transfers in Cartagena, Colombia

Earlier evaluation studies in El Salvador and the Philippines suggested that when families in low-income areas have the opportunity to invest in housing this will increase the amount of interhousehold income transfers they receive from relatives and friends. The studies also showed that transfers are "earmarked" for expenditure on basic needs consumption (food, clothing etc). These hypotheses were tested in a World Bank upgrading project in Cartagena, Colombia by combining econometric analysis and in-depth case studies. The econometric analysis of expenditure patterns during the previous month showed a statistical relationship between the proportion of household income received in the form of transfers and the proportion spent on basic needs - in other words there was clear evidence of the transfers being "earmarked". However, the case studies and participant observation revealed that families considered that the transfers they give or receive can be used for any type of expenditure. The reason for the discrepancy in the findings is that the econometric analysis focussed on one particular transfer act whereas the case studies analyzed a complex and ongoing process of social interaction involving many different transfer acts. If only the econometric analysis had been conducted it would have been wrongly assumed that transfers are isolated events, whereas if only the case studies had been used, the behavioural regularities between transfer receipts and expenditure patterns would have been missed.

Source: Kaufmann and Bamberger "Income transfers and urban projects: policy and research issues"1984.

a. Design

A sample of approximately 1800 low income households was selected in two US cities. The sample was randomly divided between an experimental group which received housing allowances and a control group which did not. A factorial design was used in which 19 treatment groups were created. Groups varied in terms of whether payments were based on the difference between their present housing conditions and the minimum satisfactory standard defined by the researchers, or whether they were paid a proportion of their rent.

Three experiments were conducted: a 3 year demand study to observe the effect of increased income (housing allowance) and reduced prices (payment of a proportion of rent) on housing consumption; a 10 year supply study to observe the impact of increased demand on prices and housing supply; and a third experiment (conducted in other cities) to assess the efficacy of different methods of administering the program. Participating households were required to provide monthly information and to respond to periodic interviews. The control group received no assistance but was asked to provide similar information, for which they received a small payment.

b. Some findings

Participation rates (the proportion of families willing to make the necessary repairs to their house in order to participate) were relatively low with only 27% participating in the demand experiments and between 33 and 42% in the supply experiments. Where subsidies were unconstrained, only 10% of the additional income in Pittsburg and 25% in Phoenix was used to increase housing consumption. The rent subsidy treatment significantly decreased the average proportion of income paid in rent from 40 to 25%. Virtually no impact was detected on housing prices and housing supply.

Although the impacts were relatively small, the EHAP program appeared to be a cost-effective way of improving housing quality for specific target groups. The observed improvements were produced significantly more cheaply than through conventional programs and it has been estimated that the implementation of the evaluation recommendations could potentially save federal housing agencies as much as $7-8 billion per year.

c. Methodological issues

The 3 year cut-off point may be too early to assess some of the impacts. There is also the problem that participants' behaviour may have been affected by the knowledge that the subsidies would only last for 3 years.

The cost of the evaluation is another issue. The experimental program, including the housing subsidies, cost over $150 million of which the evaluation component itself cost around $50 million. However, given the estimated potential annual savings of $7-8 billion, the evaluation appears to have been very cost-effective.

A final problem is that the United States experience suggests that randomized experimental designs can only be used with experimental programs which are not yet being implemented on a regular basis. It is difficult to convince politicians or project managers to accept that certain people should be excluded from access to the benefits of a regular programs so as to form part of a control group or that participants should be selected randomly rather than according to need (or some political criterion)

2. A quasi-experimental design: evaluating the impact of a sites and services housing project on employment and income in El Salvador

One of the objectives of the sites and services project was to increase household income and labor force participation rates.

a. Design

A sample of 196 future participants in a sites and services project were chosen for the study. A control group was selected from among households living in the 3 main types of low-income housing in the city, with aproximately 100 households being selected randomly from each of the 3 types. A socio-economic questionnaire was administered to both participant and control groups in 1976, just before the project began and again in 1979 and 1980. If a family moved it was replaced in the sample by the new occupant of the same structure. Multiple regression analysis was used to statistically match the experimental and control samples in terms of a set of socio-economic characteristics.

b. Some findings

It was not possible to detect any difference between the experimental and control groups in terms of the rates of change of total household income and in labor market earnings between 1976 and 1980. The only

significant difference was that the labor force participation rate of secondary workers fell by 30% for the control group but remained unchanged for participants. This suggested that during the very difficult economic and political period through which El Salvador was passing, secondary workers in project areas were more successful in keeping their jobs than were workers in the control areas. Two alternative hypotheses were advanced to explain these findings: (a) Employers gave preference to project residents because they considered that people who are buying a house are likely to be more reliable workers (b) the need to keep up mortgage payments gave secondary workers more incentive to remain in the labor force.

c. Methodological issues

The political crisis in El Salvador meant that any economic benefits the project might have produced were largely cancelled out by the economic depression. The inconclusive findings cannot therefore be interpreted to mean that the project would not have produced economic benefits under more favorable circumstances. As always with a quasi-experimental design the question arises as to how comparable the experimental and control groups really were at the start of the project. Even though the groups were matched in terms of income and other characteristics, there was no way to control for differences in variables such as motivation. The issue also arises as to how well income and employment were measured. It could happen, for example, that project participants were more willing to cooperate and provide accurate information than the control group.

3. Rapid impact evaluation employing a multi-method approach: evaluating the social and economic impacts of a cooperative program in El Salvador

The production cooperatives were designed to improve the economic conditions of communities in which housing programs were being organized and to promote grass-roots community organizations capable of understanding and improving the social and economic conditions of the communities they represented.

a. Design

The study, which took about 2 months to complete, involved one fulltime researcher and a part-time director. Interviews were conducted with leaders and community organizers in six cooperatives and secondary data was reviewed. The largest cooperative (dress making and rug-making) was

studied more intensively through: (a) participation in meetings and cooperative activities (b) preparation of intensive case studies on 12 cooperative members (c) review of secondary data from the implementing agency and the cooperative (d) open interviews with members of the directorate (e) asking all 50 members to complete a form indicating their incomes and work hours before the cooperative began and at the present time (f) interviews with project staff.

b. Some findings

The quantitative impact on the employment situation of the whole community was relatively small with only 94 people out of a total community labor force of around 2700 being employed by the cooperative. However, the cooperatives significantly improved income and work conditions for their members but the potential for growth was severely constrained by the local market.

The participatory organizational style of the cooperative made members much more aware of the political and economic reality within which they lived and increased their involvement in outside political activities such as contacts with trade unions. Finally, the cooperative increased the self-confidence of its female members many of whom became more self-assertive in their domestic relations and gained more independence in the management of their own lives.

c. Methodological issues

The estimates of project impact on income were obtained by asking respondents to recall how much they had been earning 3 years ago. This method is subject to considerable measurement error and possibly bias. It is also difficult to interpret the findings when there is no control group with which they can be compared.

Several of the cooperatives also had a strained relationship with local landowners or large businesses which may have inhibited them from freely expressing their opinions.

D. ALTERNATIVES TO LARGE SCALE QUANTITATIVE EVALUATION DESIGNS

1. The selection of the most appropriate research techniques and the use of triangulation

Annex B points out that social science has become increasingly concerned with refining research techniques, but relatively little attention has been paid to the question of how to choose the best technique for a particular purpose. The choice of research methods has a profound effect on the types of information which are obtained and the types of conclusions which are drawn about project impacts. For example, the use of a precoded questionnaire forces respondents to choose between a limited number of predetermined options (a list of community problems, reasons for migrating to the city, things they like and dislike about the cooperative). If, instead the researcher had lived in the community and had tried to observe and understand the main concerns which people expressed in their day-to-day activities and in community meetings, it is quite possible that different conclusions might have been reached. A structured questionnaire is appropriate when information must be collected on specific topics or when a hypothesis is being tested. However, it is much less effective for understanding the opinions of the community or for identifying unexpected outcomes.

No research method is infallible and consequently the evaluation design should always use the principle of triangulation and combine several data collection methods to provide consistency checks (Annex B Section 1).

2. Participant observation and related ethnographic approaches

Considerable interest has been expressed in recent years in the types of qualitative and ethnographic approaches described in Annex B. The participant observer becomes very closely involved in the community or group being studied, and either lives in the community or spends a great deal of time there. The approaches are intended to give the researcher a deeper understanding of the culture of the group and of the meanings which its members give to the subject being studied. In the case of an urban development project, the purpose is to understand how people perceive the project, who does and does not benefit and what effects it has on individual families and the community in general.

There are three main ways in which qualitative techniques can be used in impact evaluation.

a. In those cases where it is not necessary to obtain precise quantitative estimates of impacts, use may be made of the rapid evaluation techniques discussed in Chapter 3. These studies usually combine qualitative and simple quantitative techniques as in the evaluation of the cooperative program in El Salvador (Section C.3 of this chapter).

b. Unstructured interviews can be used at the start of a survey to define hypotheses and to help develop instruments. They can also be used after the statistical analysis has been completed to help interpret the findings and to provide a consistency check (triangulation). In this case the analysis is primarily quantitative with the qualitative techniques having a supporting function.

c. A complementary design can be used in which quantitative and qualitative techniques have equal importance and are used to complement each other at every stage of the research. Box 4-3 is an example of this approach.

Despite their recent popularity, qualitative methods have a number of limitations when used for impact evaluation. A recent review article summarizes some of these issues with respect to the US evaluation experience:

> "It is however,equally clear that qualitative evaluations have their limits as well. However inexpensive they may be for single, small-scale projects, they are very expensive and not very sensible approaches to the evaluation of fully developed programs that have quite specific goals. Qualitative evaluations are very labor intensive and cannot be used on very many sites except at considerable cost. Furthermore, qualitative approaches rarely provide estimates of the effects that are either very precise or free and clear of possible confounding factors. Indeed, the only large-scale programs to which qualitative approaches were applied had vaguely stated goals e.g. Model Cities(Kaplan 1973) and revenue sharing (Nathan et al 1981). In these evaluations, the findings were composed more of descriptions of program oepration than of assessments of programs effects."
> (Rossi and Wright,1984 page 343)

The issues relate to whether qualitative methods can assess impacts as well as processes, and whether they can control for possible confounding effects. There is also an issue of cost when the techniques are applied to large scale projects.

3. Rapid impact studies

Rapid studies are designed to produce rapid and economical estimates of project impact. Most of the methods combine quantitative and qualitative techniques both as consistency checks and to use the cheapest method of obtaining the required information. These techniques are discussed more fully in Annex B Sections 5,6 and 8 and in the Boxes given in Chapter 3.

A number of rapid evaluation techniques have been developed which rely largely on easily observed indicators such as the materials used in house construction or the range of products on sale in community stores. Although these indicators can be used to generate hypotheses and as a consistency check, they are difficult to use on their own. One of the most difficult problems is to know how to interpret what is observed (see Annex B)

When rapid evaluations are conducted, it is advisable to use a multi-method approach which combines several techniques of which the following are typical examples:

a. Informal interviews with representatives of all groups likely to be affected by the project.

b. Direct observation of the socio-economic characteristics of the community.

c. Use of available secondary data sets.

d. Participant observation in which the researcher attends meetings, observes the program in action, etc.

e. Rapid sample survey.

4. Simple quantitative methods

Although it will frequently be necessary to use the more rigorous techniques discussed in Section E, in many cases it is sufficient to use much simpler (and more economical) statistical procedures. For example:

a. A cross-sectional survey can be conducted at the end of the project in which a sample of project beneficiaries are compared with a control group which did not have access to the project. Annex F explains the statistical procedures such as regression analysis which can be used to

maximize the efficiency of this and similar designs. The results will be more valid if qualitative techniques are also used and if questions are included on the causes of changes (for example, participants are asked why they made investments in their house and whether the presence of the project affected their decision).

b. In projects such as squatter upgrading, where it may not be possible to identify an independent control group, it is possible to estimate project impact through the use of the regression procedures given in Annex F.

Despite their theoretical limitations, both of these types of analysis can be useful in identifying i. whether the project has had any identifiable effects and ii. which groups seem to have been most and least affected.

5. Using secondary data

Before embarking on an expensive and time-consuming data collection exercise, the evaluator should determine whether estimates of impact can be obtained through the use of existing data. For example:

a. The records of a cooperative or small business credit program may provide detailed information on the number of new businesses started and changes in employment, income, productivity etc. The problem with the use of this data is that without a control group it is difficult to know whether the changes are due to the project or whether they are occurring throughout the city.

b. Public health programs and clinics may keep records on changes in levels and types of illness and on the quality of water supply, sanitary services, etc.

c. Community organizations may keep records on membership, financial resources, numbers and types of projects organized, etc.

d. Census data or studies conducted by other organizations may provide some control data against which to compare the changes observed in the project areas.

E. QUANTITATIVE ESTIMATES OF NET PROJECT IMPACTS

The purpose of most types of quantitative impact evaluations is to estimate the net impacts produced by a project on the immediate beneficiaries or on a wider target population. The objective is to estimate "how much

better off" beneficiaries are as a result of having participated in the project. As the urban environment is constantly changing, the evaluation design must be able to isolate the effects of the many changes unrelated to the project which are taking place, so as to determine the direct impact of the project interventions being studied. In many types of analysis, impacts or benefits will be compared with costs so as to estimate the return on the resources which have been invested. In this section we describe the main stages in the design and analysis of an impact evaluation of this kind.

1. Modeling the implementation process

Whenever an evaluation of project impact or effectiveness is planned the first step should be to prepare a conceptual framework which describes the project implementation process and the design assumptions on which it is based. Annex A describes the main stages in the definition of this model and illustrates them for an artesan credit program (See Fig 4.1 in this chapter). Let us assume that when the second measurement was made 2 years after the project had begun, the evaluation showed there had been no significant increase in the output or earnings of the artesans participating in the program. What inferences can we make about the potential utility of the program? There are a number of different reasons why the project might not have produced the expected impacts:

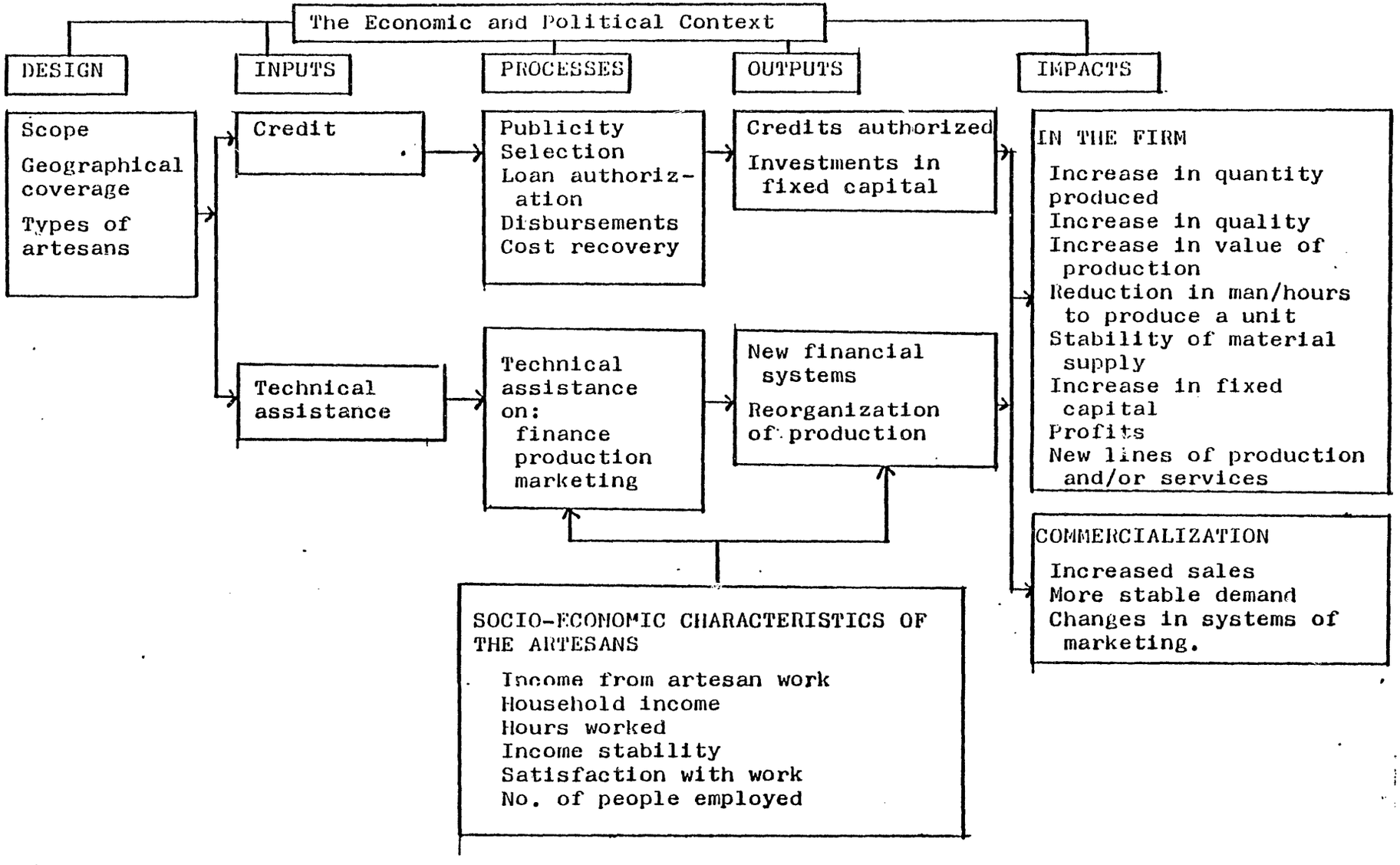
a. The design assumptions of the project may have been wrong. Artesan credit is not an effective way to improve the economic conditions of the urban informal sector.

b. The amounts or types of inputs were insufficient.

c. There were problems with the processes through which the program was implemented. Perhaps potential candidates were not aware of the program, or the loan disbursement procedures were slow and cumbersome.

d. The expected linkages between outputs and impacts may not have taken place. For example, artesans may use their credit to increase their fixed capital but not to hire more labor.

Figure 4.1: CONCEPTUAL FRAMEWORK FOR AN IMPACT EVALUATION: EXAMPLE OF AN ARTESAN CREDIT PROGRAM



e. The socio-economic characteristics of the participants may have been different from what had been expected. For example, they may have had less savings than expected or may have been involved in more capital intensive lines of production.

Only the first of these explanations would imply that the artesan credit program was not a potentially useful way to improve the economic conditions of the informal urban sector. All of the other explanations suggest that the disappointing outcomes were the result of problems in the way the project was implemented or of external pressures.

If the study had only provided statistical information on impacts, the policy recommendation would have been that the project was not effective and by implication should not be replicated. However, the above conceptual framework provides a great deal more information. Even when the expected impacts are not achieved, in most cases the policy implication will not be to terminate the program but rather to make improvements in the way it is designed or implemented.

Annex F Section 5 explains how Path Analysis can be used to statistically test the causal hypotheses described by the models.

2. Operational definitions of expected impacts

Once the model has been described, the next step is to define the expected impacts and to specify how they can be measured. In many projects this is complicated by the fact that intended impacts have only been defined in a very general way (for example "to improve the health of project beneficiaries"), or because there was an overly ambitious statement of expected impacts (for example it may be claimed that a small credit program will improve the economic conditions of the whole project population).

The evaluators should also be aware that as a result of the project intervention a number of unanticipated and often very signficant outcomes will occur. For example, work groups which were originally organized to build houses may become permanent community organizations which go on to implement other projects which were not originally planned. The evaluation design must have the flexibility to identify and measure these unexpected outcomes. In order to do this it is essential to use a multi-method evaluation design, as many of the conventional designs can only measure intended effects.

Table 4-2 gives examples of 6 types of impacts the project might wish to produce and of indicators which could be used to measure each one. The objectives must be related logically to the project model and be defined in terms of measurable indicators.

One of the objectives stated in Table 4-2 is to raise income of project households. 3 measurable indicators of income change are suggested:

a. Increased rental income from larger and better quality houses with improved access to services.
b. Increased income transfers from relatives to help with house construction and mortgage payments.
c. Through employment effects generated by increased house construction.

If the researcher is given this type of information the evaluation design can measure whether the expected changes have taken place and also indicate why the outcomes are different than expected. For example, if increased employment opportunities do not occur, it will be possible to evaluate whether this is due to macro-economic factors affecting demand for labor, to the location of the project site or to the fact that employers are less interested in hiring project participants than had been expected. Using this type of model, the information is much more useful to policy makers than a simple statement that "there was no significant impact on employment".

In addition to their direct impacts on project participants, most projects are also intended to have some wider impacts on: the total low-income population (improved access to housing, lower rents, etc.); management of the city (increased taxation, improved transport systems) and housing and urban development policy (changed attitudes to sites and services and upgrading; revision of housing subsidies; provision of new sources of housing finance; increased involvement of the private sector, etc.).

3. Experimental and quasi-experimental designs: the adaptation of textbook evaluation designs to the realities of the urban context.

In order to provide the information needed by managers and development planners, it is necessary not only to determine whether changes have taken place in the project population, but also to determine whether the changes are due to the project or to other unrelated factors. Cities are in a

Table 4.2: EXAMPLES OF 6 MICRO-LEVEL PROJECT IMPACTS AND SOME INDICATORS WHICH COULD BE USED TO MEASURE THEM

TYPE OF CONTACT	QUANTIFIABLE INDICATORS
1. Employment and Income	1. Total family income 2. Sources of income 3. Income stability 4. Type of employment 5. Number of people working 6. Labor force participation rate of a particular group, eg. women 7. Proportion of self-employed 8. Proportion working in formal and informal employment sectors
2. Demographic Characteristics of the Family	1. Family size 2. Age composition 3. Education of household head 4. Proportion of children attending school 5. Civil status of household head 6. Geographical mobility
3. Housing Costs, Quality and Value	1. House value 2. Construction quality 3. House Size 4. Access to services
4. Health	1. Infant mortality rate 2. Main types of intestinal infection 3. Time lost from work or school due to illness 4. Access to medical services 5. Amount spent on medical services 6. Anthropometric measures of weight and height
5. Consumption Patterns	1. Amount spent on housing 2. Amount spent on food 3. Amount spent on clothing 4. Amount spent on transportation 5. Amount spent on health 6. Amount saved
6. Community Participation and Attitudes	1. Number of community organizations in which families participate 2. Number of friends in the project 3. Political, social and religious organizations 4. Participation in mutual help programs 5. Satisfaction with the community 6. Satisfaction with ones own social economic and political situation

constant state of change so that project participants are subject to many other factors in addition to the project. The separation of the project impact from that of other factors requires a research design which can control for the effect of these external factors. The ideal way to do this would be to use one of the true experimental designs described in Annex E, but unfortunately the conditions required by the experimental paradigm can usually not be met in the urban context. With a few exceptions (such as the housing allowance experiments described earlier in this chapter) it is almost never possible to randomly assign subjects to the project and control groups. Subjects are either selected on the basis of need, qualification or some explicit or implicit political consideration. It is also never possible to exclude external factors such as flooding, political intervention, relocation of employment centers and macro-economic trends. Consequently an urban impact evaluation, however carefully designed, can rarely approach the statistical exactitude of the animal psychologist or the agricultural researcher.

Evaluation researchers disagree about the implications of this in terms of the objectives of urban evaluations. Some, particularly the proponents of qualitative evaluation, argue that it is not possible to measure impact or causation in any statistical sense and that the objectives of urban evaluations should be to understand project processes and the subjective reactions of beneficiaries. On the other hand many of the more quantitatively inclined researchers argue that policy makers need to know as much as possible about project impacts and the effectiveness of different components and services, and that the function of the evaluator is to introduce as much statistical rigour as possible into the complex urban context.

In order to understand the associations between project inputs and impacts on the target population, a number of quasi-experimental designs have been developed (see Box 4-4 for a typical example) which try to approximate the experimental paradigms closely as possible, within the realities of the urban context. Annex E summarizes some of the most common of these designs and uses the experimental paradigm as a yardstick against which to evaluate their strengths and weaknesses. This enables the researcher to identify the weaknesses and potential biases in each evaluation design. Other techniques can then be used as consistency checks and to compensate for some of these weaknesses.

There are a number of ways in which a badly designed evaluation can lead to wrong interpretations being made about project impacts or potential replicability. These have been classified into 4 sets of "threats to validity. Statistical conclusion validity arises when the sample is wrongly designed or an inappropriate statistical test is used. When this happens a project may appear to have no impact when in fact the sample was too small, or wrongly stratified, to have been able to detect impacts if they did exist. Internal validity problems arise when a statistical association (such as a

BOX 4-4 EXAMPLE OF A PRE-TEST POST-TEST QUASI-EXPERIMENTAL DESIGN WITH A CONTROL GROUP

Evaluating the effects of housing investments on food expenditures in El Salvador

A concern of housing planners is that the need of poor families to mobilize resources for purchasing or upgrading their housing may result in reduced expenditures on food and other basic necessities. This hypothesis was tested on a sites and services project in El Salvador by comparing food expenditures of participants and a control group before the project began and when the project had been underway for around 18 months. The results were as follows:

	Average monthly expenditures on food (pesos)		
	1977	1979	% change
Participants	177.4	240.6	+ 35.6
Control group	155.2	241.5	+ 64.2

The T-Test score of 4.63 (with 356 degrees of freedom) was statistically significant at the 0.0005 level, thus supporting the hypothesis that project participation would negatively affect food expenditures. The control group was essential to this design, to provide information on general trends in food expenditure and to control for the effects of inflation. A weakness of the design is that it is not possible to determine the extent to which the slower rate of increase in food expenditures is due to the initial differences between the two groups rather than to the effects of participating in the project.

Source: Bamberger "Statistical procedures for the evaluation of project impact." 1982.

correlation coefficient) is wrongly interpreted as proving that a causal relationship exists. Problems of construct validity of causal relations relate to the fact that different researchers may place different interpretations on the statistical findings. In the earlier example of an increase in indicators of middle class consumption patterns, one researcher may interpret this as showing that poorer families have been bought out and have left the community; whereas another may interpret the same data as proving that poorer households have remained in the project but have adopted middle class consumption patterns. Finally, external validity problems arise when generalizations have to be made from the results of the project about how a similar program would perform on a larger scale. It may be that the first project has attracted the most competent or ambitious households, and that if the program were applied on a larger scale it would probably produce less satisfactory results due to the lower potential of new participants.

4. The use of panel and independent samples

Most discussions of quasi-experimental designs assume that it will be possible to interview the same group of subjects before and after the project. Unfortunately there is a high population turnover rate in many of the areas studied, particularly among the control groups, with the result that in practice it is often not possible to re-interview the same households. Many evaluations try to ignore this problem, which can lead to serious errors in the interpretation of the findings as families who have not moved usually have quite different characteristics from movers. For example, the movers may be younger, have more education and perhaps higher income. Clearly an analysis based only on those households who have remained, will give a very misleading picture of what has happened to all families studied at the start of the project. To overcome this problem 3 alternative research designs can be used:

a. Panel design (sometimes called "related sample design") in which original households or subjects are reinterviewed. If a household has moved, it will not be replaced and the sample will become smaller. The households who are reinterviewed are a representative sample of all households who have not moved, but they are not representative of all orginal households. This design can be strengthened by comparing the characteristics in T(1) of households who moved with those who stayed. This design can be

used to estimate project impact on original households or subjects, but care must be taken in the types of inferences which are made. The advantage of this design is that "before" and "after" scores are available on the same subjects so that more precise statistical estimates can be made of project impact. This means that the statistical error arising from the chance differences in age, education etc of the two groups of subjects ("between group error") is eliminated and the level of statistical precision is consequently increased.

b. Independent samples in which a new sample is selected for the second interview. This design has the advantage that the new sample is representative of the total population being studied so that the errors of interpretation related to the previous design do not occur. The disadvantage is that the statistical analysis is more complex, and as the estimates are less precise larger samples may be required. Despite the analytical difficulties it is possible to make reliable estimates with this design.

c. Mixed sample design in which aspects of the 2 previous designs are combined. With this design a sample of structures, businesses, etc is selected in T(1). For the design of the sample in T(2) the same structures, businesses etc are revisited. If the same household or subject is still there he/she will be reinterviewed. If the subject has moved a new subject will be randomly selected. This can be considered as a stratified sample, where one stratum is a random sample of original subjects, and the other stratum is a random sample of new subjects. When the two strata are combined, this produces an approximately random sample of all subjects in T(2). Separate analysis can be conducted of: original subjects (panel sample), new subjects and all subjects. In order to be able to conduct all 3 types of analysis it is important to ensure that the sample design will include sufficient original and new subjects. Care must also be taken with respect to the procedures for selecting the new subjects.

5. Sample design

The principal objective in the design of a sample survey is to ensure that valid inferences can be made about the characteristics of a certain population. If the sample is too small, or if there are biases in the selection procedures, there is a danger of making incorrect inferences about the population. Some of the requirements for a good sample are the following:

a. The sample must cover the entire project population. Many samples inadvertently exclude groups such as renters, female headed households, the unemployed, artesans who are not registered with a certain organization etc.

b. The sample must be selected randomly so that the results can be generalized to the total population. A common error is to exclude people who refuse to cooperate or who are never at home.

c. The sample must be sufficiently large to ensure that the findings are statistically valid. Sometimes the original sample size is reduced to save money with the result that the number of cases is too small for valid statistical analysis. Another common mistake is to assume that if the sample covers 10% of the population this will be adequate. In fact the main determinant of sample size is the absolute number of interviews and in most cases the proportion of the population covered does not guarantee that the sample is sufficiently large.

d. The sample design must ensure the adequate representation of sub-groups of interest to project management. A simple random sample will frequently under-represent certain groups which are numerically not very large but which are of particular interest to management. It is usually necessary to use a stratified sample to ensure that sufficient cases are included from these groups.

e. One or more control groups should be included so as to separate the effect of the project from other extraneous factors.

In addition care should be taken to minimize costs by ensuring that the sample is not larger than necessary. Annex D describes some of the procedures which can be used to meet these requirements. A number of standard procedures exist for estimating sample size. Where possible a statistician should be consulted as the study may be almost worthless if the sample is wrongly designed.

6. Analysis and interpretation of the survey data (See Annex F)

This section presents a brief summary of the analytical procedures described in Annex F. It is strongly recommended that all of the statistical techniques are complemented by one or more qualitative procedures so as to ensure consistency checks and to help in the interpretation of the findings. The following procedures are presented in order of increasing complexity:

a. Tests of association and difference (Annex F Section 2)

The simplest type of statistical tests compare participants before and after the project with respect to variables such as household income, parasitic load or transport expenditures; or they compare participants and a control group after the project has been underway for some time. These designs are very weak and can easily lead to misinterpretation as there is no way to determine whether the observed differences are related to the project.

A more powerful design is to compare participants and a control group before and after the project (See Box 4-4). This design still suffers from the problem that the evaluator normally has no control over the selection of participants. Consequently there will frequently be important differences between the experimental and control groups at the time the project begins. When differences are found between the two groups after the project it is difficult to assess whether the differences are associated with the project or are due to the continuing influence of the initial differences between the groups (see Table F-1 in Annex F for an example). The type of statistical test to be used will depend upon whether a panel or an independent sample design is used.

b. Use of regression analysis to statistically match the control and experimental groups (Annex F Section 4)

A constant problem in the interpretation of the results of the previous types of quasi-experimental designs is the confounding effect of the initial differences in age, income, education, family size etc of the control and experimental groups. If these intervening variables are included in a multiple regression analysis, it is possible to statistically match the experimental and control groups on each variable so that there effect is "controlled". An additional dummy variable is then included to indicate whether the family was part of the control or the experimental group (See Box 4-5). If the coefficient of the dummy variable is statistically significant this means that the value of the dependent variable is signficantly different for the two groups after controlling for the effects of the intervening variables. This provides much stronger evidence of a real association between the project and the outcome variable.

The form of the analysis will be determined by whether related (panel) or independent samples are used. In the former case a single regression analysis is conducted in which the value of the dependent variable

(impact being estimated) in T(2) is regressed on a set of variables, including the value of the dependent variable in T(1). If the dummy variable for project participation is significant, this suggests that the project has affected the outcome. However, if independent samples are used, it is necessary to conduct separate analyses for the data from T(1) and T(2). The coefficients for the dummy variable are compared, and if they are significantly different, then it is assumed that the project has affected the outcome.

c. Use of regression analysis to assess the relative contribution of different project components (Annex F Section 4)

Many projects, such as upgrading, involve a large number of different components (water, roads, housing credit, technical assistance etc) which households may receive in different combinations. Regression analysis can be used to evaluate the relative contribution of each of these components to overall impact. Variables are included in the regression to measure

BOX 4-5 THE USE OF MULTIPLE REGRESSION TO STATISTICALLY MATCH EXPERIMENTAL AND CONTROL GROUPS.

Evaluating the impact of participation in a housing project in El Salvador on household income.

Table 2 in Annex F illustrates how multiple regression can be used to estimate the impact of participation in a housing project on household income. A simple statistical comparison of the control and experimental groups (using the T-Test) found that income had increased more rapidly for participants with the implication that participation in the project may have caused the difference. However, multiple regression was used in a second stage of the analysis, and participants and the control group were matched with respect to family size, age and education of the household head and household income at the start of the project. The coefficient for the dummy variable "Participation Status" was found not to be statistically significant (there was a 10% probability of the difference between the two groups having occurred by chance). This showed that the apparent effect of the project was in fact due to the initial differences between the two groups.

household access to each project service. The coefficient of each of the project component variables estimates the contribution of each component when all other variables are held constant. For example the coefficient for water in Table F-5 of Annex F is 0.2 which means that when all other components are held constant, each additional liter of water consumed will be associated with an increase of 0.2 pesos of housing investment.

This technique is potentially very powerful as it permits a cost/effectiveness analysis in which the cost of providing each component can be compared with the benefits produced.

d. Interrupted time series analysis

Where time series data is available for a period beginning before the project started and continuing through the life of the project, it is possible to compare the trend (slope) before the project began and after it has started to determine whether a "jump" can be detected. This can either be done through simple observation, or if enough observation points are available, through the use of regression analysis.

e. Path analysis to describe the causal process through which impacts are produced (Annex F Section 5)

Path analysis consists of developing a diagram to model the causal process through which it is expected that impacts will be produced. Regression coefficients are obtained for each path, and in this way it is possible to evaluate the direct and indirect contribution of each component to the production of the final outcomes. Critics of path analysis stress that it assumes a recursive model in which all causality is one way, and where there is no interaction between different components. Although this is true path analysis is still a very useful first approximation and has the great advantage over more complex models that the analysis is very easy to conduct and intuitively simple for people who are not statisticians to understand.

f. Hedonic price analysis

Annex F describes ways in which Hedonic Price Analysis can be applied to project evaluation. This can be used in the evaluation of shelter projects to estimate the benefits produced by each project component.

F. COMPARING THE EFFECTIVENESS OF DIFFERENT PROJECTS

One of the limitations of net impact evaluation for the policymaker is that it estimates benefits without regard to their cost. It is implicitly assumed that any project which produces the desired benefits should be financed. This approach does not directly help the program planners whose main concern is how to choose between alternatives so as to maximize social utility with limited resources. Three approaches will be described which can provide an objective basis for choosing between alternative projects: cost benefit analysis; cost-effectiveness analysis and cost-utility analysis. The reader is referred to Levin, 1984 for a more comprehensive discussion of the three approaches. Table 4.3 summarizes the main stages in the design and application of each approach. Each approach implies a comparison of costs with a measure of effectiveness (benefits, output or subjective utility) and before describing each technique we will discuss some of the general issues involved in the measurement of costs and effectiveness. It is important to appreciate that the purpose of these techniques is to provide guidance to the decision maker, but they should never be used as the only factor on which decisions are based.

1. Issues in the definition and measurement of costs

a. It is important to define whether costs are estimated from the point of view of society (social or opportunity costs) or from the point of view of a particular organization or group. For example, the time of a volunteer or the value of land donated by another agency, are not costs for the implementing agency, although they do represent a cost for the volunteer or the agency donating the land. The concept of social or opportunity costs is important, as a project which is cost-effective for the implementing agency may have a high social opportunity cost and thus be less desirable from the social point of view than another project.

b. The issue of costs which cannot be monetized is frequently a serious problem. For example, it is difficult to measure the impacts of an urban development project on health, or on the quality of the environment. Costs which cannot be monetized are frequently ignored with the result that some projects can appear more attractive than they really are.

c. Costs occur at different points in the project cycle so that it is necessary to use a discounting procedure to estimate the present value of all costs. The choice of discount rate, which is a policy rather than a

Table 4.3 Main stages in the 3 approaches to effectiveness analysis

	Cost Benefit Analysis	Cost-Effectiveness Analysis	Cost-Utility Analysis
1. Define the discount rate	1	1	1
2. Define all costs			
Identify project "ingredients"	2	2	2
Define who pays each cost	3	3	3
Compute shadow price (opportunity cost)	4		
Estimate cost streams by year	5	4	4
Define policy for non-monetized costs	6	5	5
Compute present value of each cost (PVC)		6	6
3. Define benefits/outputs/utility			
Identify all benefits	7		
Decide which benefits can be monetized	8		
Estimate benefit streams by year	9		
Monetize benefits	10		
Compute net benefits by year	11		
Define main outputs of project		7	7
Decide which output to use		8	
Define how outputs will be scaled (O)		9	8
Assign subjective weights to each output (W)			9
Estimate probability of occurence (P)			10
Compute utility index (U)			11
4. Compute effectiveness indicators			
Compute present value of net benefits	12		
Compute internal rate of return	13		
Assign social weights and compute social analysis	14		
Conduct sensitivity analysis where required	15		
Compute cost-effectiveness indicator as PVC/O		10	
Compute cost-utility index as PVC/U			12

technical decision, can have a significant effect on the results of the analysis.

d. The cost/effectiveness analysis is frequently conducted on a pilot project to help management decide which option to chose for a larger future project. It is difficult to generalize from a pilot project as costs are likely to be significantly different from the costs of replication on a larger scale. On the one hand the pilot project may be relatively more expensive as it must include research and development costs which would not have to be repeated. On the other hand a pilot project may appear to be relatively cheaper as there may be a number of social costs not covered by the executing agency. Other agencies may provide free services such as architects or lawyers, and there may also be volunteers the value of whose labor is not imputed to the project. Land or buildings may also be provided below cost, particularly if the government wishes to use this as a show project. Finally there is the problem that many of the factors of production (land, materials and some types of staff) are in short supply so that their cost is likely to increase sharply if the scale of operations increases. For all of these reasons it is difficult to use the analysis of the pilot project to make inferences about the relative cost-effectiveness of the different options when they are implemented on a larger scale.

2. Issues in the definition and measurement of benefits and effectiveness.

a. It is important to clarify the difference between outputs and impacts. Outputs are usually the immediate products or objectives of the project (provision of credit, equiping of classrooms and provision of teachers, meetings with community groups etc). These are intended to produce certain benefits or impacts on the target population (for example, generation of employment, improvement of reading skills, greater social cohesion etc). The fact that the outputs have been produced does not necessarily mean that the project will be successful in producing the intended benefits. As cost-effectiveness and cost-utility analysis both use measures of outputs, it is important to keep the distinction between outputs and impacts clearly in mind.

b. Cost-benefit analysis requires that benefits be monetized so that benefits and costs of different projects can be compared. Two sets of problems constantly occur. First, many benefits are difficult to monetize and are either ignored or are included in a very arbitrary way. Benefits such as

improved health, improved environmental conditions, psychological satisfaction, reduced crime or violence are frequently left out of the calculations, with the result that total benefits may be significantly underestimated. Second, economists frequently use imputed rent as a proxy for total benefits. This is based upon a set of assumptions about perfect consumer knowledge and free markets which are usually far from true. In many cases it is also difficult to obtain a reliable estimate of rent as projects operate in areas where there is not a free housing market. For all of these reasons there is frequently a much greater margin of error in the estimation of benefits than in the estimation of costs, so that the results of the analysis may be seriously distorted.

c. There are also a number of problems in the measurement of ouptuts. First is the problem of scalability. The way in which cost-effectiveness is estimated usually assumes that outputs are measured on an interval scale (for example that an increase from 5 to 10 points on a reading aptitude test, is the same as an increase from 50 to 55 points). Many outputs are in fact measured on an ordinal scale where these assumptions about equal intervals are not valid. Second, many projects have multiple outputs and there is a problem of deciding which output or set of outputs to measure.

3. Cost benefit analysis(described in Annex F)

All costs and benefits are identified and their incidence for each year of the project life plotted. Table 4.4 gives a simple worksheet which can be used with each of the 3 approaches to identify and assign costs. A monetary value must then be assigned to each cost and benefit, so that net benefits can be computed for each year. A crucial step is then to decide the discount rate to be used to estimate the present value of the net benefit stream. The choice of discount rate can significantly affect the results of the analysis so care must be taken in deciding the appropriate rate. The economic efficiency of each project is then defined either as the Net Present Value (discounted net benefits) or as the Economic Rate of Return (the discount rate which makes net benefits equal to zero). If a number of projects are being compared, the project with the highest ERR or NPV will normally be recommended. If only one project is being studied, it is considered economically viable if the ERR is higher than the selected discount rate.

Table 4.4 Cost Worksheet for Cost-Effectiveness Analysis

Item	Accounting Cost	Cost to client	Cost to other public agencies	Cost to other group	Shadow cost	Cost stream over project
Personnel						
Land						
Rent, purchase or construction of buildings						
Urban infrastructure						
Materials and equipment						
Loan funds						
Contracting professional services						
Other (specify)						
Total cost (Present value)						

Adapted from: Henry Levin "Cost-Effectiveness Analysis in Evaluation Research (1975)

The analysis can be refined by using social weights to assess the distribution of benefits between different population groups (rich and poor, young and old etc) or between different types of expenditures(savings and consumption). It is also possible to use sensitivity analysis to assess the effect on ERR of changing the way in which costs or benefits are measured.

Where it can be used, Cost Benefit Analysis (CBA) is a very powerful analytical tool. One of its greatest values is as a teaching device which forces planners to think more carefully about the assumptions on which the project is based and on how they define the costs and the benefits. One of the most common problems is that important benefits, and sometimes important costs, are excluded as they cannot be monetized. This can lead to spurious precision and very misleading results. In general it is more difficult to estimate benefits than costs, and the outcomes of the analysis can vary greatly depending on the set of assumptions used to measure benefits. Because of the problems of measuring benefits CBA has had very limited application in many social fields such as health, community development and education.

Another serious problem relates to the interpretation of the findings. The cost assumptions used for the pilot project will frequently not hold true for a much larger project so that the pilot project with the highest ERR may not be the best option to choose for larger scale replication. CBA methodology has difficulty in coping with this problem of replication.

A final problem is that the methodology assumes a perfect market and perfect knowledge on the part of consumers. In the analysis of many types of social programs, these assumptions are not valid and this can seriously affect the application and interpretation of the analysis.

4. Cost-effectiveness analysis

The crucial step in this analysis is the definition and measurement of the effectiveness indicator (output). It is normally only possible to use one indicator so it must be selected carefully. Ideally it should be measurable on an equal interval scale. The present value of costs is computed in the same way as for CBA, and cost-effectiveness ratios are computed by dividing costs by the amount of output produced. It is assumed that the option with the highest effectiveness ratio is the one which should be selected.

One of the problems with this approach is the comparability of the options being studied. Most projects produce a number of different outputs, and to compare them in terms of only one output can be misleading. Scalability is another serious problem. For example, televison might produce larger gains in reading scores for average and superior children, but the classroom situation might be more effective for below average children. If the project is concerned about backward children, it might be necessary to give a higher weight to improvements in reading scores at the lower end of the scale. The problems of scale and the difficulties of generalizing from small pilot projects also exist with this approach.

5. Cost-utility analysis

Costs are computed in the same way as for the previous approaches, but the method for estimating utility is different. All major expected outputs are identified for each project being compared. Managers and policy makers are asked to rate each outcome in terms of its importance on a scale of (for example) 1 to 10. The weight assigned to each output is the mean of all the ratings. In some cases the managers and policy makers will also be asked to estimate the probability of the outcome being achieved or the amount of improvement which is expected (for example the probability that households will install a pit latrine) The utility score for each output is the weight multiplied by the probability. Some advocates of this method recommend combining the utility scores for each output so as to obtain a total utility score for each project. The Cost-Utility ratio is then computed as cost divided by utility.

The main advantage of this approach is that it provides a more rational way for planners to select between projects which have several outputs. The problem is that the concept of utility is by definition subjective and difficult to measure, so that the effects of multiplying and adding utility scores may produce a very misleading result. The danger is that some policy makers may assign more credibility to the cost-utility coefficients than they deserve.

G. CHOOSING THE APPROPRIATE STRATEGY FOR IMPACT EVALUATION

The purpose of an impact evaluation is to assist policy makers in the selection of future projects and the definition of investment

strategies. In order to select the appropriate type of impact evaluation it is therefore essential to understand the future decisions which have to be made and the types of information on which these decisions will be based. The following are some of the key factors which affect the choice of impact evaluation design:

1. What are the future options currently being considered and what are the key factors which will affect the decision.

The following are some of future policy options which may be under consideration, each of which has different implications for the evaluation design:

a. A decision is to be made whether to replicate the pilot project on a larger scale or to discontinue it.

b. The project will be compared with other options in terms of its ability to achieve one or more clearly defined objectives. The "best" option will be selected for replication in a second project.

c. A decision has already been made to continue the project and the key issues relate to improving efficiency and effectiveness and ensuring the benefits reach the target population.

In addition to knowing what types of decisions have to be made, it is also necessary to know some of the key factors on which the decision will be based. For example, how important is the successful production of impacts in determining whether a project should continue? How important is the cost-effectiveness of the project in producing certain types of outputs? How important are political decisions relating to, for example, the distribution of the programs throughout the country, or to the accessibility of the project to certain population groups? All of these factors can affect the choice of evaluation design. If it is essential to rigorously measure the extent to which certain impacts (increased income or employment) have been achieved then one of the designs discussed in Section 5 may be appropriate. If, on the other hand, the concern is mainly to assess the ability of the project to produce certain outputs (training courses, administration of credit) at an economical cost, then cost-effectiveness may be more appropriate. In those cases where a number of different projects are being compared in terms of their economic impacts on a target population, cost benefit analysis may be an option to consider.

2. Time scale

Many evaluations are started when a project is nearing completion which obviously limits the types of design which can be used. It is also common for management to require the results within a very short period of time (say 3 months) which leaves very little option other than to use one of the rapid evaluation approaches discussed in Section D.

3. Available resources

Many types of impact evaluation are relatively expensive, so that where only a small evaluation budget is available, the options will again be limited.

4. Scale and complexity of the project

The larger and more complex the project, the greater the need for a rigorous statistical analysis. Whereas the likely impacts of a small project with limited geographical distribution can often be assessed with a simple and rapid study; this is usually not possible when the project involves a number of different components and affects tens of thousands of beneficiaries with a wide geographical dispersion.

5. The nature of the project and the complexity of its objectives.

Some projects are mainly concerned to produce a limited number of easily measurable outputs (for example reading improvement or small business credit). On the other hand, other projects seek to produce a broad range of social and economic changes (for example a squatter relocation program). In the latter case it is more important to include a qualitative component in the evaluation to ensure that the complex social and psychological effects of the project are well understood.

6. The need for a multi-method approach

The need for a multi-method approach is as important for impact evaluation as for any other type of study. Many of the approaches use relatively sophisticated types of analysis (multiple regression, economic rates of return, cost-utility ratios etc) all of which are based on the assumption that impacts, benefits, outputs or utility have been accurately measured. In most cases the indicators are extremely difficult to measure and

estimates can vary widely according to the definitions and measurement techniques used. It is strongly recommended that all key variables are measured by at least 2 independent methods. It is also recommended that statistical analysis always be complemented by qualitative techniques in order to understand the meaning of the statistical results and their implications for the target populations affected by the project.

CHAPTER 5: MANAGING THE EVALUATION

This Chapter discusses the main issues involved in defining the appropriate organizational structure for monitoring and evaluation at the level of the implementing agencies, the local coordinating agency, specialized sectoral agencies and national development and financial agencies. Some of the issues discussed include: defining who should conduct the evaluation, the role of consultants, the appropriate organizational location of the monitoring and evaluation units, the role of a steering committee, and the distribution of evaluation responsibility between the national, sectoral and local agencies. There is no best organizational structure and the decision in each case is determined by the scope and complexity of the project and the relative size and research experience of the different agencies involved. Guidelines are provided for estimating the likely resource requirements for different types of monitoring and evaluation programs. The final section discusses common problems in the organization of an evaluation program and some of the possible solutions.

A. THE MONITORING AND EVALUATION FUNCTIONS AND NEEDS OF DIFFERENT ORGANIZATIONS

Figure 5.1 illustrates the main organizations involved in the monitoring and evaluation of a typical urban development project:

1. Most urban development projects come under the responsibility of a government ministry (Planning, Interior, Urban Development) which is responsible for coordination and supervision. In some cases the ministry may be a passive recipient of monitoring data, but in others it may be actively involved in designing and helping implement a wide range of monitoring and evaluation studies.

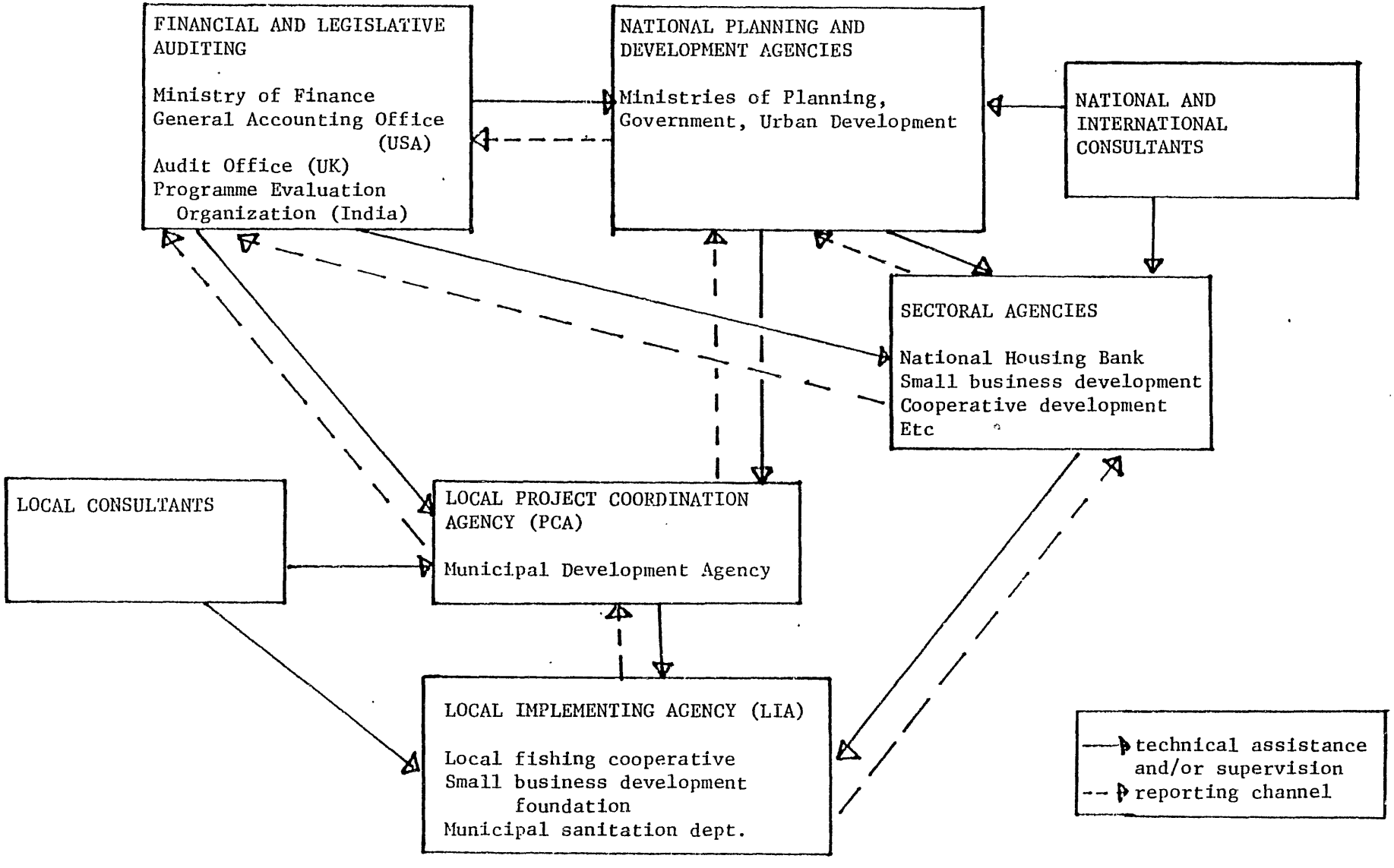
2. In all countries there is a government agency (usually the ministry of finance) responsible for monitoring financial performance of all projects. Monitoring is often limited to standard auditing but in some cases the cost-effectiveness techniques described in Chapter 4 are used. In a number of developing countries the legislative branch is beginning to create its own watchdog agencies such as the Programme Evaluation Organizations in India (see World Development Report, 1983). As pressure grows to achieve maximum utility from scarce resources, cost-effectiveness will become an increasingly important tool for project monitoring.

3. Many projects involve a large number of specialized agencies in areas such as housing, low-cost sanitation, small business development, community health, transport and cooperatives, each with some monitoring and evaluation responsibilities.

4. Many government agencies subcontract all or part of their evaluations to national or international consultants, and when interest in evaluation grows there will usually be a corresponding increase in the number of consulting organizations.

5. Multi-city urban projects are usually coordinated and managed by a local project coordinating agency (PCA) such as a municipal development corporation which, in a large project, may be responsible for monitoring several hundred project components Once the performance monitoring system is

Fig 5.1 PRINCIPAL NATIONAL, SECTORAL AND LOCAL AGENCIES INVOLVED IN THE MONITORING AND EVALUATION OF A TYPICAL INTEGRATED URBAN DEVELOPMENT PROJECT



in place, PCA management becomes interested in receiving rapid feedback on problems arising during the implementation process.

6. The execution of the project is the responsibility of a series of local implementing agencies (LIA) such as a small business development foundation, the municipal sanitation department, local office of the national housing bank etc. These agencies are required to provide monitoring reports to the coordinating agency (PCA) and often to the appropriate sectoral agency.

7. Finally use is often made of locally hired consultants who assist in the design of the monitoring system or who conduct evaluation studies for the local implementing or coordination agencies.

The following sections will discuss the issues and options involved in the design and organization of monitoring and evaluation for each of these groups. However, before beginning this discussion, the question must be asked as to whether it would be possible to centralize monitoring and evaluation in one specialized agency rather than having so many different organizations involved. There are several factors which make it difficult for monitoring and evaluation to be completely centralized:

a. Each agency requires information it can trust and which it knows has been collected to respond to its own particular needs. The cooperation required to study delicate issues such as interagency conflicts, poor organizational performance and lack of community support, will only be given if the agency is able to control how the information will be used and disseminated.

b. Experience has proved that there is no such thing as a comprehensive objective study which can cover all issues in an unbiased way. A study must be designed for a specific purpose if it is to provide the information needed by a particular client. A general purpose study usually does not fully satisfy the needs of any of the prospective clients as the right information has not been collected and the right analysis prepared.

c. There are frequently potential areas of conflict between agencies due to competition for resources or different political allegiancies. Under these circumstances, each agency seeks to control access to information on its performance.

d. Finally, the evaluation program must have the speed and flexibility to respond rapidly to the information needs of the organization it services. In practice this cannot be done, unless the evaluation unit is directly controlled by the project manager.

B. ORGANIZING MONITORING AND EVALUATION AT THE PROJECT IMPLEMENTATION LEVEL

Urban projects usually comprise a number of different components such as sites and services, maternal and child health or artesan credit, each of which is implemented by a different project implementing agency (PIA). This section discusses how the monitoring and evaluation of each component should be organized.

1. The use of consultants

(Note: the following discussion is equally applicable to the use of consultants by other types of agencies)

a. General guidelines on the use of consultants

The manager must decide which parts of the monitoring and evaluation should be conducted in-house, and which should be subcontracted either to local consultants or to another government agency. It is recommended that Performance and Implementation Monitoring should normally be conducted in-house, although this does not preclude some involvement of consultants.

The use of consultants brings a number of advantages, including access to specialized research skills, a greater degree of objectivity, and greater flexibility in the employment of staff and the use of financial resources. From the administrative point of view, the use of consultants also has the great advantage of avoiding the creation of a large permanent staff. The benefits are discussed in more detail below.

However, the use of consultants also has a number of potential disadvantages (see Box 5-1). Outsiders are often perceived as a threat and it is difficult for external consultants to establish the rapport

and daily contact with the project staff which can be achieved by an internal unit. An external evaluation is usually much more expensive, and does not develop an internal research capability, hence creating a continuing need for outside consultants in future projects. Finally, as consultant contracts usually specify very precisely the scope and duration of work, many consultants do not have the flexibility to adapt to changing management information needs (Box 5-2).

BOX 5-1 TYPICAL PROBLEMS WHEN MONITORING AND EVALUATION IS SUBCONTRACTED TO OUTSIDE CONSULTANTS

In one of the early World Bank urban projects, the design and implementation of the monitoring and evaluation system was subcontracted to a newly formed group of local consultants. The quality of many of the reports was technically very good, but in most cases had very little operational impact for the following reasons:

a) The executing and coordinating agencies had no mechanisms for reviewing the reports so that long delays occurred before the recommendations were acted on.

b) The government agencies did not have an experienced researcher who could coordinate with the consultants and interpret their work for the executing agencies.

c) To avoid controversy the consultants avoided making direct recommendations or drawing attention to delicate issues. For example, they left it to the reader to assess the policy implications of the high level of illegal subletting which was detected.

d) The reports were not presented in a very operational way and contained more detail than the average manager was able to absorb.

e) The consulting group, which was quite small, encountered personal and financial problems which resulted in them losing some of their key staff. Due to this they were not able to complete their contractual obligations or to prepare a final report.

f) None of the staff of the executing or planning agencies received any training or guidance in evaluation techniques from the consultants, and 5 years after the consulting contract began, it was still necessary to look to private consultants to conduct even the most basic evaluation studies.

The following are some of the ways in which consultants can be used within a general policy of developing in-house evaluation capability:

i. A full or part-time adviser may be contracted, often as part of the technical assistance package provided by a lending agency or donor. It is important to ensure that the advisor does not become the de facto director of the evaluation unit. An advisor may impose his or her own research scheme, which, although technically sound, does not respond to the

BOX 5-2 PROBLEMS WHICH CAN ARISE WHEN A LONGITUDINAL IMPACT EVALUATION IS SUBCONTRACTED TO OUTSIDE CONSULTANTS

A local university group was contracted to design and implement an impact evaluation for a sites and services project. A relatively complex design was developed and discussed with the executing agency. After the initial design had been approved the executing agency suggested that a number of additional questions be included on beneficiary preferences for design features such as the size and location of windows. The consultants declined to include these questions, firstly because the purpose of the impact evaluation was not to study the details of project design, but more importantly because they felt the "objectivity" of the evaluation would be compromised if the executing agency was involved in the design of the study. After the application of the baseline survey the executing agency contacted the consultants on several occasions to ask for help on operational issues, but were advised that this was not included within the terms of the contract. When the follow-up survey was conducted, consultants complained that the validity of the evaluation design had been seriously affected by delays in project implementation and by changes in the communities covered. When the reports were received, management felt they were too academic and that they did not provide any useful guidance on policy issues or the design of future projects.

The lessons to be drawn from this experience are that it is necessary to define clearly the objectives of an evaluation, and that closer contact must be maintained between consultants and the executing agency. Management was expecting a process evaluation to help with implementation, whereas consultants understood that their function was to evaluate long-term impacts.

needs of the organization or the technical capabilities of the evaluation unit. The ability and willingness to train local staff should be an important factor in the selection of the adviser. Box 5-3 gives an example of the use of long-term consultants in an IDRC-World Bank sponsored 3 country evaluation.

ii. Consultants may sometimes be contracted at the beginning of an evaluation to assist in the design of the instruments and the systems for data analysis.

BOX 5-3 AN EXAMPLE OF THE USE OF LONG-TERM RESIDENT CONSULTANTS TO ASSIST IN THE DESIGN AND IMPLEMENTATION OF THE EVALUATION

The IDRC-World Bank evaluation of urban shelter programs

In 1975 the IDRC and World Bank agreed to finance a 5 year evaluation of 3 of the first Bank fianced urban shelter programs. In each country (Zambia, Senegal and El Salvador) an evaluation unit was established within the project executing agency, and an expatriate consultant was contracted to be resident advisor for a period of between 1 and 3 years. Some of the consequences of the long-term involvement of the advisors were the following:

(a) More sophisticated evaluation designs were used than would have been developed by local researchers.

(b) The initial evaluation designs responded more closely to the interests of the World Bank than to those of the executing agencies. In particular more emphasis was placed on longitudinal impact studies and less on process evaluation.

(c) The presence of the advisor provided easier access to senior management and gave the evaluation more prestige and influence than it would have had under the direction of a less trained local researcher.

(d) There was a feeling, in some, but not all, of the countries, that the evaluation unit was reporting to the World Bank and not to project management, and in one case this resulted in considerable distrust and marginalization of the evaluation.

(e) The advisors were able to provide a more rigorous training to local researchers than would have been possible through periodic visits. This also had considerable multiplier effects as the advisors became involved with local universities and in providing assistance to a wide range of national organizations.

(f) The advisor was able to act as an intermediary between the executing agency and the World Bank on a number of operational and research issues.

iii. A consultant may be contracted to provide technical assistance on specific topics such as sample design, data analysis or report writing. The objective should again be for the consultant to help train local staff to take over these activities.

iv. Consultants may assist in the selection and training of evaluation staff.

v. Specific studies may be sub-contracted. Examples of such studies include:

- Studies which require special methodologies such as sophisticated economic or financial analysis.
- Studies which require the conducting of large sample surveys which are beyond the experience or resources of the evaluation unit.
- In-depth anthropological studies of particular communities.
- Studies intended to provide an independent "objective" perspective on the project (see following section).

The amount of outside consulting should be kept to a minimum until the evaluation unit has had time to establish itself. If a new and inexperienced evaluation director has a more experienced adviser looking over her/his shoulder, the director may never gain sufficient confidence to develop her/his own research program and priorities. The best approach in the early stages may be to have occasional assistance on specific topics, either through periodic visits from an expatriate or through occasional (say monthly) meetings with a local consultant. The consultant should provide advice and assistance on an established evaluation program rather than being the person who determines the program itself.

b. <u>Using consultants to provide an independent ("objective") perspective</u>

One of the potential dangers of an in-house evaluation is a loss of objectivity. Organizations tend to perceive the world in terms of their own objectives, and although they may conduct very rigorous analysis of how well these objectives are achieved, they may never question their appropriateness. For example, a shelter project may rigorously evaluate the

speed and cost of production of shelter units, but may never question the value of these units as an effective way to improve the welfare of low-income households. An important function of outside consultants is to evaluate the objectives and underlying assumptions as well as the operational procedures.

Another role of the consultant is to help assess the actual and potential impacts of the project on the low-income population and on local and national development policies. Project management is frequently so involved in the day to day project activities, that they find it difficult to stand back and review potential replicability. For example, a special project unit may be set up to ensure that the pilot project is successful. This unit may create antagonisms (because of higher salaries or priority attention) which may make it difficult for the project to be replicated. An outside consultant will often be able to see this conflict more clearly than project management. Organizations often enter into alliances and conflicts with other agencies and groups, and it is easier for an outsider to assess the implications for project implementation.

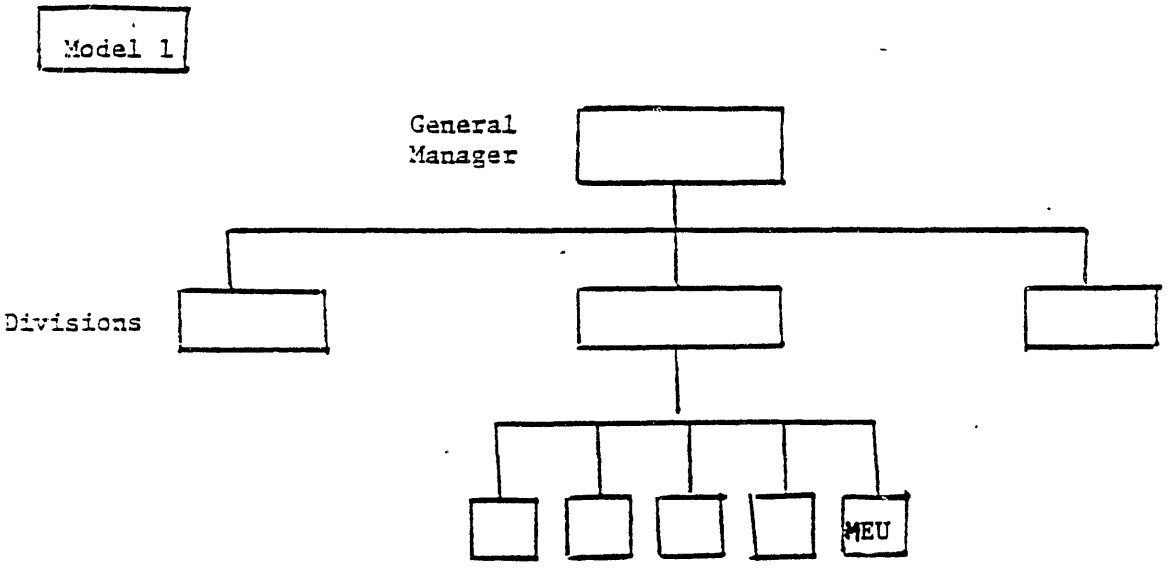
The consultant can also serve as a communication link between implementing agencies and project beneficiaries. Managers frequently receive limited and distorted feedback on beneficiary reactions often through the members of a small and unrepresentative community junta, or from project technical staff who see their job as selling the community on an already designed project rather than listening to the community. Managers frequently react defensively to any criticism and may discourage feedback through bureaucratic procedures which make beneficiaries feel uncomfortable (for example all meetings may take place in the project office in the city center). The consultant is not subject to these organizational constraints and is able to listen to and understand the point of view of all of the main sectors of the affected populations and communicate these views to management.

2. Locating the evaluation unit in the organization

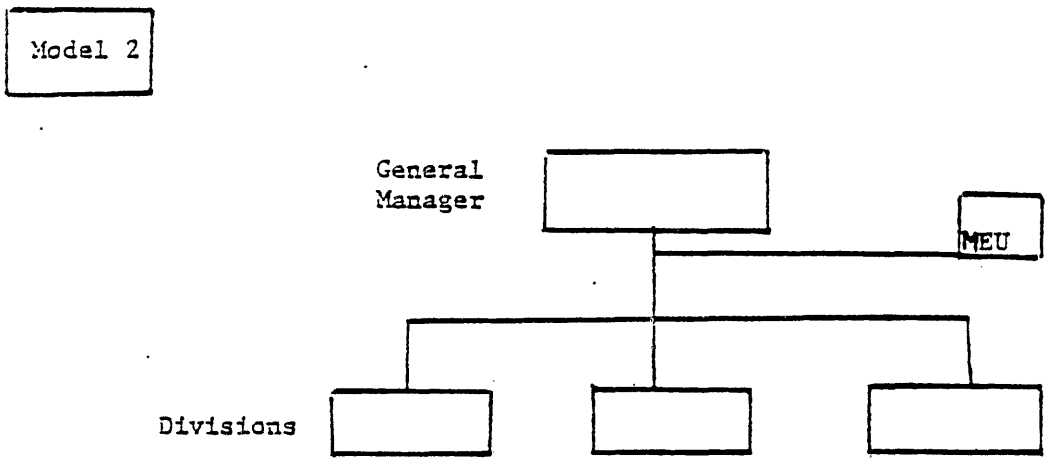
a. Alternative organizational structures

Figure 5-2 presents 4 organizational options for monitoring and evaluation. In Model 1, there is a single monitoring and evaluation Unit

Fig 5-2: Alternative organizational structures for the monitoring and evaluation of a one city project implementing agency

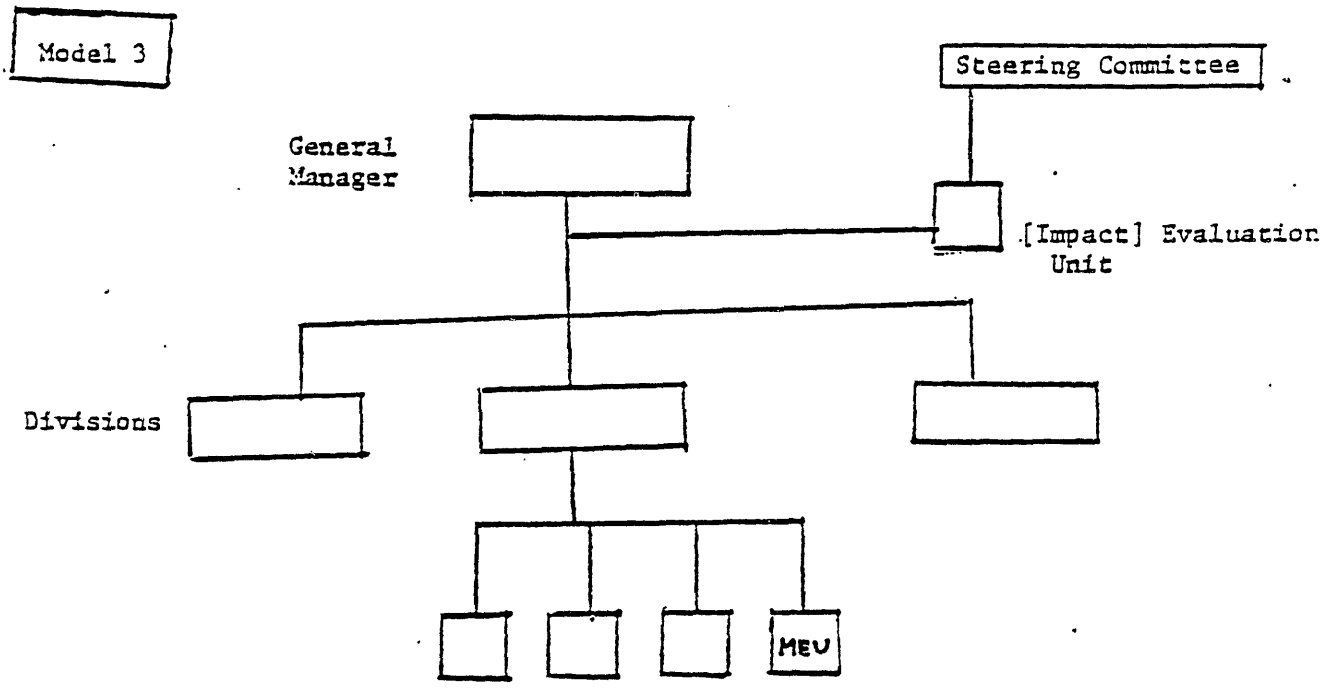


Monitoring and Evaluation Unit (MEU) reporting to one of the operating divisions



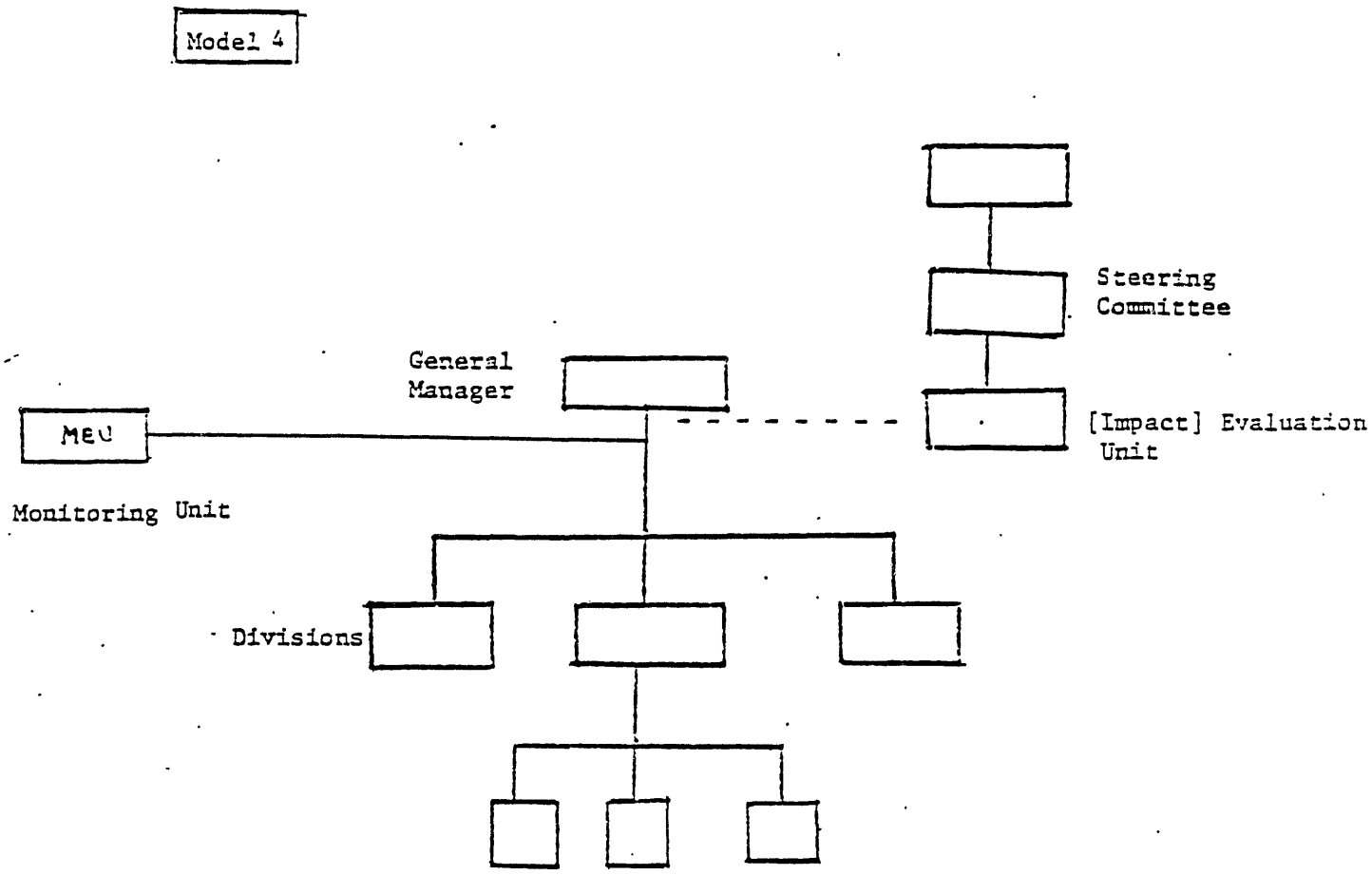
Monitoring and Evaluation Unit (MEU) reports directly to the General Manager

Fig 5-2 (continued)



[Impact] Evaluation Unit reports directly to General Manager and is sometimes guided by a steering Committee. The Monitoring Unit reports to one of the operating divisions.

Fig 5-2 (concluded)



Separate Impact Evaluation and ← {Monitoring} Units, each of which reports directly to the General Manager.

which reports to one of the operating divisions. The problem here is that the research will be subordinate to one division and will not be able to examine broader issues affecting other divisions, or to influence general policy of the organization.

In Model 2, the monitoring and evaluation unit reports directly to the general manager. This option permits the research unit to conduct general studies of a project's efficiency and impact. A potential problem is that the evaluation may become too remote from the operating divisions.

In Model 3, two separate units have been created; an impact evaluation unit reporting directly to the general manager and a monitoring unit under one of the departments. This has the advantage that the monitoring unit can respond directly to operational needs, while the impact evaluation unit can conduct more general and long-term studies. A potential danger is that the separation of functions may lead to a lack of coordination, with either a duplication of functions or certain important types of study not being conducted. For example, potentially important medium-range studies such as reasons for poor cost recovery or efficiency of mutual help construction groups, may be difficult to conduct because they do not fit into the terms of reference of either unit.

In Model 4, both the evaluation and monitoring units report directly to the general manager. This model permits greater objectivity and flexibility of hiring staff, but it runs the risk of making the evaluation too remote from the operational needs of the organization. This model can also be used when part of the evaluation (usually the impact studies) is subcontracted to consultants.

In selecting the appropriate organizational model, the following guidelines should be kept in mind:

** For the evaluation to be operationally useful, senior management must participate actively in the planning and review of the evaluation program. If the evaluation is relegated to a lower level of the organization or is located within one of the specialist divisions it will automatically lose most of its effectiveness.

** In a small organization the evaluation unit should probably report directly to the General Manager.

** In a larger organization, such as a national housing authority, which is responsible for a number of different housing projects, the decision is more difficult. If the evaluation unit reports to the General Manager (or Executive Staff) it may be too far removed from the project. If, on the other hand, it reports to a specific project manager or department head, then access to senior staff is reduced. Where there are two or more separate evaluation exercises relating to different projects, the solution may be to have a small central advisory staff reporting to Senior Management, and a special evaluation unit reporting to the manager of each project being evaluated.

** A balance must be achieved between the requirement of objectivity (and, hence, a certain distance from the day-to-day activities) and the need to maintain close contact with operational activities.

** The evaluation must be able to conduct both basic monitoring studies and more general impact and policy studies. This can either be achieved by having two units, or by specialization within one unit. Where there is specialization it is important to avoid problems of coordination and communication.

** Autonomy may be required to provide the necessary administrative flexibility to contract high grade staff, but it is essential to avoid forming an elite which creates resentment and which is hard to reintegrate into the organization in the future.

** Box 5-4 compares projects which combined and had separate monitoring and evaluation units so as to demonstrate the relative merits of each approach.

BOX 5-4 THE MERITS OF SEPARATING OR INTEGRATING MONITORING AND EVALUATION UNITS

Comparing the experience of urban projects in Zambia and Senegal

** In the evaluation of the First Urban Project in Zambia, an evaluation unit was established with World Bank-IDRC funding. At the same time, an already existing monitoring unit reported directly to the project director. The main function of monitoring was to provide rapid studies, usually conducted within a month, which were a direct response to management requests. These studies tended to be descriptive with very little analysis, but were considered by management to be very useful. On the other hand, the evaluation unit, which had less direct contact with management, was more concerned with longitudinal impact and more general studies. An advantage of the separation of the two units was that the functions of each unit were clearly defined which helped develop expertise in particular fields and the production of a high output in each area. Some of the disadvantages were a lack of coordination between the two units, and relatively little cooperation on studies or exchange of information. It also meant that neither unit systematically monitored the efficiency of project implementation.

** In the First Sites and Services Project in Senegal, the Bureau d'Evaluation was responsible for both monitoring and evaluation, and was able to produce a considerable output of both types of study. Partly due to this integration, there were a number of process evaluations conducted, which in fact proved to be some of the most useful studies. Data from monitoring studies was built into the impact evaluations, and vice versa.

b. Should a special evaluation unit be created?

It is sound organizational procedure to avoid the creation of additional units which increase overheads and produce bureaucratic delays and problems of control. In addition, if a special evaluation unit is created with outside funding and special employment conditions, it may prove difficult to re-integrate the staff of this unit into the organization when the pilot project ends.

A further problem is that the relative autonomy of an independent unit may also mean that it is financially weak, as it may not have direct access to government funding. In several cases these special units have experienced long delays in paying staff and meeting other expenses while waiting for the special funding to arrive.

On the other hand, there are a number of reasons why an autonomous evaluation unit may be recommended. Where the evaluation program is very large it becomes almost essential to create a special unit. The autonomy of the special unit can also provide more independence and objectivity as the evaluation is not controlled by any one department. The greater autonomy may also permit greater flexibility in hiring staff (offering more attractive salaries, etc.). Although this may create long-run problems if the unit has be be re-integrated, there are many countries where it would be impossible to attract good research staff if only public administration salaries could be offered.

c. The Role of a Steering Committee

A Steering Committee, which oversees Monitoring and Evaluation, can have a number of important functions. The first is to provide technical guidance in the design and review of studies by including representatives from the local universities as well as researchers from other government agencies. Second, many projects have four or more agencies involved in different stages of the implementation process and the Steering Committee can provide a way to ensure their interests are represented in the research. Third, representatives can be invited from organizations interested in using the evaluation findings (Ministry of Planning, etc.). Finally, the Steering Committee can ensure a certain degree of objectivity of the research. It is important not to make the Steering Committee so large as to become ineffective; nor should it put unnecessary barriers in the way of report publication.

C. ORGANIZATION OF MONITORING AND EVALUATION FOR THE PROJECT COORDINATING AGENCY (PCA)

1. Distribution of monitoring and evaluation functions between the project coordinating agency and the implementing agencies.

There is no single best way to distribute monitoring and evaluation functions between the PCA and the implementing agencies because projects are organized in different ways; vary in size, areas covered, and in the number and research capacity of the organizations involved. (Box 5-5 illustrates the organizational complexities of monitoring and evaluation in a large, multi-city project in Brazil, and identifies some of the many factors

which influence the distribution of evaluation responsibilities). Where implementing agencies have only a limited research capacity the PCA will be required to assume a greater responsibility for monitoring and evaluation. This may be done by designing manuals and guidelines, by contracting consultants or by directly conducting some of the studies. PCA's also vary in their research capacity so that in some cases they may be a relatively passive recipient of monitoring studies, whereas in others they may initiate a wide range of impact and cost-effectiveness evaluations. The functions of the PCA are likely to include:

a. Assistance in designing the monitoring and evaluation system.

b. Providing technical assistance and possibly direct staff and financial support for conducting studies.

c. Assistance in selecting and supervising consultants.

d. Direct responsibility for conducting some of the more complex studies, or other types of needed studies which are outside the capacity of the local agencies.

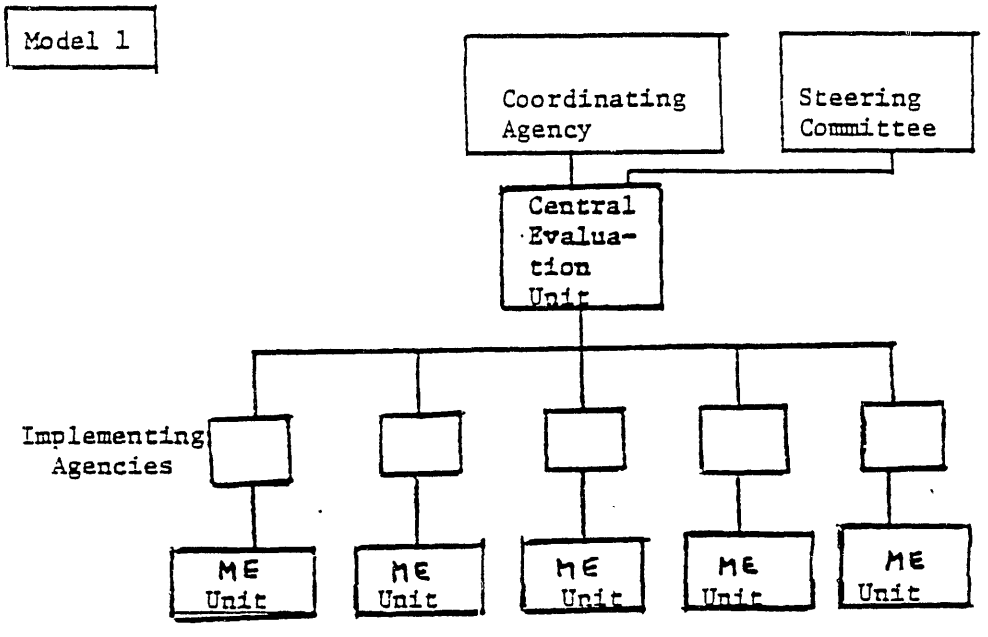
e. Training of evaluation staff.

f. Assistance with report preparation and dissemination.

Figure 5-3 presents two possible ways to organize monitoring and evaluation at the level of the PCA. In Model 1 each implementing agency has its own monitoring and evaluation unit which conducts performance and implementation monitoring, and possibly impact evaluation studies. The scope and complexity of the studies is likely to vary considerably from one agency to another within the same project. With this model the functions of the central evaluation unit (located in the PCA) may be limited to coordination and the preparation of summary reports or the data from the PIA's may be used as part of a more complex monitoring and evaluation program conducted by the PCA.

In Model 2 the PCA plays a more active role and has both a monitoring and an evaluation unit. It may be more directly involved in the design and implementation of the monitoring studies, as well as with impact evaluations and cost-effectiveness studies.

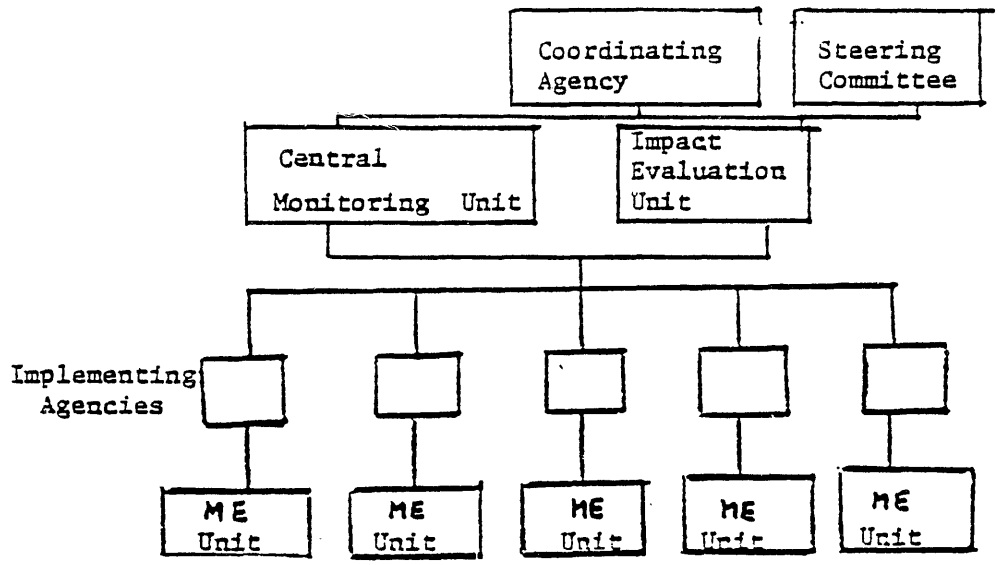
Fig 5-3 Alternative organizational structures for the monitoring and evaluation of a multi-component urban project



A Central Evaluation Unit, guided by a Steering Committee, is responsible for the design and coordination of the evaluation of the program and for the preparation of the program evaluation reports. This unit reports directly to the Coordinating Agency. A monitoring unit should be established in each implementing agency. The monitoring units are likely to vary in size and technical competence and technical assistance and direct logistical support may be provided by the Central Evaluation Unit.

Fig 5-3 (continued)

Model 2



Similar to Model 1 except that two separate units, the monitoring unit and the impact evaluation unit, now report to the Central Coordinating Agency. This model will typically be used in larger and more complex programs where more complex impact studies may be undertaken.

BOX 5-5 THE ORGANIZATION OF A MONITORING AND EVALUATION SYSTEM FOR A COMPLEX, MULTI-CITY PROJECT

The First Medium Cities Project in Brazil

The Medium Cities Project (CPM) is a complex project covering more than 10 cities throughout Brazil. The project is coordinated from Brasilia by the Urban Development Council (CNDU) of the Ministry of the Interior, with technical inputs from various federal sectoral agencies in the fields of labor, fishing, small businesses etc. In each city there is an executing agency (UAS) which is responsible to a city or state level agency. Each UAS has an evaluation unit and a budget allocation whose use has to be approved by CNDU. The evaluation staff are directly responsible to the director of the UAS but receive technical assistance and some supervision from CNDU. Consulting services are provided on design and analysis by locally contracted consultants and through 2 World Bank consultants who coordinate with CNDU. Some of the lessons from this experience are the following:

(a) Evaluation organization must be flexible and adapt to varying conditions at the city level.

(b) Flexibility and speed of implementation of the evaluation studies was considerably delayed by the need to seek approval for all studies from CNDU and often one of the federal sectoral agencies, both in Brasilia. The studies would have been much more responsive to local management needs if the UAS could have had direct control over a discretionary budget.

(c) The Bank consultants had a useful role in giving priority to more rapid, operationally useful studies and in speeding up the approval process. Their role was also useful in "selling" many of the more controversial studies to management and in giving support to the evaluation units.

(d) It proved difficult to establish clear evaluation priorities as each UAS, federal agency and CNDU had their own list of priority issues. The latter two tended to independently recommend studies to the UAS so that the evaluation units tended to become overloaded with data collection and developed a back-log of unanalyzed data, and uncompleted reports.

(e) Conferences and periodic meetings of evaluators from each city were organized by CNDU and proved useful in seeking uniformity of methods and procedures.

The major issue with both of these models is to determine how evaluation responsibility should be divided between the coordinating agency and the executing agencies. The decision will partly depend on the technical capacity and interest of each agency. The following guidelines are suggested:

a. at least basic monitoring activities should be assigned to each executing agency as it is important for them to consider the evaluation as a management tool rather than as external supervision.

b. technical assistance and possibly staff may have to be provided to help the executing agencies with the design, analysis and possibly the implementation of monitoring and evaluation.

c. the amount of monitoring and evaluation activity is likely to vary from one agency to another. In some cases, only very brief statistical documentation will be prepared, whereas other agencies may be able to undertake more complex research.

d. overall coordination and technical assistance should probably be the responsibility of the central coordinating agency. This agency should also be responsible for any more complex studies.

e. the guidelines presented in Section A can be applied.

D. MONITORING AND EVALUATION BY SECTORAL AGENCIES

Sectoral agencies such as a national housing bank, or small business development agency, are responsible for design of new projects, monitoring of implementation and evaluation of performance. These agencies must both monitor and provide technical support to the implementing agencies and conduct comparative and impact studies to help define future policy. In some cases they may initiate rigorous impact or cost-effectiveness studies. Some of the alternative modes of operation are the following:

1. The agency will have its own staff, including interviewers, who conduct and analyze the studies. This permits the agency to develop specialized skills, provide policy guidance to other organizations and conduct systematic comparisons of project performance in different cities and economic contexts. In order to develop a standardized methodology the sectoral agency must be prepared to take an active role and initiate a certain number of comparative studies. This approach has the problems that centrally designed

studies tend not to be responsive to local conditions, and that local agencies resent not being sufficiently involved.

2. The agency subcontracts major studies to outside consultants such as local universities. This has the advantage of involving universities in project development and can also provide access to experts in particular fields. One potential problem is the lack of involvement of local implementing agencies which can lead to resentment as well as affecting the quality of the work if the local experience is not used. Secondly there is a potential danger that the studies may become too academic and not sufficiently operational.

3. The studies are conducted by the local implementing agencies with technical assistance and some direct staff inputs from the sectoral agency. This approach has the advantage of fully involving the local agencies and makes it more likely that the results will be used. A common problem is that inadequate support is provided by the sectoral agency and the local agency is not able to adequately conduct and analyze the study. The greatest difficulty is often in the analysis stage, and it is common for many reports never to be produced due to lack of experience or resources to conduct the analysis. Another problem is that scarce local agency resources may be diverted to the evaluation, with the result that project implementation is negatively affected.

The best evaluation strategy will usually combine elements of these three approaches. The local agencies should be actively involved in the design, supervision and review of all studies, but a careful review should be made of their level of direct involvement in the implementation. Some involvement is a useful learning experience, but it is unrealistic to assume that local staff with limited research experience will be able to completely manage a complex study. It is particularly important for the sectoral agency to ensure that adequate provision is made for analysis and report preparation. An important advantage of the independent outside study is that it can permit a greater degree of objectivity, which can be very important when the performance of the project is affected by a set of local circumstances which it may be difficult for local staff to perceive or comment on.

E. ORGANIZATION OF MONITORING AND EVALUATION AT THE NATIONAL LEVEL

1. National urban development agencies

The responsibility for urban development policy usually involves the Ministries of Planning, Interior and in some cases Urban Development. Traditionally the responsibilities of these agencies were limited to the coordination and monitoring of ongoing projects with new projects evolving as an uncoordinated response to current political and economic pressures. However, many countries are now seeking to develop more coherent and integrated national development strategies and consequently the functions of these agencies are being broadened. In the example given in Box 5-5 a National Urban Development Council, with overall responsibility for the definition and supervision of a national urban development strategy, was established within the brasilian Ministry of the Interior Some of the typical functions of these development agencies are the following:

a. Developing and supervising a monitoring system to be implemented by the PCA or implementing agencies and which can provide rapid feedback to local and national management on all project components. This has been a difficult task as many monitoring systems have proved time consuming and cumbersome to use, with the information being produced too slowly to have much practical utility to project managers. A useful function of the national agency can be to sponsor special studies to test new approaches (such as participant observation or methods for conducting rapid feedback studies).

b. Where local resources and experience are very limited, an evaluation coordinator may be assigned by the national agency to work in the PCA in a particular city. This person would be trained, paid, and, to some extent, supervised by the national agency. It is important that the work of the coordinator is directly supervised by the local PCA and that as much data as possible is provided by staff of the local agency.

c. Coordination is a vital, but frequently neglected function, which needs to be carried out at various levels. First, the efforts of the sectoral agencies must be coordinated to ensure uniform approaches, avoid duplication and ensure that excessive numbers of studies are not being requested from the local project units. Second, all evaluation efforts must be coordinated within each city to avoid each component being studied in isolation and without consideration of how it affects, and is affected by, the

other components. Third, meetings and other forms of interaction must be arranged between evaluation units in each city.

d. Developing and implementing studies to evaluate the impact or cost-effectiveness of different projects. This is an essential function if systematic information is to be obtained on which an urban development strategy can be based. Examples of these studies can include:

i. Impact evaluations to determine whether projects are able to produce the intended effects on the socio-economic conditions of the target population. For example, a study to estimate the impacts of squatter upgrading on health, housing quality and household income.

ii. Cost-benefit analysis to compare the costs and benefits of alternative strategies. An example is the study of low-cost housing options in El Salvador in which the economic rates of return of 9 projects were compared.

iii. Cost-effectiveness analysis in which projects are compared in terms of their ability to achieve a stated output. For example, different methods of health delivery can be compared in terms of the cost per patient attended.

2. Finance ministries and legislative watchdogs

In all countries the ministry of finance or the treasury has basic monitoring functions related to costs and expenditures of projects, and to their implications for cash flow and receipts of disbursements from international loans. In some countries these functions are conducted by the executive branch, but in an increasing number of cases the legislature has created its own monitoring agency. Probably the most systematic legislative watchdog, in this rspect, is the US General Accounting Office (GAO) which is charged with the responsibility of monitoring and evaluating the use of all federally authorized expenditures. The GAO initially used traditional accounting and auditing approaches but since the Sixties has conducted rigorous program review to measure the impacts as well as the cost-effectiveness of government expenditures. Box 5-6 is an example of a typical GAO impact evaluation report. This evaluation, of the effects of budget reductions (Budget Reconciliation) on incomes and labor force participation of families with dependent children, showed that the Act had eliminated benefits for almost one half a million families. In most cases the families had not

been able to compensate by increased labor force participation so that the result was a net decline in their household incomes. The report was based on an analysis of 10 years of experience with family assistance programs. Studies of this kind are requested by the US Congress and used to assess the actual and potential impacts of legislation and the effectiveness of federally supported programs.

A number of developing countries have created somewhat similar agencies. One of the best known is the Programme Evaluation Organization created to oversee India's community development programs. PEO is an agency of the Planning Commission and reports to Parliament. It is considered to have been so successful that similar agencies have been set up to evaluate

BOX 5-6 EXAMPLE OF AN IMPACT EVALUATION CONDUCTED BY THE US GENERAL ACCOUNTING OFFICE

Effects of the 1981 Budget Reconciliation Act on Aid to families with dependent children

The following is the summary of the findings of the above study:

"The Omnibus Budget Reconciliation Act of 1981 (OBRA) made major changes in the Aid to Families with Dependent Children (AFDC),particularly in regard to AFDC benefits for many working recipients, and they reduced benefits for many others.

From its survey of state public assistance agencies and an analysis of 10 years of ... program data, GAO estimates that when the declines in caseload and outlays stabilized, OBRA had decreased the national AFDC caseload by 493,000 cases and monthly outlays by $ 93 million. However, because the caseload rose faster than predicted after this point, long-term effects are less certain.

..... These evaluations indicate that by fall 1983, most working recipients who lost benefits because of OBRA had not quit their jobs and returned to AFDC..... Although earnings increased for many who remained in the labor force, the respondents as a whole ... experienced significant income losses in all five sites. Apparently they did not make up the loss of income from AFDC and food stamps by working."

Source: General Accounting Office "An evaluation of the 1981 AFDC changes: initial analysis." April 2, 1984.

education, health and rural development programs. Agencies of this kind differ from most monitoring and evaluation at the national level in that they are responsible to the legislature rather than to the executive branch.

Unlike the urban development agencies which seek to develop evaluation capacity within local coordination and implementation agencies, financial and general audits are usually conducted by the central government agency or its consultants. One of the purposes of the audits is to check on the performance of the implementing agencies and to do this it is necessary to obtain independent data to avoid any deliberate or unintended distortions in data produced by the agency being evaluated.

F. RESOURCE REQUIREMENTS FOR MONITORING AND EVALUATION

1. Resource requirements for local implementing agencies

Table 5-1 shows that a monitoring system can frequently be implemented by a local implementing agency with as little as one half time professional and one part-time research assistant together with secretarial support. In addition several interviewers, who can often be borrowed from another department within the organization, will be required periodically. With this level of staffing it will usually be possible to produce the quarterly progress report, to conduct occasional special studies and to prepare an interim and final report on the evaluation. It should be feasible to obtain this minimum level of staffing within almost every urban project. It is important, however, to ensure that these staff members are permitted to work their assigned time on the evaluation and that they are not constantly being diverted to other activities within the organization.

The staffing levels indicated in Table 5-1 will only permit the preparation of the basic monitoring reports. Where there is a need for more reports or more detailed analysis it will be necessary to increase the staff level. Table 5-2 estimates the numbers of staff weeks typically required to conduct an average size impact study with a sample of about 600 households. It is suggested that the following resources will be required:

35-40 senior staff weeks
50-55 assistant staff weeks
55-60 interviewer and coder weeks
Secretarial assistance
Computer time
Publication costs
Travel costs
Possibly consultants must also be hired.

Table 5-1 - Minimum Monitoring Resource Requirements for a Local Project Implementing Agency Agency

Personal/Resource	Qualification	Functions
1. Half-time professional staff member (Director of the Evaluation)	Bachelors degree in one of the social sciences and ideally some research experience and/or experience in data analysis. 3. Design and implementation of occasional special studies. 4. Preparation of interim-final reports.	1. Prepration of financial status and project status reports. 2. Coordination with other departments, government agencies and communities.
2. Research Assistant (at least half-time)	Completed high school. Ability to conduct basic computations. If possible some experience with financial analysis.	1. Assist with analysis of surveys and field data. 2. Help conduct interviews with community leaders and residents. 3. Analysis of financial and other monitoring data from within the organization.
3. Secretary (at least half-time)	Normal secretarial skills. Experience in preparation of tables.	Preparation of reports and normal secretarial duties.
4. Interviews (often on loan from other departments	Completed high school. Ideally prior experience with interviews or social work. different aspects of the project. Interviews for special studies.	Periodic interviews with community leaders and residents to determine their opinions on
5. Transportation and driver	In some cases it will also be necessary to transport interviewers.	To permit visits to projects.

Table 5-2 Resource Requirements for the First Round of a Typical Impact Study With a Sample of 600 Households

PHASE		TIME (Weeks)	RESOURCES
1.	Planning and design	4-8	4-8 senior staff weeks
2.	Sample design	4-12	4 senior staff weeks 4-12 assistant weeks
3.	Open interviews to prepare instrument	2-4	1-2 weeks senior staff 2-4 interviewer weeks
4.	Drafting instrument	2-4	2 weeks senior staff 2-4 weeks assistant
5.	Pilot interviews (25-50)	2-4	1 week senior staff 4 weeks interviewer
6.	Preparation final instrument and guide	2-4	1-2 weeks senior staff 2-4 weeks interviewers
7.	Interviewer training	1	1 week senior staff 1 week assistant
8.	Interviewing (600)	4-5	1 week senior staff 10 supervisor/weeks 200 interviewer/days
9.	Developing coding frame	2-4	1 week senior staff 2-4 weeks assistant
10.	Coding	2-4	1 week senior staff 2-4 weeks supervisor 40 coder/days
11.	Inputting data to computer	1-2	1-2 weeks computer operator
12.	File creation and data cleaning	2-4	1 week senior staff 2-4 weeks assistant
13.	Basic descriptive analysis	4-8	2-4 weeks senior staff 2-4 weeks assistant
14.	Hypothesis testing and more complex analysis	4-8	4-8 weeks senior staff 4-8 weeks assistant
15.	Draft report	4-6	4 weeks senior staff 4-6 weeks assistant
16.	Review and revision of draft	4-6	2 weeks senior staff 2-4 weeks assistant
17.	Preparation of final report	2-6	2-6 weeks senior staff 2-6 weeks assistant
ADDITIONAL RESOURCES			Secretarial staff Consultants Travel funds Printing and publication costs Computer costs

The purpose of this table is to provide a checklist of the staff and financial costs which will typically be required for an impact evaluation. One important point to note is that the direct costs of interviewing will typically only represent between one quarter and one third of the total study budget and may only require between 6 and 8 weeks out of a total estimated study time of between 9 and 12 months. The lesson from this table is that large scale surveys tend to last longer and to cost more than the inexperienced researcher may expect.

Table 5-3 presents the recommended staffing level for a more comprehensive monitoring and evaluation program for a typical component of an urban project in one city. With this level of staffing it will be possible to prepare detailed quarterly progress reports, to conduct regular special studies and to carry out at least a simple longitudinal impact study. The characteristics and qualifications of the staff are described below:

a. Research Director: Responsible for designing the research, supervising analysis, preparing reports and coordinating with the different organizations involved. Ideally, an MA in sociology or a related field is required, together with some previous experience in research and evaluation. It is important to find someone with interest in operational issues to avoid making the research too academic. If more complex studies, such as longitudinal impact analysis, are to be conducted, the Director would also require experience in survey design and analysis. Experience with electronic data processing is useful.

b. Assistant Director: Responsible for coordination within the executing agency and for interviews with project staff and community leaders. Responsible for supervision of data analysis which may involve the use of a computer, and may be involved in participant observation studies or other qualitative studies. The success and utility of the evaluation will depend on maintaining close relations with the different divisions of the executing agency, so it is important to find someone with good human relations and the ability to understand the point of view of technical staff in various different fields. A BA is required, with some

Table 5-3 - Adequate Staffing Level for a One City Evalution

Position	Functions	Qualifications
Director	Coordination with management Supervision Report preparation Supervision design Coordination with other organizations	Preferably M.A. Minimum of BA with some practical research experience Experience with survey design and data analysis will be required for the more complex evaluations.
Assistant Director	Coordination with divisions of executing agency Interviews with project staff and community leaders Supervision of data analysis and report preparation Participant observation Training and supervision of interviewers	BA some background in research and data analysis.
Clerical Assistant	Revision and tabulation of data from project records Basic data processing	Completed High School
Secretary	Normal secretarial activities	Completed High School
2-5 Interviewers	Interviewing project participants Application of observation guide Coding/analyzing interview	Attending or completed High School
Driver	Taking staff to and from project sites	

experience in data analysis. An additional area of responsibility might be the training and supervision of interviewers.

c. Clerical Assistant: Responsible for revision and tabulation of data from project records (numbers and types of applicants and people selected, cost recovery, amount of material loans given, etc.). Also responsible for data processing, either manually or using a computer. The minimal requisite is completion of high school.

d. Interviewers: Two to five full-time interviewers will probably be required. For more complex evaluation programs, this number may increase up to 10-12. They will be responsible for conducting regular interviews with community residents, as well as for the application of observation guides and participant observation. The interviewers will also be involved in the coding and possibly basic analysis of the surveys. There are no general rules for selecting a good interviewer, and a team will have to be built up by trial and error. Students of architecture or engineering may be useful in studies of house construction. Trainee nurses often work well on studies of health and family-related matters, including household expenditure patterns. Older women are sometimes the best people to interview mothers, as many mothers do not like to discuss health and children with a man or with younger women.

e. Driver: If some project sites are inaccessible, it may be necessary to hire a full or part-time driver.

2. Resource requirements for the local project coordinating agency (PCA)

In planning staff requirements for Project Evaluation, it is necessary to define the staff required for both the Coordinating Agency and the Implementing Agencies. The numbers and levels of staff will depend both upon the size and complexity of the project, and upon which of the organizational scenarios is selected. The requirements for each implementing agency were discussed in the previous section.

There will be many projects in which the Coordinating Agency is only able to assign very limited resources to the evaluation. The following represents the minimum level of resources required to comply with the basic monitoring and evaluation requirements (Table 5-4):

a. One full-time professional staff assigned exclusively to work on the evaluation.

b. One full-time assistant for both clerical analysis of reports from the project agencies and for visits and interviews in the communities and with staff of other agencies.

c. Full-time secretary

d. Occasional access to interviewers

e. Access to transport

f. Small budget for hiring consultants.

With this level of resources, the Coordinating Agency should be able to:

- Help project agencies set up the basic reporting system.
- Prepare the quarterly progress report.
- Organize regular meetings of the inter-agency coordinating committee.
- Conduct occasional special studies.
- Prepare the interim and final reports.

For larger or more complex programs, or where a more sophisticated program of research is required, the types and numbers of staff will have to be increased. Some of the typical appointments to be made as the program grows include:

a. Deputy directors for monitoring and impact evaluation.

b. Field directors/coordinators for each city of a multi-city program.

c. Data manager. In most larger evaluations, the data will be processed by computer and the data manager will be responsible for the collection and analysis of all quantitative data. In most cases the data manager will have two sets of responsibilities. The first is to develop a management information system so that monitoring information from the implementing agencies can be routinely processed. In small programs, computer

Table 5-4 - Staff Requirements for Program Evaluation: Requirements for Implementing Agencies and Project Coordinating Agency

	For Each Implementing Agency	Central Coordinating Agency
Minimum Requirements	1/2 time professional Research assistant (at least 1/2 time) Secretary (at least 1/2 time) Occasional interviewers Driver and transport	1 fulltime professional 1 fulltime research assistant 1 fulltime secretary Occasional interviewers Driver and transport Consultant budget
Adequate Staff for More Complete Evalution Program	Research director (with Masters Degree in social science) Assistant director Research assistant 2-5 fulltime interviewers Secretary	Evaluation director (with at least Masters Degree) Deputy directors for monitoring and for impact evaluation One or more research Assistants Team of interviewers (number depends on size of program) One or more specialists in areas such as finance, engineering, statistics etc. Data manager/computer programmer. One or more secretaries One or more drivers

analysis will probably be conducted by one agency, but in larger programs, several agencies may have their own computers and the data manager will be responsible for developing computer interfaces so that data can be transferred from each agency to the central computing facility.

The second responsibility will be the analysis of surveys. Where longitudinal impact evaluations are conducted, survey analysis may become the main task of the data manager.

d. Permanent interviewers.

e. Specialist staff in fields such as financial analysis, engineering, sociology, and statistics.

3. Evaluation resources for the national development agencies

It is difficult to provide guidelines on evaluation resource requirements at the national level as the variation is so great. Some agencies, whose main function is to compile monitoring data received from the field, may only have 2 or 3 staff assigned to monitoring and evaluation. On the other hand, organizations such as the U.S General Accounting Office and the Program Evaluation Organizations in India may have several hundred staff (the GAO, for example, produces an average of 4 evaluation reports every working day!).

4. Fitting staff into civil service categories

Many evaluation positions are difficult to describe in terms of existing Civil Service grade levels, and this has often made it difficult to offer competitive salaries to well-qualified staff. If salary scales do not correspond with current salaries being paid to this type of person elsewhere, it may be necessary to consider special procedures to resolve the problem. The resolution of these personnel problems can take a considerable amount of time, and for this reason the discussions with the Personnel Management Department should begin early in the project cycle.

Another related problem is that personnel departments are usually reluctant (with good reason) to create new permanent positions for a program which only lasts a few years. As posts once created are permanent, this means

that other positions would have to be found for these staff members once the evaluation ends. Again, it will be necessary to discuss this at an early stage so as to permit time for finding the appropriate solution. One method of circumventing this problem is to contract consultants for certain key positions as there is usually more flexibility in terms of their salary and employment conditions.

5. Staff training

Evaluation research is a relatively new field, and training programs are frequently not offered in the universities, the question arises as to how to provide staff with the necessary training. Monitoring does not normally require exceptionally high skill levels so that the problem is less severe. However, if the general upgrading of staff is a project objective, it may be useful to consider a scholarship program to enable evaluation staff (as well, of course, as other staff) to participate either in short three- to four-month courses, or to study for an advanced degree in a field related to the evaluation.

6. Temporary assignment of other project staff to the evaluation

An effective way to reduce the evaluation budget is to make maximum use of staff already in the organization. For example, community promoters can be used as interviewers. In addition to reducing costs, this also has the advantage of integrating the evaluation into on-going project activities.

Another possibility is to assign a new senior staff member to work with the evaluation team as a way of getting to understand the workings of all aspects of the project. This could be a very good introduction for someone who will later assume a senior position in the organization.

7. Using consultants to overcome salary constraints

A problem with staff recruitment in many countries has been the fact that civil service salaries are often not attractive to well-qualified researchers. One reason for this is that university researchers often earn part of their salary through outside consulting, something which will usually not be permitted while working full-time for the government.

A frequent solution has been to hire the research director or a senior staff member as a consultant as there are usually not the same types of constraints on consultant salaries. This is not an ideal solution, as the long term problem has not been resolved.

G. THE ROLE OF DONOR AGENCIES IN THE EVALUATION

Donor and lending agencies often have an important influence on the design and implementation of program evaluation. In many cases the implementation of an evaluation will be one of the loan effectiveness conditions, and consequently the donor may have considerable influence on the contents of the evaluation.

The role of the donor can be both beneficial and detrimental to the evaluation. The positive aspects arise from the incentives provided to initiate an evaluation and often the provision of technical assistance. In many cases it is likely that the evaluation would not have started had it not been for the donor's encouragement. The negative aspect is that the evaluation may be oriented more towards the interests of the donor than to the needs of the borrower. The donor may require, for example, that an expatriate consultant be appointed, and he or she will often impose a very different approach than would have been used by the borrower. In some cases, for example, the evaluation has become more complex or has focused more on long term impact studies than would have been the case had it been designed by the borrower. A related problem is that the evaluation unit may be confused as to whom they report to. Although they may formally report to the borrower, in practice they may receive more guidance and even direct supervision from the donor. This may also cause the evaluation unit to be mistrusted as they are perceived to be spying on behalf of the donor.

Another issue is the question of how the evaluation should be financed. Although the evaluation may be financed by the borrower under the loan, the evaluation unit may be encouraged to conduct more complex and expensive studies to respond to questions which are primarily of interest to the donor. In this case it can be argued that the donor should share the cost of the studies.

The following are some general guidelines which can be used to define the role of the donor/lender in the evaluation:

a. A clear distinction should be made between the types of information which must be collected as a loan requirement, and any additional

studies which are proposed by the donor. The latter should only be conducted after careful review by the borrower and in many cases with the donor covering at least part of the costs.

b. The management of the executing agencies should take an active role in the design and implementation of the evaluation.

c. The borrower should not expand the scope of the evaluation at the suggestion of the donor unless a very clear agreement is reached on how the additional data will be analyzed and used. There are many cases in which expensive studies have been conducted at the suggestion of donors but where the data has never been analyzed as no resources were allocated for this purpose.

d. Where a consultant is contracted, the guidelines suggested earlier in this chapter should be followed.

H. POTENTIAL PROBLEMS AND POSSIBLE SOLUTIONS

The following are some of the problems which typically arise in the management of the evaluation. In each case possible solutions are proposed (indicated by an *):

1. The process of recruiting staff is very time consuming so that by the time staff are hired the project is already well advanced. As a consequence the earlier stages of the project are not evaluated.

* The organizational structure and staff requirements for the evaluation should be defined at the time of project appraisal and recruitment should begin as soon as the loan is approved.

2. It is difficult to hire experienced researchers under normal civil service salary and employment conditions.

* Informal consultations should be held with the appropriate personnel departments at the appraisal stage, and the possibility of making special arrangements should be discussed. Some arrangements which have been used in the past include the contracting of staff as consultants rather than permanent positions; special salary scales for the project unit; special payment arrangements etc. These requirements are much easier to arrange at the project negotiation stage than once the project is underway.

3. The evaluation unit often becomes isolated and loses contact with the other departments of the implementing and planning agencies.

* Care should be taken in defining the organizational structure to ensure that the evaluation unit reports directly to senior management. A review process should also be instituted to ensure that all evaluation reports are discussed at management meetings.

4. Evaluation staff are frequently coopted to assist in a wide range of organizational activities unrelated to the evaluation, with the consequence that evaluation studies fall behind schedule.

* There is no simple answer to this as crises and unexpected activities constantly arise. As the evaluation work is less urgent than many other activities there will be a strong temptation to draw on the evaluation staff. Ironically the more successful the evaluation studies, the more likely management is to draw on the evaluation team to help in other activities such as writing reports and proposals and meeting with visitors.

* One advantage of using consultants is that they are not subject to the same distractions as regular staff. If the resources are available it may be a wise precaution to overbudget the evaluation staff on the assumption that a certain amount of time will be used in other activities.

5. The feasibility of conducting the evaluation studies may be affected by interagency conflicts. Agencies frequently resent the fact that they are being "evaluated" by another ministry, and will consequently not cooperate.

* As far as possible the implementing agencies should be responsible for preparing their own monitoring reports and the role of other agencies should be to provide technical assistance.

6. The practical utility of the evaluations may be significantly reduced by the slowness of receiving information from the different implementing agencies. By the time information has been received and compiled, a number of months may have gone by so that the information is out of date before it is published.

* One of the best ways of avoiding this problem is to keep the information to be collected as simple as possible. The implementing agencies should also be provided with rapid feedback so that they see that the information is in fact being used. Collection and processing can also be streamlined if a study is conducted to identify the main bottlnecks and to propose alternative methods of collection and analysis.

CHAPTER 6: ISSUES AND APPROACHES IN EVALUATING NON-SHELTER RELATED URBAN PROJECTS

This chapter reviews some of the non-shelter related urban projects which have evolved in recent years and discusses the extent to which the monitoring and evaluation framework discussed in earlier chapters is applicable to them. The 4 types of projects which are discussed are: income and employment generation, health, transport and urban and municipal development. It is concluded that the techniques of performance evaluation can be readily applied to all of these new types of projects, and that process evaluation can be easily applied to the first three and with some difficulty to municipal and institutional development. The main difficulties occur in the evaluation of the impacts and cost-effectiveness of the projects. Problems arise because the size and scope of many projects makes it difficult to identify a control group, to specify and measure impacts or because the project does not have a single set of outcomes and impacts which can be clearly defined and measured. Strategies are recommended for the application of each type of monitoring and evaluation study to each of the 4 project areas.

A. NEW DIRECTIONS IN URBAN DEVELOPMENT PROJECTS

The first generation of World Bank financed urban projects, beginning in the early Seventies, were mainly small scale shelter projects designed to demonstrate the feasibility of more economical approaches to the provision of shelter. The two main approaches were sites and services, in which new units were provided, and squatter upgrading in which urban services, infrastructure and house contruction credit were provided. In most of the countries where the first projects were developed, the concepts were relatively new and the projects were designed to test the approaches and to convince policy makers of their practical utility. Part of the process of testing and replicating the projects involved the implementation of monitoring and evaluation systems which would permit a more precise measurement and documentation of the results of these projects.

After a few years two things became apparent. First, well run sites and services or upgrading projects were potentially able to provide shelter at a much more economical cost than conventional housing programs, and that these solutions were accessible to large sectors of the low-income urban population who had previously been excluded from public housing. Second, it also become clear that these approaches, however successful, were only one component of a general strategy for improving the living conditions of the urban poor.

There were a number of reasons for this conclusion. First, intensive Bank involvement in the design and execution of specific projects was too slow and cumbersome an approach to make any quantitative impact on the shelter supply. Second, shelter is only one of a series of related services which are needed to make significant improvements in the living conditions of the urban poor. Income and employment generation, health, food distribution and marketing and transport are some of the other areas needing to be addressed. Third was the need to improve the capacity of municipal authorities to select, plan, finance and implement the development projects. Finally, and related to this, was the need to build the institutional capacity of regional and national agencies to plan and manage multi-component projects, which could be implemented simultaneously in a number of different cities.

The purpose of this chapter is to discuss some of the key issues relating to non-shelter urban projects and to suggest some of the monitoring and evaluation approaches which can be used in their evaluation. The discussion is by no means exhaustive, largely because sufficient evaluation experience does not yet exist. In each section we will discuss some of the key research issues and will consider how far they are unique to this particular type of project, and the extent to which the designs discussed in the previous chapters can be used. Four types of projects will be discussed: income and employment generation; health; transport; and municipal and institutional development.

B. INCOME AND EMPLOYMENT GENERATION PROJECT COMPONENTS

1. Key policy issues

Income and employment projects seek either to integrate workers into the formal labor market (mainly through training and job placement) or to stimulate the growth of the informal sector. Programs to develop the informal sector have been given more attention in urban projects. The main approaches have been to provide credit and technical assistance, serviced land for businesses purposes and access to new markets. The following discussion will mainly address projects designed to develop informal sector and micro-businesses.

One of the issues which continually arises is how to ensure that projects are accessible to low-income beneficiaries whilst developing rapid and economical techniques for selection, technical assistance and supervision of credit. Most projects involve banks and other financial and training institutions who are used to working with much larger businesses. Their appraisal and processing procedures are designed for working with large companies and in most cases the procedures are not applicable to small bsuinesses who have very limited experience in basic management procedures such as accounting, stock control and marketing. One of the consequences is a tendency to exclude the smaller businesses and to work mainly with the larger and better organized applicants. A major evaluation issue is to determine how well the programs are reaching the target population, and what are the main factors limiting access.

The problems do not all lie with the lack of willingness of banks and other institutions to work with the small businesses. Another set of problems relate to the social and economic costs for a small business to participate in the project. Many businesses, due to their quasi-legal status, avoid many taxes and are not subject to many of the health and labor regulations. In order to apply for a loan, the businesses would in most cases be required to go through the process of legalization (which is costly and complex) and would have to pay taxes and comply with other regulations. For many businesses the costs of participation may outweigh the benefits. Projects participation may also imply social costs. For example, for some families the money lender is an important resource in times of family emergency and borrowing money from the project rather than from the money lender may result in their losing access to the emergency assistance which the money-lender could provide to his or her regular clients.

The social, economic and political world within which the small businesses operate, is often not well understood by the designers of the small business projects and this may lead to poor project design. For example, the political influences within a cooperative may exclude certain groups or result in loans being directed towards larger businesses. The political contacts of suppliers and middlemen may affect access to markets and may force the new businesses to seek some form of accommodation with these groups.

Many projects involve intermediary organizations such as cooperatives, community groups or federations of small businesses. It is essential to understand how well these function as intermediaries and communication links, as in many cases the intermediaries only represent certain sectors of the beneficiaries or provide inadequate information on the opinions and needs of beneficiaries.

Another issue relates to the difficulties of conducting efficiency and cost-effectiveness analysis of many of the programs due to problems of getting accurate information on costs, income and profits.

A final issue relates to the difficulties of projecting the potential project impacts on the target population. Many businesses which have been successful on a small scale find that once their size increases they must enter directly into competition with much larger and better organized competitors and find their growth potential curtailed. Businesses directed to low-income households also confront the low purchasing capacity of this

sector. It is much easier to train people to make shoes than it is to increase consumer income so that demand for shoes will also increase. Factors such as these make it difficult to project from the findings of a pilot project to what would happen if the project were replicated on a larger scale.

2. Evaluation issues

a. Design issues

Most small businesses operate in an environment where costs, prices and markets are subject to continual fluctuations. In order to evaluate the impacts of a small business development project it is therefore important to have a control group which can be used to provide information on market trends. Unfortunately it is difficult to locate and use control groups, and there are also political and ethical problems involved in withholding project benefits from people in the control group. In most cases there is very little systematic information available on the numbers, types and location of small businesses so that it is difficult to locate a representative sample of businesses not affected by the project. It is also likely that the project will attract the most dynamic businessmen so that the socio-economic characteristics of participants may be quite different to those of the control groups. Finally it is extremely difficult to isolate and control for the impacts of external factors on the progress of the project. For example, the production costs and sales prices of many products fluctuate continually, businesses may be affected by the entry of new larger competitors, government agencies may decide to place large orders at uneconomic prices for political reasons, and businesses are continually affected by changes in the macro-economic environment. Not all businesses are affected in the same way by these factors so that it is difficult to assess how much of observed changes can be attributed to the project.

Another analytical problem is caused by the wide diversity in the types of businesses. Even a small program may give credit and assistance to businesses as diverse as carpenters, shoe-makers, dress-makers, food retailers and artesans working in metals. Each of these types of businesses should ideally be examined separately to understand the factors determining

its success or failure. This makes it difficult to generalize from the findings of programs which only affect a few hundred businesses (this is not a problem for the few larger programs which affect larger numbers of businesses).

b. Measurement issues

The evaluation of the impacts of a small business program requires that accurate information be obtained in expenditures, income, stock levels and profits. This information is frequently difficult to obtain, either because the business does not keep the necessary records or because the businessman does not wish to make the information available. Most small businesses do not keep accurate records. For example, small food stores, as well as other types of small businesses, frequently mix the household budget with that of the store; basic necessities are purchased with whatever money is available at the end of the day, and if anything is left over it is used to buy more stock or materials. No systematic records are kept and the owner would find it very difficult to estimate how much profit had been made.

The problem of obtaining accurate information is particularly difficult for the control group. Many projects seek to provide technical assistance in book-keeping and stock-control so that the quality of beneficiary records is likely to improve as the project progresses. However, this assistance is not given to the control group so there may be differences in the accuracy of information from participants and control group. Control businesses also have less incentive to cooperate, which again may affect the quality and detail of information provided.

Problems also exist in obtaining information on sources and costs of credit. Many credit arrangements are informal and small traders often think in terms of how much has to be repaid at the end of the day or week, rather than of interest terms. Similarly the cost of providing short-term credit to customers may not be computed. Although interest rates can often be computed from this type of information, there is a potential for a larger margin of error in the estimations.

3. Applicability of monitoring and evaluation designs discussed in previous chapters

Although specific measurement issues exist, in general the monitoring and evaluation strategies discussed in earlier chapters can be used. Some of the similarities and differences are as follows:

a. Business programs often require more extensive monitoring information from beneficiaries than is the case for many other types of projects and this facilitates conducting the types of performance monitoring described in Chapter 2. The main problem is to ensure the reliability of the information.

b. The methods of process monitoring discussed in Chapter 3 are directly applicable. It is important to develop a model of the process of project implementation, so as to be able to assess the weaknesses and potential problems. Several of the examples given in Chapter 3 in fact refer to small business programs. In principle, it is easier to measure efficiency and cost-effectiveness of small business programs as there is general agreement on which indicators should be used.

c. The evaluation of project impacts can also use the approaches discussed in Chapter 4. One of the problems in applying quasi-exerimental designs is the effect which the wide diversity of types of business has on sample size. Qualitative techniques are an important complement to quantitative analysis given the diversity of types of businesses and the importance of understanding social and political factors.

d. Difficulties in obtaining reliable information on the performance of the control group is likely to be a serious constraint on impact evaluation. For reasons discussed above, the information is likely to either not be available or for there to be serious questions about its reliability.

C. FAMILY AND PUBLIC HEALTH PROGRAMS

1. Key policy issues

There are 5 main types of urban health programs: (a) Maternal and child care (b) Nutrition (c) Water supply and sanitation (d) Solid waste disposal and (e) family planning.

A first general issue which relates to all of these projects is cost-effectiveness. All projects have to be paid for, either directly by beneficiaries or indirectly through local or central government. Costs affect accessibility and replicability and it is important to monitor carefully the different cost components and to compare them with alternative programs. The application of conventional cost-effectiveness techniques is extremely difficult. A major problem is the difficult of defining and measuring the

quality of outputs. There is no generally accepted way to define a good family planning program or an effective maternal and child care health project. This makes it extremely difficult to develop a standard measure of effectiveness and hence to be able to compare the cost-effectiveness of different programs.

A second related issue is cost-recovery. The efforts to make the project affordable and financially self-supporting are largely wasted if the beneficiaries are unwilling or unable to repay. The issue of cost recovery is complicated for many types of public health programs due to the difference between social and private preferences. Many families are unwilling to pay for disease prevention programs such as improved sanitation or solid waste disposal as they do not perceive how they themselves will benefit. Where cost recovery problems exist, it is important to assess whether these are mainly due to: (a) lack of capacity to pay (b) dissatisfaction with the services and lack of willingness to pay (c) political pressures discouraging an effective collection system or (d) inefficient adminstration. Box 6-1 illustrates the interaction between these factors in a project to provide drinking water in Lusaka. Assumptions about capacity and willingness to pay vary greatly from one project to another, and in many cases seem to be somewhat arbitrary. This is an area in which further research is clearly needed.

A third factor relates to the involvement of beneficiaries in the design, implementation and maintenance of the project. Many projects are designed without consulting the community and are then "sold" by technical staff to the community. Advocates of community participation argue that this lack of community involvement will often negatively affect usage, cost recovery and maintenance. However, the issue is complicated by the fact that many people question the efficacy of using free community labor in project implementation. It is difficult to supervise voluntary labor so that the quality of the project may suffer. There is also the question of whether it is justified to make the poor work to obtain services which higher income groups may receive free. Many of the problems related to project implementation are due in part to inadequate communication links between the community and the implementing agencies, and some of these problems may be reduced by more effective community participation.

BOX 6-1 COMPLEXITIES OF ENSURING COST-RECOVERY FOR PUBLIC HEALTH SERVICES.

Cost recovery in a water supply project in Lusaka.

The First Urban Development Project in Lusaka, Zambia included filling-in insanitary wells and the construction of communal standpipes to be shared by groups of 25 families. In an attempt to use group pressures to encourage payment of service charges, the water would be cut-off if any one of the families fell behind in their payments. With a few exceptions, the repayment rates were very low, despite the fact that water was cut off in some cases. Two negative effects of cutting off water were that families made illegal connections and in some cases opened up the cholera infected wells.

The following were some of the reasons for poor cost recovery:(a) Many local politicians opposed poor families having to pay for water and tried to stop water being cut off (b) The project accounts were several months in arrears so that families sharing a standpipe did not in fact know who had not paid (c) the office where payments had to be made was pooring supervised and was often not open when families tried to make payments (d) the idea of group pressures to ensure cost-recovery was only introduced at a late stage when defaults had already reached a high level. The process had not been explained to families and many neither understood nor accepted the concept of communal responsibility (e) Many families were recent migrants from rural areas where water was considered to be a gift from God which was obtained free from rivers or wells. Many did not understand why they should be expected to pay for water.

Source: Bamberger, Sanyal and Valverde "Evaluation of Sites and Services Projects: The Experience from Lusaka, Zambia". World Bank Staff Working Paper No. 548. 1982.

Fourth, technical problems frequently limit access to certain water supply and sanitation systems. The efficient operation of the system may be determined by the water table, types of sub-soil, elevation and gradient, and consequently certain communities may be excluded on technical grounds.

A fifth issue relates to the accessibility of the project to all sectors of the target population. Access may be affected by costs, cultural and political factors or the way in which the project is designed and implemented (Box 6-2). Frequently project services are mainly enjoyed by the better off or more influential sectors of the community. For example, houseowners may receive improved water and sanititation, but the benefits may not be shared with renters; ethnic or religious minorities might be excluded; or benefits may go largely to the supporters of a particular political group.

BOX 6-2 CULTURAL FACTORS AFFECTING ACCESS TO PUBLIC HEALTH PROGRAMS

a) In Mombasa the traditional form of low-income housing is the Swahili household in which 10 or more families live in a compound and share toilets and water supply. In order to provide a cost-effective sanitary unit, toilets and standpipes were constructed outside the compound. However, most of the families were strict Muslims and women were not allowed to leave the compound after dark with the result that they were not able to use the new toilets.

b) In Indonesia difficulties were encountered with the acceptance of communal toilets, due to the fact that different cultural and religious groups in the communities had different sanitary practices and were not willing to share a toilet.

c) Water collection and washing of clothes are frequently communal activities which provide an opportunity for women to meet and exchange gossip. In traditional societies this may be one of the few times that women are permitted to leave their compound, and consequently many women have not wished to accept individual water supply or washing facilities as this would cut them off from an important social activity.

d) In a classic study in Peru (Rogers, 1962) slides were shown to illustrate the types of bacteria which live in water and to explain the importance of boiling drinking water. The project had very limited success, firstly because traditionally only sick people drank warm water, and secondly because the bacteria shown in the slides seemed very large and no-one could see them when they examined their own water. The only people to initially boil water were a small number of social outcastes such as a woman who had been abandoned by her husband.

A sixth issue relates to the way in which the projects are managed. Many implementation agencies are not aware of the special problems involved in the implementation of social projects and they are managed in the same way as a civil works project. The effects of this approach on design, cost recovery and accessibility/acceptability for different sectors of the target population, must be assessed.

A final, but extremely important issue is to assess the impacts of the projects on health. Although the stated objective of most projects is to improve health, there are extremely few projects in which any successful estimates of health impacts have been made. Although health impacts are extremely difficult to measure (see following section) the lack of adequate information is also due to the fact that very few studies have tried to measure health impacts in any systematic way. In some cases this is because it is considered obvious that improved water supply or sanitation will improve health, in other cases it is because health impacts are considered too difficult or expensive to measure. For all of these reasons a key issue is to determine what impacts the projects actually do have on health, which groups do and do not benefit and what are the factors affecting outcomes. The key question for the development planner is to identify the most cost-effective package of services to produce a given set of health outcomes.

2. Evaluation issues

a. Design issues

The success of health programs appears to be very sensitive to the project delivery system, and for this reason the first stage of the evaluation design must be to develop a model of the implementation process. Widely accepted efficiency indicators exist for many of the specific technical components (for example the cost-effectiveness of different types of water supply or the numbers of patients per doctor in different types of clinic), but issues relating to community involvement, the effects of cultural and political factors, and the access of different groups to the project, may be less well understood. As mentioned earlier the management style of many health projects tends to focus on technical, rather than cultural and political issues.

Relatively little progress has been made in developing satisfactory models for evaluating the impacts of health programs. Some of the design issues include:

i. Defining an adequate set of indicators of health outcomes. For the evaluation of most types of health project a basic conceptual issue relates to the selection of indicators of project success. What should be measured? For example, which of the following indicators of health impacts should be used?: days absent from work due to illness, infant mortality, morbidity rates for particular population groups, anthropometric measures, parasitic load etc. A clearer conceptual framework is often needed to decide on the appropriate set of indicators. The indicators must be reliable, cheap to apply, logically related to the project and capable of detecting changes with the size and type of sample being used.

ii. Developing a model of the project implementation process, including the effects of household characteristics and the external environment.

iii. Specification and quantification of project treatments (water, sanitation, health education etc).

iv. Specification and quantification of the intervening variables (household and community attributes).

v. Development of a quasi-experimental design in which participants are compared with a control group, and in which measurements are made at several points in time.

vi. Application of multivariate techniques (see chapter 4) to control for differences between the experimental and control groups and to estimate the contribution of each project component.

b. Measurement issues

Health indicators are subject to at least 3 types of measurement related problems. First, a number of indicators are derived from information provided by a respondent on health status, illnesses and causes of death. There is considerable evidence as to the large margins of error in many of these indicators (particularly types of illness and causes of death). There is considerable disagreement among experts as to how reliable mothers reports are on the frequency of diarrhoea, which is unfortunate as this is a potentially economic way to obtain information on parasitic infection.

A second issue relates to cost, as many of the more reliable and analytically useful indicators, such as analysis of stool samples, medical inspections and measuring the height and weight of children, are relatively expensive and time-consuming.

A third issue relates to the size and type of sample which is required for the use of different indicators. Infant mortality rates are potentially a very good indicator of health status, but normally a sample of several thousand households would be required to obtain statistically reliable indicators of project impact on mortality rates. Many other indicators such as parasites, specific illnesses and height for weight are subject to seasonal fluctuations so that ideally measurements should be repeated several times during the course of a year.

3. Applicability of the monitoring and evaluation framework

Monitoring and evaluation techniques can be applied in much the same way as described in Chapters, 2,3 and 4. Issues of cost-effectiveness, accessibility and overall efficiency can be studied in the way described in Chapter 3. Similarly, the discussion of impact evaluation and the use of quasi-experimental designs is directly applicable to health issues. Measurement and sample size need to be given special attention but the general principles are the same. It is again essential to combine quantitative and qualitative techniques and to develop a general model of the project implementation process.

One special issue is that in many projects most of the information for both monitoring and evaluation will be collected through project health staff such as health visitors, social workers and nurses. This means that the evaluation designs must be adapted to the level of experience and the time-availability of the staff. This contrasts with many other evaluations which are conducted by specially hired interviewers.

D. URBAN TRANSPORT

1. Key policy issues

Although forming a relatively small part of early urban projects, transport has come to represent a significant part of urban lending. The main components include: upgrading and maintenance of existing roads, railway lines and stock, buses and buildings; construction of new roads, both community access and trunk roads; acquisition of new vehicles for road and rail transport; traffic management, including traffic light systems; stimulating the growth of private transport companies through credit, training and legislation; and measures to economize the use and importation of fuel.

Urban transport is a field in which precise internal performance indicators have been developed and one of the key issues is to determine how well projects have performed in comparison with these indicators. Some of the common indicators include: passengers per bus per day; kilometers per bus per day; kilometers per bus per day; total staff and staff per bus; light or dead mileage; breakdowns as percentage of buses in operation; buses in service as percentage of total fleet; fuel consumption; and the ratio of revenues to total costs.

In addition most projects have a set of external objectives such as the proportion of the total urban population served, responding to urban growth patterns and economizing fuel costs. More complex objectives may relate to impacts of the transport on urban growth patterns, the distribution of the population and location of employment.

An issue in many projects is the accessibility of the transport programs to the urban poor. This involves factors such as passenger costs, penetration of low-income areas, and directness of transport routes between low-income residential areas and centers of employment.

Some projects seek to stimulate the entry or growth of private transport companies, particularly small bus operators. The evaluation must study the extent to which they do enter, the special problems faced by small private companies and also a cost-effectiveness comparison of public and private bus companies.

A final set of issues relate to the evaluation of the impact of different types of transport projects on the low-income population. (Box 6-3). A number of impacts should be assessed, incuding the following:

a. What happens to low-income families who are displaced as a consequence of road construction? Are they given compensation? Where do they move to, and how are their living conditions affected by the project? It is important to evaluate the impacts on renters and squatters as well as on house owners.

BOX 6-3 SOME IMPACTS OF TRANSPORT PROJECTS ON THE LOW-INCOME URBAN POPULATION

a) Low-income communities must frequently be relocated to permit the construction of a road. As the agencies responsible for road construction often do not coordinate with housing programs, many families will not be provided affordable alternative accommodation. Frequently no records will be kept on what happens to these households so that no-one is even aware of the magnitude of the problem.

b) In Jakarta, a number of kampungs (urban villages) had existed for 50 years or more. Urban development programs frequently included the construction of vehicular roads through these communities. Frequently the roads changed the whole composition of the community, making it dangerous for children to play, bringing in many more outsiders, increasing land values (with possible negative impacts on poor renters) and providing easier access to the city center. However beneficial economically, the project impacts on the existing cultures were more difficult to assess.

c) In Lusaka it was found that the construction of penetration roads had a number of impacts on squatter communities. Stores were able to sell a wider range of products as larger vehicles could now deliver; garbage could be collected by the city trucks and it became easier to work in the city center due to more rapid access.

b. How do new transport systems affect the access of the low-income population to centers of employment?

c. Many roads pass through or near to low-income communities and this will often affect land values. What impact does this have on different sectors of the low-income population?

d. The construction of vehicular roads through communities which previously only had footpaths, can have a dramatic impact on the cultural and economic life of the community. These impacts should be assessed.

2. Evaluation issues

a. Design issues

The nature of most transport projects makes it extremely difficult to make a comparison with a control group. For example, the areas of the city affected by a trunk road are likely to be very different from all other sectors of the city so that comparisons are difficult. Many of the

projects are so large that it is impossible to isolate their effects from the many other changes which are simultaneously taking place. This is complicated further because the projects are often a response to changes which are already underway. For example a trunk road is constructed because the city is already expanding in certain directions or because industry is beginning to move out from the city center.

On the other hand, internal performance and cost-effectiveness comparisons of different transport systems and modes is often easier to conduct as the required information is more accessible.

A difficult but important issue, which is usually not studied, is the impact of major tranport projects on the low-income population. The construction of new roads will often require the relocation of low-income communities. Even though in some major cities low-income households may have been forced to move 3 or 4 times as the city expands, there is usually no record of where they move to or how their living conditions have been affected by the move. An important evaluation design issue is to identify the families who are forced to move and to determine where they move to and how their new living conditions compare with their previous ones. The study can be done in two main ways. On the micro level the attempt can be made to follow individual households, and to revisit them in their new location. In practice this is extremely difficult to do on a large scale. The second option is to request the national statistical office or one of the universities to include questions on migration patterns within the city in future city-wide studies. In this way it may be possible to determine (say 2 years later) in which sectors of the city these families are now living. This approach can only be used when a very large number of families have been relocated, as the sample would not pick up a sufficient number of households from small communities.

In order to evaluate impacts on communities which have been penetrated by, or are close to new roads, a number of research options are available. One option is for a participant observer to live in the community and describe the process of change. A second is to select a panel of households who will be periodically revisited over a period of years. A third is to interview a sample of households before the project begins and then once or twice after the project is underway. Where a quantitative approach is to be used, this should be combined with qualitative techniques to help understand the meaning of the results and to ensure reliability of the data.

b. Measurement issues

Well established techniques exist for collecting and evaluating the types of data required for performance monitoring and cost-effectiveness analysis. Although care must be taken to check the reliability of data, it is expected that the quality of the data will in general be more reliable for transportation than for many other types of projects.

More difficulties potentially exist in trying to obtain data on the impacts of the projects on the affected populations. We have already mentioned the problems of locating families who have been forced to move. In many cases relocation is a controversial political issue so that respondents may be reluctant to give information for fear of reprisals.

3. Applicability of the monitoring and evaluation framework

With respect to monitoring, the standard procedures described in Chapters 2 and 3 are directly applicable. As in previous cases it is important to develop a model of the project implementation process which can serve as the conceptual framework for design and interpretation of both process monitoring and impact evaluation. It is likely to prove difficult to use a quasi-experimental design for impact evaluation, due to the size and complexity of many projects and the difficulties of locating a comparable control group. This means that in many cases, reliance will have to be placed on qualitative and descriptive techniques.

In contrast to health and employment projects, major transport projects are designed to affect the social and economic life of the whole city. This means that greater attention must be paid to macro-level analysis than was the case with the previous types of project.

E. MUNICIPAL AND INSTITUTIONAL DEVELOPMENT

1. Key policy issues

As projects move out from the capital into secondary cities, the issues of the capacity of the municipality to manage and finance the projects becomes crucial. A number of inter-related issues are involved:

a. How can the municipality generate the required counterpart funding? Many municipalities have traditionally had very limited sources of financing and have been largely dependent on transfers from the central government. This was often a deliberate policy to ensure control by

the national government. Consequently many cities have had limited experience in how to generate and manage their own revenue sources. In many cases the available financial sources are limited because central government has pre-empted many sources such as income tax and taxes on sales and businesses. Frequently the only option will be some type of land or property tax. The evaluation must study the effectiveness with which new revenue sources are identified and the efficiency with which they are utilized.

(b) The problems encountered in the attempt to use these new sources of revenue should be assessed (Box 6-4). In some cases the problems are technical (for example lack of qualified land surveyers) but in many cases they are also political. Local politicians may oppose the application of land taxes to the poor, and the wealthy are often able to sabotage or evade efforts to tax them more heavily.

BOX 6-4 EXAMPLES OF PROBLEMS WHICH ARISE WHEN MUNICIPALITIES TRY TO DEVELOP NEW REVENUE SOURCES

a) Muncipalities are constantly being urged to use tougher sanctions to pressure families to pay for services or housing. One of the most logical is to threaten eviction or cutting off of services. However, these measures tend to be strongly opposed by local politicians who will use their influence to stop the measures being implemented.

b) Many municipal finance departments do not have the resources to keep service payment records up to date. This means that it is not possible to identify and sanction people who are in arrears, or even to make it easy for people to pay on time.

c) In many cities the major defaulters on payment for water, light and other muncipal services are other public agencies. Often there are no sanctions which the service agencies can use against these powerful offenders.

d) In many cities, efforts to raise bus fares or food prices to economical levels have resulted in major riots or strikes which have resulted in the measures being withdrawn.

c. How can the incidence of new sources of revenue generation on different socio-economic groups, be assessed? Given high levels of tax evasion and the difficulties in developing effective collection systems, the incidence of the tax burden may be significantly different from what had been intended.

d. The technical capacity of the municipality to design and manage complex development projects.

e. Problems in recruiting, training and retaining qualified staff.

f. How can the technical capacity of a municipality be assessed, and do objective criteria exist for comparing the capacity of different municipalities? Can changes in municipal efficiency be assessed over time?

g. Is it possible to develop a typology of municipalities (based on size, revenue etc) so as to facilitate more systematic comparisons between municipalities?

Urban development projects also require that the national institutions responsible for planning and implementing projects be strengthened. In many countries urban development policies have never existed in any practical sense and each ministry or agency has been responsible for developing its own sectoral programs with very little overall coordination between them. Often the Ministry of the Interior or the office of the president is responsible for resolving conflicts between the different ministries but does not take the initiative in developing overall policies. One of the main problems in the implementation of more complex, multi-city programs has been the inability of the central government agencies to manage the projects and serious delays or bottlenecks have developed. Some of the issues involved in the promotion of institutional development are the following:

a. Description of the current relationships between the main agencies involved in urban development. How are their efforts coordinated and what are the main problems which currently exist in the planning and implementation of multi-sectoral projects?

b. What are the main causes of bottlenecks and delays? To what extent are the problems technical, caused by lack of staff or political?

c. To what extent are the financial problems of the municipalities caused by the way in which the national agencies operate, and what could be done to alleviate some of these problems?

2. Evaluation issues

a. Design issues

Monitoring the impact of municipal and institutional development issues is frequently complicated by problems of access to the required information. Municipalities have limited capacity for data collection and documentation so that even basic monitoring information may not be available or is subject to long delays. There is also a conceptual problem in that this field is relatively new and there is little agreement on the appropriate set of indicators for measuring muncipal effectiveness.

Assessing the impacts of municipal development projects is even more difficult. A municipality is subject to many political and economic influences and is in a state of constant change. Consequently it is extremely difficult to identify the contribution made by any particular set of interventions. For example, to what extent is improved revenue generation a result of the training and technical assistance received under the project, and to what extent is it caused by: the election of a new mayor, influx of new industries, a political conflict with the state government, purchase of a new computer and a prolonged drought in the interior of the state which has accelerated rural to urban migration? When only one city is being studied, much of the analysis will innevitably be based on descriptive and qualitative analysis. However, in projects covering a large number of cities it is possible to isolate the impacts of these factors. Box 6-5 gives an example of how this approach was used for the analysis of a 23 city project in Colombia. It then becomes possible to examine the way in which variations in these factors affect indicators of efficiency.

Techniques such as participant observation and intensive case studies help understand the political dynamics of the city. However, in many cases the techniques are more difficult to apply than in the evaluation of shelter projects as it is much more difficult to obtain access to the subjects being studied. Thus, although it may be relatively simple for an observer to rent a room in a squatter settlement and become part of the community, it is much harder to gain access to the city council and to participate in the political deliberations which are being studied.

b. Measurement issues

There is relatively little consensus on what are the best indicators of municipal efficiency, and even less on how the information should be collected. Record keeping is frequently very poor and very slow so that it may be very difficult to obtain the quantitative information on indicators such as progress of cadastral surveys, tax collection performance and the registration of families who will be relocated.

BOX 6-5 ANALYSIS OF THE IMPACT OF CITY CHARACTERISTICS ON THE PERFORMANCE OF URBAN DEVELOPMENT PROJECTS IN 23 COLOMBIAN CITIES

Most analyses of the performance of municipal governments focus on factors such as local political conflicts, the personality of the mayor and the historical and cultural characteristics of the city. The analyst often implicitly assumes that the situation of each city is unique. In a project covering 23 medium sized cities in Colombia, socio-economic characteristics of each city (per capita income, literacy rate, economically active population etc) were compared with indicators of project performance (no. of houses constructed, children attending project schools, numbers of residents participating in community organizations, no. of people employed in the production cooperatives etc). A number of statistical regularities were found. For example: there was a positive correlation between house construction and per capita city income, literacy rates and the proportion of houses having access to piped water. Similarly there was a significant association between the economically active population and the number of patients treated by the community clinics. These examples show that project performance is strongly influenced by some of the structural characteristics of each city.

Source: Instituco Ser de Investigacion "Diagnostico Evaluativo del Avance del Program IPC" Bogota. Colombia. 1984.

It is even harder to obtain reliable information on decision making processes and the reasons for adopting particular policies. Many decisions are highly political and the real reasons are likely to be very different from the official ones. Tax authorities may not readily explain the types of pressures put on them to provide tax exemption for certain industry, nor will the mayor always admit the way in which his upcoming re-election affected decisions about which communities to upgrade. The unravelling of these issues often requires the services of political scientists or other specialists.

3. Applicability of the monitoring and evaluation framework

The indicators and issues involved in evaluating municipal and institutional development are quite different from those discussed in previous sections, and consequently it is necessary to develop a new set of monitoring and evaluation approaches.

Performance monitoring can be applied in the same way as in other types of projects although with problems and delays in obtaining the information, but the approaches to process monitoring are likely to be quite different. One of the main differences is that a municipality is not a single organization with a limited set of objectives. Consequently it is not possible to apply the same type of models to describe the project implementation process.

Municipalities have many different functions so that it is not possible to define a specific set of quantitative objectives which the municipality seeks to achieve, so it again becomes much more difficult to evaluate impacts. The evaluation must be conducted on a more descriptive and policy oriented level.

F. SUMMARY AND CONCLUSIONS

Despite the specific characteristics and issues connected with each type of project, the monitoring and evaluation framework discussed in the earlier chapters would seem to be generally applicable. Municipal and institutional development programs would seem to offer the greatest difficulties in the application of the framework, as their objectives and methods cannot be so precisely defined.

Performance monitoring can be applied relatively easily to all 4 types of projects. In each case it involves the definition of inputs and the specification of intended outputs (objectives).

Process monitoring can be applied relatively well to employment, health and transport but is more difficult to apply to municipal and institutional development, as the latter involves a wide range of different activities and cannot be easily reduced to a single model. It is possible to study the implementation process for different municipal activities but is much more difficult to study for the overall process of municipal government. This does not, however, mean that the analysis will not be useful as important lessons can be learned from the analysis of particular municipal activities.

One approach would be to select a number of typical activities such as: interaction with low-income communities; interagency coordination; tax collection; and supervision of large infrastructure projects. Analysis of these activities would throw very useful light on the general management and organizational style of the municipality.

Impact evaluations face different issues and problems in each area. The quasi-experimental approach is relatively simple to apply in small business programs as their scope is limited and their objectives clearly definable. The main requirement is that the sample of businesses be sufficiently large to adequately cover all of the different types of businesses. The evaluation of health impacts presents more complex conceptual and measurement issues. There is relatively little agreement on indicators or designs so that any evaluation design has relatively little previous experience to guide it. In the field of transport the main difficulty is that most projects are intended to affect an area of the city which is either very large or which is unique, so that it is very difficult to design any type of control group. Also the types of changes which a transport project may produce are largely macro-economic and social and are consequently much more difficult to evaluate as there are very large numbers of changes taking place simultaneously. Probably the most difficult area in which to conduct impact evaluation is for municipal and institutional development projects, as their objectives are very difficult to define and quantify.

ANNEX A: DEFINING AND USING A MODEL OF THE PROJECT IMPLEMENTATION PROCESS

1. DEFINING THE PROJECT MODEL AND ITS ASSUMPTIONS

(a) The elements of the model

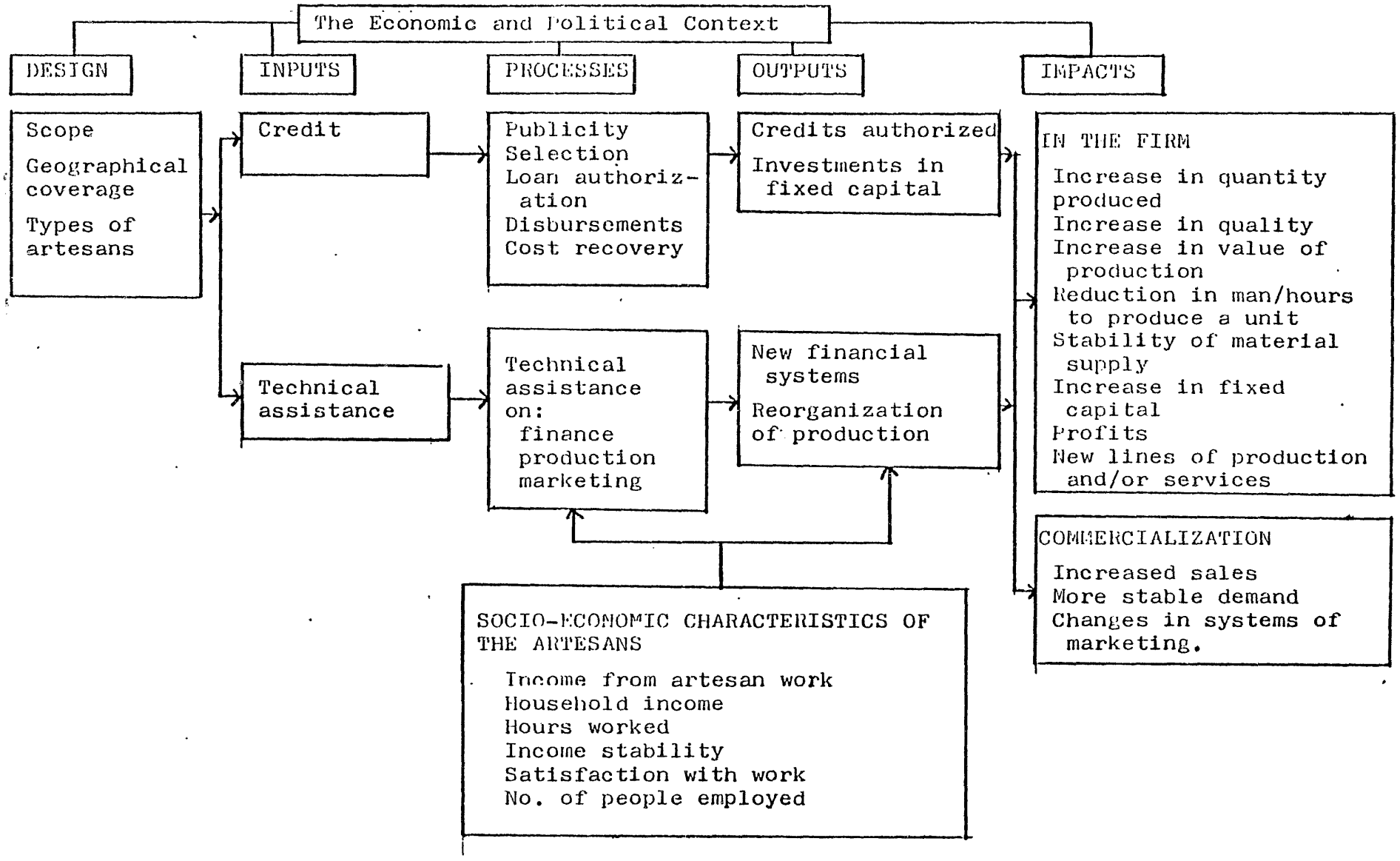
The design of every urban project is based on a model of how resources are used to achieve desired goals. In some cases, the model and its assumptions are clearly stated; in others, the model is implicit. Before a monitoring or evaluation system can be designed, it is essential for the researcher to have a clear definition of the underlying model and its assumptions.

Table A-1 illustrates the essential elements of a project model for a small business credit program. The project is based on a number of design assumptions about scope, geographical coverage, types of businesses to be helped etc. There are a certain number of financial, material and human resources which are the project inputs. In the present example these include credit and technical assistance. It is intended that these inputs will produce a set of outputs which include purchase of fixed capital, new forms of organization and improved financial systems.

The project design specifies a number of processes through which the inputs are transformed to produce the outputs. For example, the loans may be processed through banks or directly by the agency and technical assistance may be given individually or through courses Although they are often not made explicit, there are a number of assumptions which led the project designers to choose one set of implementation processes rather than another. For example, it may be assumed that providing technical assistance through group sessions may have some additional social impacts (such as using group pressures to ensure loan repayments).

The project is not implemented in a vacuum, and the design also includes assumptions about the ways in which the implementation process and the final outcomes will be affected by the project environment and by the characteristics of participating families. For example, designers often make assumptions about the prior experience of participants in business, their need for certain types of technical assistance, and their ability and willingness to pay for different types of services. Similarly, there are assumptions about the likely reactions of local political and community organizations, the forms of cooperation from other government agencies and the effects of other market trends.

Table A-1 CONCEPTUAL FRAMEWORK FOR AN IMPACT EVALUATION: EXAMPLE OF AN ARTESAN CREDIT PROGRAM



The physical and financial outputs are not the final objective of the project, but rather a means to achieve a certain set of impacts. Some of these impacts will occur at the level of the target population, (impact on income, employment, health, etc.) whereas others are intended to have a more general impact on the city. Examples of the latter are the impacts on the low income population in general, on wholesale markets and on prices of raw materials. Although many of these objectives will be explicitly stated in the project design, others are implicit and have to be identified and made explicit by the researchers.

The model also shows that there are assumptions about the ways in which these impacts will be affected by the characteristics of participant households and by the external environment. For example, it may be assumed that female and male headed households will respond differently to the project or that household income will affect the types of outcome. There are similar assumptions about the ways in which the external environment affects outcomes.

Before the evaluation begins, it is essential to prepare a detailed and very explicit model. This model should:

(i) List all inputs.
(ii) Identify underlying assumptions which guided the choice of inputs.
(iii) Specify the processes by which the inputs will be transformed into outputs.
(iv) Identify assumptions which led to the choice of implementation methods.
(v) Identify all assumptions about how the processes and outcomes will be affected by household characteristics and the external environment.
(vi) Specify all outputs.
(vii) Specify all of the impacts the project is intended to produce on the target population.
(viii) Identify assumptions about how project impacts will be affected by household characteristics and the external environment.
(ix) Specify all impacts which the project is expected to have on the low-income population, urban development, and housing policies.
(x) Identify the ways in which the impacts will be affected by household characteristics and the external environment.

(b) Sources for specifying the model and its underlying assumptions

Information on inputs and outputs, and to some extent, on impacts and processes, can be obtained from project documents. In addition, it is necessary to conduct interviews with key project staff. In many cases staff will find it difficult to make explicit many of the assumptions on which the project is based. It will usually prove helpful to construct a model similar to that give in Table A-1 and to gradually fill in the details. At each stage, project planners are asked:

- why were these particular inputs chosen?
- why did you decide to use this process rather than another?
- how do you expect this particular impact to be produced?
- how do you expect the characteristics of households to affect their participation in the project or the impacts the project produces on them?

Once the model and its assumptions have been outlined, a meeting should be held with planners to review the model and its assumptions. Often this will be an iterative process with a number of modifications being made before general agreement is reached.

(c) The use of the model in different types of evaluation

Table A-2 shows the different parts of the model which are used in the main types of monitoring and evaluation:

(i) Performance monitoring is mainly concerned to monitor the use of inputs (were contractors hired? were materials delivered to the stores, etc.), and outputs (how many houses have been built, how many loans given, how many small business started, etc.).

In this case, there is no need to test assumptions but simply to check on whether resources have been mobilized and the required outputs produced. When it is found that either of the above have not happened (or have not happened according to schedule) then other types of evaluation will be required.

(ii) Process monitoring is concerned to assess the efficiency with which project activities have taken place, and how this has affected the achievement of the specified outputs. Some of the typical questions which are asked in this type of study include:

* how efficiently were project participants selected?
* did the selection criteria and procedures ensure that the project was accessible to all sections of the target population.
* how efficient was the use of community labor in house construction?
* how well did cost recovery operate and what were the main causes of delays and defaults in repayments?

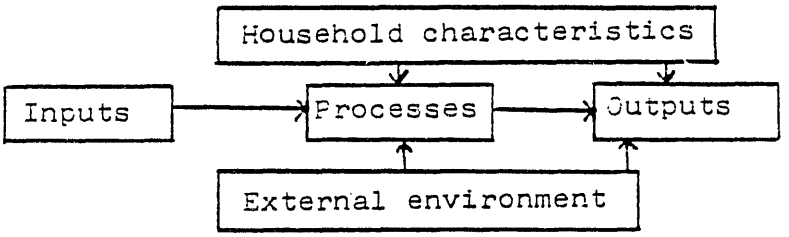
In the process evaluation it becomes necessary to look at the effects of household characteristics and the external environment. For example:

* Are there any differences in loan repayment rates by male and female headed households?
* Does household income affect willingness to request a materials loan?
* How does the action of local political groups affect the land acquisition process?

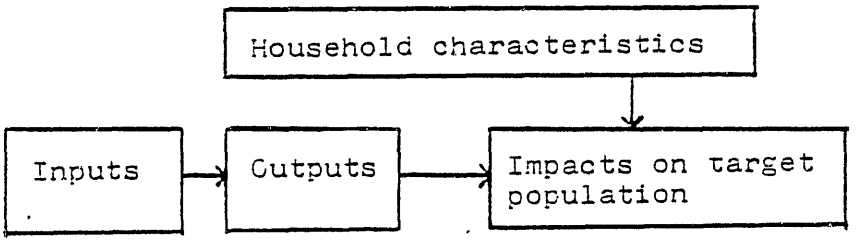
Table A-2 THE COMPONENTS OF THE PROJECT MODEL WHICH ARE STUDIED IN DIFFERENT TYPES OF MONITORING AND EVALUATION

1. IMPLEMENTATION MONITORING

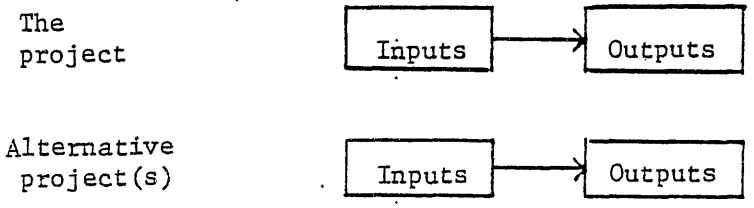
2. PROCESS MONITORING



3. IMPACT EVALUATION



4. COST-EFFECTIVENESS ANALYSIS



(iii) Impact Evaluation is concerned to estimate the impacts which the project has produced on the social and economic conditions of participating households. This type of evaluation tries to relate the degree of access to project inputs and outputs on the changes produced. It also focuses on the effects of household conditions on the degree and direction of changes produced.

(iv) Cost-effectiveness analysis compares alternative projects in terms of the amount of output produced for a given cost.

2. DEFINING MEASURABLE OBJECTIVES

A frequent problem in performance monitoring is the lack of any clearly defined objectives against which actual performance can be measured. In some cases, the organization may prefer its objectives to be kept vague in order to avoid accountability, but in most cases, the problem is that no one has ever tried to specify exactly what the program hopes to achieve. One of the first tasks of the evaluator must be to ensure that agreement is reached on a clearly definable and measurable set of objectives.

The objectives should be defined with relation to the program model discussed in the previous section. Based on this model, 4 main sets of objectives will be identified:

* Implementation objectives
* Process objectives
* Micro-level impact objectives
* Macro-level impact objectives

(a) Implementation objectives

Every project has a set of objectives related to: the use of inputs, the production of outputs, and the time periods and costs related to the process of implementation. Each of these sets of objectives should be clearly specified in terms of the model:

Inputs: The project has available financial (money to subcontract parts of the construction, credit lines for material loans, house purchase, etc.); physical (land, materials, vehicles, etc.); and human resources (administrators, advisers, community development workers, etc.). All of these have been authorized in specified amounts and over specified periods of time. All of these resources should be listed together with the amounts of each and the periods over which they are to be used.

Outputs: Table A-3 gives an example of the way in which the intended project outputs are specified in the loan agreement.

Implementation timetable: For a project to be successful the specified outputs must be achieved within the specified timetable. Delays will increase costs or will slow down the efforts to increase scale to a level comensurate with the city shelter deficit. For this reason, it is essential to specify the implementation timetable. Annex H proposes a network based monitoring system in which the time allotted to each activity is clearly specified, together with the amount of slippage which is permitted before delays will be caused in other components.

Financial objectives: Annex I proposes methods for using network based monitoring to specify and monitor all cost-bearing project activities. The financial monitoring must be integrated with the physical implementation monitoring as delays have serious cost implications.

(b) Process objectives

Table A-4 presents a simplified version of the logical network presented in Annex H. This shows the main stages in the implementation of a sites and services project. The complete network contains 36 steps, but they have been simplified for the present purpose to only include the 10 main components:

* Project design
* Land acquisition
* Offsite infrastructure (roads, water and sewage trunk systems, etc.)
* Tendering for infrastructure and house construction
* Construction of core unit
* Selection of participants
* Material loans
* Completion of habitable units
* Occupation
* Start of cost recovery

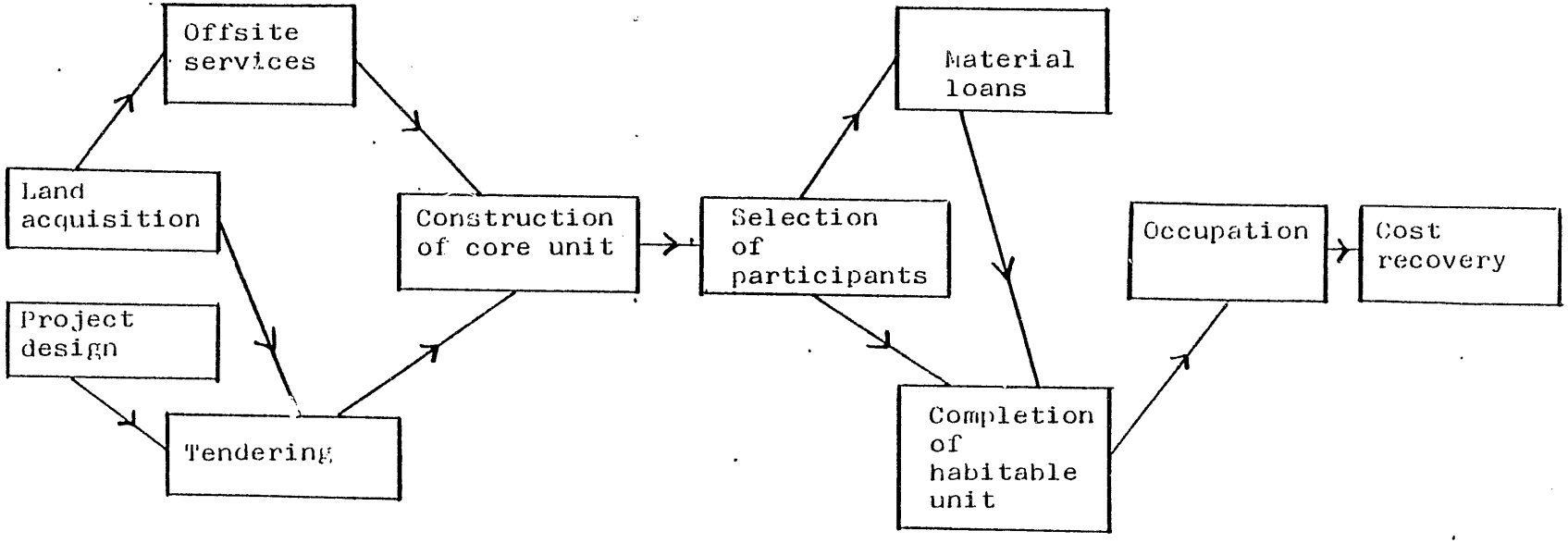
Each of these components has a timetable, and delays in any one will cause delays in overall project implementation. In addition, each component has its own specific objectives relating to questions such as cost, quality, accessibility to the target population. In some cases there will also be objectives related to community participation. The objectives of each component must be specified in a measurable way so that their progress can be

Table A-3 THE SPECIFICATION OF INTENDED OUTPUTS - THE EXAMPLE OF THE LUSAKA UPGRADING AND SITES AND SERVICES PROJECT

COMPONENT	TARGET AT THE TIME OF PROJECT APPRAISAL
PHYSICAL IMPLEMENTATION	
1. Residential units serviced	
a. Squatter upgrading	16924 plots
b. Overspill areas	7588 "
c. Normal low-cost sites	1204 "
d. Normal medium-cost sites	1938 "
TOTAL	28851
2. Community facilities	
a. Schools	20
b. Health centers	3
c. Markets	17
d. Community centers	17
SOCIO-ECONOMIC IMPROVEMENT	
1. Residential units allocated	
a. Squatter upgrading	16924
b. Overspill areas	7588
c. Phased normal (basic sites)	1197
d. Normal low-cost sites	1204
e. Normal medium-cost sites	1938
2. Allocation of sites and services plots to target group	
a. % plots allcoated to households earning less than Kwacha 70 permonth	50%
3. Core units erected	
a. Overspill area	7588
b. Sites and services areas	4339
4. Value of building material loans	Kwacha 5,609,000

Source: Bamberger, Sanyal and Valverde "Evaluation of sites and services projects: the experience from Lusaka, Zambia." World Bank Staff Working Paper No. 548 1982.

Table A-4 SIMPLIFIED VERSION OF A LOGICAL NETWORK SHOWING THE MAIN COMPONENTS (PROCESSES) OF A SITES AND SERVICES PROJECT



monitored. Table A-5 gives examples of how objectives could be defined for the selection of participants and the completion of the habitable unit. The criteria which are used are the same as those used to measure efficiency namely:

* Output
* Speed
* Cost
* Accessibility to the target population
* Replicability in future projects

In addition, organizational objectives, which may not be quantifiable, are included. For example, in the selection of participants it may be an objective to hold group meetings so that participants understand the social as well as the economic objectives of the project. The monitoring reports should address these organizational, or non-quantifiable, objectives.

(c) Micro-level impact objectives

Micro-level impacts are the changes the project hopes to produce in participant households. Table A-6 gives examples of 6 types of impacts the project might wish to produce and of indicators which could be used to measure each one. It is important in the formulation of impact objectives that the objectives are related logically to the project model and are defined in terms of measurable indicators.

We will take the example of employment and income to illustrate these two points. It is hypothesized that the project may affect income through:

(i) Increased rental income from larger and better quality houses with improved access to services.
(ii) Increased income transfers from relatives to help with house construction and mortgage payments.
(iii) Through one of the employment effects described below.

The project may affect employment in the following ways:

(i) Short-term employment opportunities during the process of house construction. It is expected that a high proportion of households, particularly high income and self-employed, will subcontract parts of the construction.
(ii) Increased business for local stores and services due to the higher average income of households in the project areas.

Table A-5 EXAMPLES OF THE DEFINITION OF PROCESS OBJECTIVES FOR 2 COMPONENTS OF A SITES AND SERVICES

PROJECT: SELECTION OF PARTICIPANTS AND COMPLETION OF HABITABLE UNIT

COMPONENT	TIMETABLE	OUTPUT	SPEED	COST	ACCESSIBILITY	REPLICABILITY
Selection of participants	Weeks 101-114	15,000 applicants 5,000 pre-selected 2,000 selected 500 reserves	Average times per applicant: * Pre-sel = 15 mins. * Selection= 90 mins Elapsed weeks: * Receipt to pre-sel = 4 * Pre-sel to sel = 4	150 staff/ weeks Other direct costs = 10,000 pesos	25% selectees have monthly incomes below 250 pesos. - Maximum income = 1000 pesos	Development of computer processing and selection system which could be replicated with same time/speed/ accessibility constraints for projects with 50,000 applicants per year.
Completion of habitable unit	Weeks 126-138	2000 habitable units all of which occupied	Average of 24 weeks to complete and occupy unit from approval of material loan	Costs to household should not exceed 2000 pesos of own resources + loan.	No households should fail to complete due to lack of finance.	Loan and technical assistance procedures streamlined to reduce processing time and cost in future projects.

(iii) Increased business due to improved vehicular access and, hence, more outside customers and easier delivery of supplies.

(iv) Project located closer to employment centers so that employment opportunities increase.

(v) Local business may give preference to job applicants from the project as they believe that house-owners are more reliable workers.

The ways in which these changes can be measured are indicated in the table. Once the researcher understands these objectives and assumptions it is possible to design a study to both measure whether the changes have taken place and to indicate why the outcomes are different than expected. For example, if increased employment opportunities do not occur, it will be possible to evaluate whether this is due to macro-economic factors affecting demand for labor, to the location of the project site or to the fact that employers are less interested in hiring project participants than had been expected. Using this type of model, the information is much more useful to policy makers than a simple statement that "there was no significant impact on employment."

(d) Macro-level impact objectives

In addition to their direct impacts on project participants, most projects are also intended to have some wider impacts on:

(i) Total low-income population (improved access to housing, lower rents, etc.).

(b) Management of the city (increased taxation, improved transport systems, etc.).

(c) Housing and urban development policy (changed attitudes to sites and services and upgrading; revision of housing subsidies; provision of new sources of housing finance; increased involvement of the private sector, etc.).

Table A-6: EXAMPLES OF 6 MICRO-LEVEL PROJECT IMPACTS AND THE INDICATORS WHICH COULD BE USED TO MEASURE THEM.

TYPE OF IMPACT	QUANTIFIABLE INDICATORS
1. Employment and Income	1. Total family income 2. Source of income 3. Income stability 4. Type of employment 5. Number of people working 6. Labor force participation rate of a particular group, eg. women 7. Proportion of self-employed 8. Proportion working in formal and informal employment sectors 9. Geographical location of employment
2. Demographic Characteristics of the Family	1. Family size 2. Age composition 3. Education of household head 4. Proportion of young people attending school 5. Civil status of household head 6. Geographical mobility
3. Housing Costs, Quality and Value	1. House value 2. Construction quality 3. House size 4. Access to water, light and sanitation
4. Health	1. Infant mortality rate 2. Main types of intestinal infection 3. Main types of illness 4. Time lost from work due to illness 5. Time lost from school due to illness 6. Access to medical services 7. Amount spent on medical services 8. Anthropometric measures of weight and height
5. Consumption Patterns	1. Amount spent on housing 2. Amount spent on food 3. Amount spent on clothing 4. Amount spent on transportation 5. Amount spent on medicine and medical treatment 6. Amount saved
6. Community participation and attitudes	1. Number of community organizations in which families participate 2. Number of other families in the project known, and considered as friends 3. Membership in particular types of community organization like political, social and religious, etc. 4. Participation in mutual help programs 5. Satisfaction with the community 6. Satisfaction with ones own social, economic and political situation

ANNEX B: METHODS OF DATA COLLECTION FOR MONITORING AND EVALUATION

1. The importance of the choice of research methods

> "A knowledge of technique needs to be complemented by an appreciation of the nature of research as a distinctively human process through which researchers make knowledge. Such appreciation stands in contrast to the more common view of research as a neutral, technical process through which researchers simply reveal or discover knowledge. Such appreciation requires that we reframe understanding and debate about research in a way that goes beyond considerations of method alone."
>
> Gareth Morgan "Beyond Method" (1983) page 7.

Social science has become increasingly concerned with refining research techniques, but relatively little attention has been paid to the question of how to choose the best technique for a particular purpose. The decisions about which research methods to use can have a profound effect on the types of information which are obtained and the conclusions which are drawn. For example, many studies, for reasons of convenience of analysis, use a precoded questionnaire in which respondents are asked to indicate whether they agree or disagree with a particular statement, or to indicate which of a list of community problems they consider most important. This technique forces respondents to choose between a limited number of predetermined options, and does not permit them to indicate that they are more concerned with other issues or even that the question itself does not make any sense to them. If, instead of using a questionnaire, the researcher had lived in the community for three months and had tried to describe the main concerns which people expressed in their day-to-day activities and in community meetings, it is quite possible that some very different conclusions might have been reached.

In an interesting book edited by Morgan (1983), twenty one authors describe the methods they would use for research on organizations and the assumptions underlying their choice of methods. The book shows how the choice of research method influences the issues studied and determines outcomes, often in ways which are not intended or appreciated by the researcher. As one of the authors states "We never talk about the world - social or physical - only about our construction of it." (Bourgon,1983). The researcher must be aware of the strengths and weaknesses of each method and must constantly seek

ways to validate the conclusions drawn from using a particular technique. In the following sections some of the main research methods are described, together with some of their strengths and weaknesses (See Table B-1).

Many researchers have their preferred methods which they seek to apply to whatever problem they are studying. It is essential that the evaluator adapt his/her methods to the problem rather than seeking a problem which can be studied with the preferred research methods. There is no one ideal method and all techniques have their strengths and weaknesses. Consequently the evaluation researcher is strongly advised to combine several different methods in the evaluation design.

A continuing debate in the evaluation literature concerns the relative merits of quantitative and qualitative methods. Journal articles and books have been devoted to advocating one of these methods over the other and many of the leading evaluation practitioners are identified with one or other of these two, supposedly conflicting approaches. In a recent article Reichardt and Cook (1979) show that the supposed correspondance between qualitative and quantitative research paradigms and qualitative or quantitative methods does not in fact hold in practice. Advocates of the subjectivist qualitative paradigm will frequently use small sample surveys to complement their participant observation and many practitioners of "hard" statistical analysis will use open interviews to help develop and validate their instruments and to assist in interpreting their statistical findings. The two approaches complement each other and should be used together for most evaluation purposes. Quantitative and qualitative methods normally have different objectives as well as having different strengths and weaknesses.

Sections 3 onwards describe the principal methods, starting with qualitative methods and moving through the spectrum to the more quantitative approaches.

2. The use of triangulation to build in consistency checks

No research method is infallible, and consequently the evaluation design should always include consistency checks in which at least two independent methods are used. This technique of developing consistency checks through independent measures is called triangulation. The reason for this name can be seen in Table B-2. In example 1, changes in household income are estimated both from a sample survey and from observing changes in the range

TABLE B-1: APPLICATIONS AND LIMITATIONS OF SOME OF THE PRINCIPAL EVALUATION RESEARCH METHODS

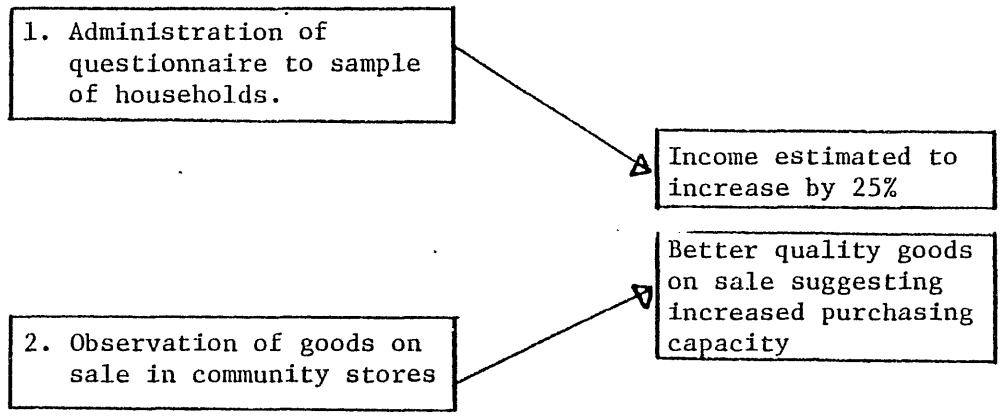
METHOD	APPLICATIONS	LIMITATIONS
1. Quasi-experimental designs	Statistical estimation of project impact. Eliminates many of the spurious claims of project impact which can arise when no control group is used.	Difficult and expensive to use in practice. Does not provide information on causation or the effects of the project implementation process on outcomes.
2. Sample surveys using structured questionnaire	Statistically reliable estimates of the attributes or attitudes of different groups. Eliminates many of the incorrect inferences which can arise from studies based on non-randomly selected small groups.	A structured quetionnaire forces respondents to use a conceptual framework imposed by the researcher. There are many types of information which cannot be obtained reliably with this type of instrument.
3. Intensive case studies of households or groups	Provides a more complete understanding of how a group or household operates and what they feel about key issues. Illustrates how the group/family interacts with other community groups and with external organizations.	The studies can be very time consuming. As the number of cases will often be very small it can be difficult to generalize and to know how representative the findings are. Validity problems exists as much of the information will come from subjective opinions of the observer.
4. Observational and related techniques	Many of these techniques are economical and fast to apply. They can also provide consistency checks on survey and other methods.	There is often a tendency to rely on whatever indicators can be easily observed, even though these may not be the best for a particular purpose. There are often difficult problems in interpreting the meaning of what has been observed.

Table B-1 (continued)

METHOD	APPLICATIONS	LIMITATIONS
5. Secondary data	Existing secondary data can often answer many of the questions which would be covered by a new and expensive survey	There are often problems of reliability or coverage and the information may be out of date.
6. Participant observation	Living in, or maintaining close contact with a community over a period of serveral weeks or months can provide an understanding of how the community perceives a project in a way which could never be obtained with other more formal approaches. Can also provide an understanding of many delicate or complex issues which would not be detected by surveys.	Can be relatively time consuming and expensive. Problems of validity of subjective perceptions of the observer. How representative are the small number of cases selected.

Table B-2: Use of Triangulation to check the consistency of estimates of project impact on household income

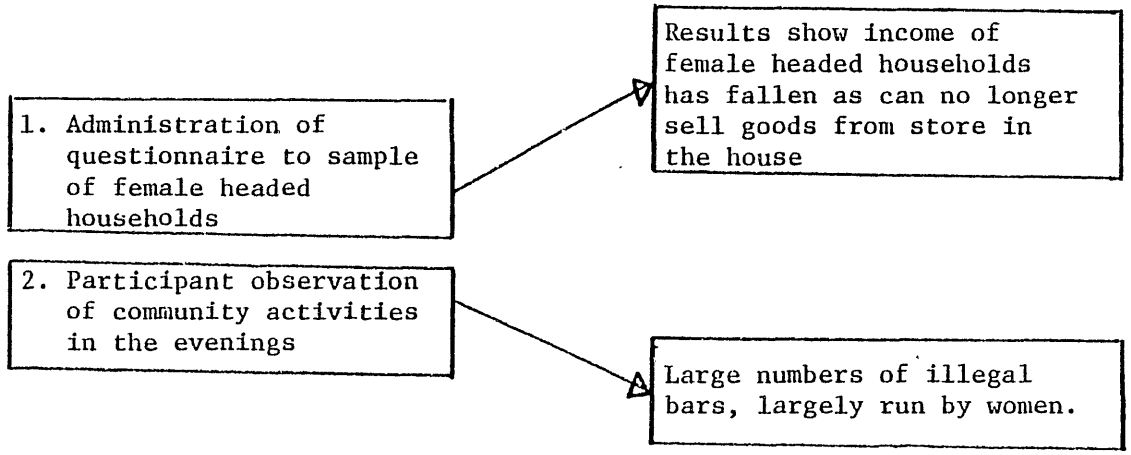
Example 1: Survey and direct observation provide converging (consistent) estimates.



Conclusion

Both estimates are consistent and show income has increased

Example 2: Survey and direct observation provide diverging (inconsistent) estimates.



Conclusion

Women receive a considerable proportion of their income from illegal sources which they do not declare to interviewers or project management. This suggests incomes of at least some women have fallen less than it at first appears.

and quality of goods on sale in community stores. In this case both estimates converge or triangulate towards the same conclusion - namely that household income has increased.

In the second example, a sample survey and participant observation are used to estimate changes in the income of female headed households. According to the survey, the income of female headed households has declined. However, the participant observation study reported that women engaged in a number of illegal or quasi-legal forms of income generation such as beer brewing and the running of bars. In this example, the two estimates are divergent, with the survey stating that incomes have gone down and the participant observation suggesting there may be important sources of income which have not been reported. As beer brewing and the ownership of unregistered bars are both illegal, it is not surprising that neither of these sources of income was reported in the survey. This is an example where the researcher would be advised to consider modifying his survey instrument or using alternative methods of estimation.

3. Participant observation and related ethnographic techniques

Partly as a reaction to some of the problems encountered in purely quantitative evaluations, a number of qualitative evaluation approaches have become popular in recent years. One of the best known is participant observation which is defined by Bogdan and Taylor (1975) as:

> "Research characterized by a period of intense social interaction between the researcher and the subjects, in the mileau of the latter. During this period, data are unobtrusively and systematically collected."

The evaluator lives in or is closely involved with the community or group being studied over a considerable period of time in order to understand the culture of the group and the meanings its members give to the subject being studied. In the case of an urban development project, the purpose is to understand how people perceive the project, who does and does not benefit and what effects it has on individual families and the community in general. The approach is based on a long sociological tradition of "verstehen" (the orginal term used by Max Weber), in which the researcher seeks to understand the subjective meanings which people give to the organizations and activities in which they are involved and to the world in which they live.

Most of the data is obtained through unstructured interviews and through direct participation in, and observation of group activities. The researcher may relate to the community or group as a normal group member (for example the study by Lisa Peattie, 1969 on the way in which the community in which she was living in Venezuela organized itself to combat the construction of a sewage outlet); as a clearly identified outsider (for example the study by Salmen, 1984 in which he lived in a community in La Paz which was in the process of being upgraded); or under some false pretext (as when researchers pretend they wish to join a religious sect they are studying).

An interesting approach to participant observation in urban projects is the study by Salmen (1983) in which he lived in World Bank upgrading and sites and services projects in La Paz and Guayaquil. Among the many insights his study was able to contribute are the following:

* It was found in one project that community leaders were opposed to the project on political grounds and were deliberately misinforming the community as to the nature of the project. Project management had not been aware that this was the reason it had never been possible to initiate the project.

* In another project the community leaders, although well intentioned, were all property owners from higher income groups and did not represent the interests of poorer renters

* A much higher level of frustration with project delays was found than had been expected by project management.

* In the unstable economic and political contexts of these two cities, property ownership offered one of the few opportunities to achieve a sense of personal security. Consequently, an extremely high value was placed on ownership.

* Living in one of the newly designed houses revealed many of its design shortcomings (lack of ventilation, poor drainage, high noise level from neighbors) in a way which was never appreciated from casual visits.

Participant observation has a number of advantages. First, the observer is better able to understand the meanings which individuals and groups give to the subjects being studied (for example, attitudes to the installation of piped drinking water, a community development center etc) and is less likely to fall into some of the kinds of misinterpretation which may arise if the study is based simply on responses to a structured quesitonnaire. Second, the observer is better able to study delicate issues (such as attitudes to

community leaders or the origins of community conflicts) which people may be unwilling to respond to directly in a questionnaire. Third, the observer is able to observe and evaluate processes as well as particular events. This is extremely difficult to do with a questionnaire. Fourth, it is possible to study feelings and attitudes which respondents may have difficulty in verbalizing. Finally, the approach is flexible so that it is possible to adapt to changing circumstances in a way which cannot be done with structured surveys.

Despite its recent popularity, participant observation also suffers from a number of limitations. First, it is difficult to identify and control for observation bias introduced by the researcher. As the observer usually works alone it is much harder to introduce the kinds of consistency checks which can be used to control for interviewer bias in sample surveys. One aspect of this problem is that the researcher may bring his/her preconceptions to the study and may inadvertently structure the conversations or the observations to support these preconceptions. Second, it is difficult to evaluate the distortions caused by the presence of the observer. Behaviour is likely to be changed by the presence of an outsider in ways which are difficult to assess. Third, the observer can only observe part of the reality being studied. It is not possible to attend all meetings, or to be present during all activities of the subjects being studied. In fact there is likely to be selectivity as it is much easier to observe certain activities (for example public meetings) than others (negotiations between the leaders of rival political groups for example). Again it is difficult to be aware of or to control for the biases which occur in this way. Fourth, the largely subjective way in which information is collected and reported makes comparability difficult. Thus it is difficult to systematically compare reports produced by different observers. This makes quantification difficult and limits the utility of participant observation in large scale projects. Difficulties of comparability and quantification lead to a fifth problem which is the inability to use many of the statistical techniques used to match groups and control for spurious causality. For example, the observer may notice that two groups of residents respond in different ways to a project and

he/she may hypothesize that the differences are due to cultural factors. However, statistical analysis might have shown that after controlling for socio-economic variables such as income, age and education the differences no longer exist.

A recent review article on the US experience in the evaluation of social development projects during the past 20 years (Rossi and Wright, 1984) highlighted some of the limitations on the use of qualitative methods for impact evaluation.

> "It is however,equally clear that qualitative evaluations have their limits as well. However inexpensive they may be for single, small-scale projects, they are very expensive and not very sensible approaches to the evaluation of fully developed programs that have quite specific goals. Qualitative evaluations are very labor intensive and cannot be used on very many sites except at considerable cost. Furthermore, qualitative approaches rarely provide estimates of the effects that are either very precise or free and clear of possible confounding factors. Indeed, the only large-scale programs to which qualitative approaches were applied had vaguely stated goals e.g. Model Cities (Kaplan 1973) and revenue sharing (Nathan et al 1981). In these evaluations, the findings were composed more of descriptions of program operation than of assessments of programs effects."
>
> (Rossi and Wright,1984 page 343)

4. Direct Observation

In their classic study, Webb and Campbell (1966) suggested a wide range of "unobtrusive measures" which could be used as indicators of more complex processes or events. Examples of such indicators might include:

* the amount of wear on steps as an indicator of the use which is made of a community facility;
* analysis of garbage dumps as a source of information on consumption patterns.
* Using the number of tin roofs or tin utensils as an indicator of wealth.
* Using the types of washing left out to dry as an indicator of consumption patterns and wealth.
* Types of commodities on sale in local stores as an indicator of the economic level of the community.

The problem in the use of these indicators is: How do we know what is a good indicator? For example, a common issue in urban housing projects is to determine whether the project benefits are reaching the intended income groups or whether many of the benefits are going to higher income groups. A number of indicators have been proposed which reflect the consumption patterns and economic level of the project population. Some of the possible indicators which can be used to detect the presence of higher income households in the project include: iron window grilles, exotic plants which do not naturally grow in the region and which have been purchased, and expensive furniture. The problem is that there are at least two possible interpretations of the meaning of an increase in the number of observable middle class symbols. The first possible explanation is that middle-class (higher income) households have moved into the project. However, a second possible explanation is that poorer households may have begun to adopt some of the consumption patterns of their better off neighbors.

In cases such as this, it is usually necessary to investigate further in order to determine what the indicators really mean. This does not invalidate the usefulness of these indicators, but simply points out that their use and interpretation is often more complex and expensive than some writers would suggest. The use of rapid observation techniques should not be used as an excuse for avoiding scientific rigor. The researcher should always state the assumptions upon which the selection and interpretation of indicators was based, thus permitting independent verification. Where a number of similar projects are planned it becomes worthwhile to invest resources in the careful selection and validation of a set of indicators which can then be used to evaluate a number of similar projects.

Observational indicators are useful for the evaluation of the physical conditions of the community. An observation guide can be constructed with items such as the following:

* Conditions of the streets.
* Cleanliness of streets and public areas.
* Maintenance and use of community facilities.
* Construction materials used in houses.
* Proportion of houses which are occupied, complete, in process of construction.
* Frequency of public services such as street lights, bus stops, public telephones, etc.

When judgements have to be made, for example on cleanliness of the streets, it is important to give precise instructions and if possible to provide photographs illustrating what is meant by "clean", "reasonably clean", etc.

The inclusion of photographs in the evaluation report is also a useful way to illustrate the conclusions. A good photograph is often much more effective than a table or a page of text. However, there is a danger of selecting photographs which are dramatic rather than typical so care must be taken to avoid the use of photographs which can misrepresent.

5. Informal group discussions

The organization of informal group meetings can often produce information not easily elicited from individual interviews. The group discussions are less directed by the researcher and hence more spontaneous. Participants tend to stimulate and also correct or challenge each other so that many issues are brought up which would not have been addressed in an individual interview. It also becomes easier to address organizational issues and group attitudes, as the information and points of view of many different people can be compared.

One method is based on asking a set or "What if?" questions to gradually focus on the rules governing organizational or community behaviour. For example, the following questions could be use to provide insights into the ways in which organization "X" interacts with the community:

"What would happen if the roads in this sector were constantly flooding"?

"Would you contact agency X"?

"What if the flooding was in sector Y; would that make any difference"?

What if the flooding occurred in the year of an election"?
Etc., etc.

6. Unstructured("open") interviews

Unstructured or "open" interviews are a technique in which informal conversations are conducted without the use of a questionnaire. The interviewer has a checklist of questions which are covered during an informal conversation rather than in a question and answer session. The interviewer may refer to the checklist and possibly even take a few notes, but usually the interview report will be written up afterwards so as to keep the situation as

natural and informal as possible. In those cases where it is necessary to analyze language usage, the interviewer may decide to use a tape recorder. Whenever a formal questionnaire is to be used, unstructured interviews should always be included as a preliminary stage of the research design to understand the key issues as they are perceived by the population being studied and to ensure that the right questions are being asked. Further unstructured interviews may be conducted at the end of the study to help interpret the findings.

Unstructured interviews usually cover "key informants" representing the main groups involved in or affected by the project. A number of "ordinary people" should also be included so as to ensure that their point of view is included as well as that of the more vocal leaders. Typical key informants might include:

* local community and political leaders
* local religious leaders
* storekeepers and other local businessmen
* local government officials
* academics
* newspaper reporters

The problem with the use of key informants is that information almost always contains a bias. Political leaders, for example, are likely to praise programs when their party is in power, and criticize activities of the opposition. Similarly, government officials may play down community dissatisfaction. Minorities and weak groups will tend to be under-represented. The following are useful guidelines in the interpretation of data from key informants:

* Always note the position of the informant (local politician, storekeeper, property owner, etc.) and try to estimate how this position may bias the information.

* Obtain several independent opinions on all major topics. Where the information is inconsistent, try to understand why this is so.

* Randomly select people who are <u>not</u> key informants. Information from these people should be compared with opinions of supposed community leaders. Also try to learn from these respondents how they feel about the leaders.

* Analyze the characteristics of key informants to determine which groups are not represented. For example:

- Have women been interviewed?
- Do interviews include low-income as well as better off respondents?
- Have all tribal, ethnic, or religious groups been covered?
- Have people from all geographical sectors of the community/ city been consulted?

* Remember that people who claim to be community leaders or spokesmen almost never represent the whole community but only some part of it. Never accept the opinions of leaders about a project or similar issue as truly representing the opinions of all of the community.

7. Structured questionnaires

A structured questionnaire is one in which all or most of the questions are precoded so that the respondent must select from among a number of predetermined categories. For example, during an unstructured interview the respondent may be asked to state what he/she considers to be the main problems facing the community; whereas the structured questionnaire would ask the respondent to indicate which of a list of predetermined problems he/she considers most important.

For an impact evaluation the questionnaire will usually contain 4 types of questions:

(a) Classification information: participant status, type of project house, who is being interviewed etc.

(b) Exposure to project variables: indicators of the types and amounts of services received or to which exposed (for example, amount of housing credit, distance from communal water tap etc.).

(c) Outcome variables:indicators of project impact.

(d) Intervening variables: attributes of the household, the business, the community etc which might affect project outcomes.

Questions can be asked in a number of different ways. They can be open or closed; they can be asked verbally by the interviewers or the respondent can be asked to complete the information by indicating his choice from a list or by writing replies on the questionnaire. In more sophisticated questionnaires it is possible to include an attitude scale where the respondent is asked to chose from a number of alternative responses which have been ranked or where he has to indicate the degree of agreement or disagreement with a statement. A well constructed attitude scale permits a much greater degree of sophistication in the

analysis. However, in order for the results to be valid and meaningful a great deal of time and effort must be put into the design of the scale. Many attitude scales appearing in evaluation questionnaires are of almost no use as they have not been designed properly.

The process of designing and validating a questionnaire is also time consuming. The main stages are the following:

(a) Define carefully the objectives of the study, the key issues and the types of variables to be studied.

(b) Conduct unstructured interviews, participant observation and direct observation to understand the meanings which respondents attach to the concepts being studied, and to test out various ways of asking the questions.

(c) Prepare a first draft of the questionnaire and check the list of questions against the objectives of the study. This is to ensure that all questions are covered and also to eliminate any unnecessary information.

(d) Prepare an interview guide explaining how the survey is to be conducted and how each question is to be asked.

(e) Conduct a pilot test of the questionnaire. The survey designers must conduct some of the interviews themselves.

(f) Discuss the results of the pilot study in detail with the interviewers and make whatever changes are suggested in the instrument.

(g) With a complex questionnaire it may be necessary to conduct several pilot tests.

(h) Organize a training session for the interviewers. This should include conducting a number of test interviews, the results of which will be reviewed in the training session.

The use of a structured questionnaire offers a number of major advantages in evaluation research:

(i) It permits a rigorous comparison of responses obtained by different interviewers or in different communities or cities. Techniques exist for testing the reliability of the information and for determining whether bias is being introduced by any of the interviewers.

(ii) The information can be quantified, thus permitting the use of more sophisticated analytical procedures.

(iii) The comparability of responses means that changes can be measured over time or comparisons made between different groups.

(iv) Many types of spurious causal relationships can be controlled for statistically in a way which is usually not possible with unstructured methods. For example, it is possible to control for social and economic attributes of the household or community to determine whether observed differences are in fact due to the project or simply to the initial differences between the groups.

Despite its wide use and potential advantages, structured questionnaires have a number of problems:

(i) The framework of the instrument is very rigid and it is usually not possible for the interviewer to record any deviations from the standard questions. Respondents may wish to give replies other than those in the list of options or they may wish to give information not included in the survey at all. Usually this information cannot be analyzed. In extreme cases this could mean that even though all of the questions have been answered, the information is almost useless.

(ii) An interview is a very unnatural situation, not conducive to gaining confidence. In many development situations the interviewer will be classifed as a representative of the government who is collecting information which may be used to the disadvantage of the respondent. These factors make it difficult to gain the respondent's confidence and to obtain his/her full cooperation.

(iii) Due to this lack of cooperation, some of the key information may be inaccurate or falsified. Respondents may not, for example wish to give accurate information on their incomes or how long they have been living i the city. Unless the interview is complemented with other techniques, there is a danger of inaccurate information.

(iv) It is particularly difficult to obtain information on subjects which are socially delicate (such as sexual behaviour or child rearing) or which are politically sensitive (such as opinions about community leaders or political participation).

(v) A questionnaire forces people to verbalize their responses. This can be very difficult, particularly for groups with a low educational level. Consequently questions such as: " How do you feel about .....", "If you were able to choose between these types of house

which would you prefer .....", "What are the main things you dislike about this community..." may be very unreliable.

(vi) The questionnaire is not able to compare people's behaviour with their verbal responses. A mother may assure the interviewer that children are only fed the correct types of food. A participant observer might have observed that in fact childrens' diets were much less carefully controlled.

(vii) The design and application of a large scale survey will often be expensive.

(viii) Analysis of the results may be very time consuming and expensive. This can be particularly problematic in countries where survey analysis infrastructure does not exist.

8. Secondary Sources

Useful information can frequently be obtained from sources such as project records, government agencies and previously conducted studies. In some cases, newspapers may also provide useful information (for example, prices of rental property, political, religious and social activities, etc.). The researcher should always check for the existence of secondary sources before planning to collect new information.

FURTHER READING

1. Participant Observation

Oscar Lewis The Children of Sanchez (1961). A classic study in which participant observation and tape recorded interviews are used to describe the life of a family in a Mexican slum.

Eliot Liebow Talley's Corner (1967). Study of the life of a group of black males in Washington, DC. The structure of the book around topics such as work and family life provides an easily digestible way to present the findings of the study.

Lisa Peattie The View from the Barrio (1969). Description of the life and organization of a low-income community in Venezuela in which Peattie lived while working as a consultant on the development of a new city. She becomes involved in the efforts of the community to obstruct the construction of a sewage outlet which would contaminate their water supply. The study is interesting as she was actively involved as a community resident rather than just as an outside observer.

George McCall and J.L.Simmons (Editors) Issues in Participant-Observation: A Text and Reader (1969). Comprehensive review of methodological and validity issues as well as an extensive presentation of case material.

Lawrence Salmen Participant Observer Evaluation of Urban Projects in La Paz, Bolivia and Guayaquil,Ecuador (1984). Summary of the experiences of the first systematic attempt by the World Bank to use participant observation in the evaluation of urban shelter projects

2. Unobtrusive Measurement and Direct Observation

E.Webb and Donald Campbell Unobtrusive Measures: Non-Reactive Research in the Social Sciences. (1966). Presentation of a wide range of unobtrusive methods which can be used for evaluation and other types of social research.

Michael Patton Qualitative Evaluation (1980). One of the leading, and most readable, exponents of qualitative evaluation.

Joseph Valadez Non-Survey Techniques in the Evaluation of Urban Shelter Programs (1982). An application of the Webb and Campbell approach to the evaluation of urban projects.

3. Questionnaire Design

Stephen Malpezzi, Michael Bamberger and Stephen Mayo Planning an Urban Housing Survey: Key Issues for Researchers and Program Managers in Developing Countries (1982). Explanation of how to design an urban housing survey. Includes a model questionnaire.

Michael Bamberger and Julie Otterbein Designing a Questionnaire for Longitudinal Impact Studies.(1982). Explanation of the stages involved in the design, testing and application of an impact survey. Includes a model questionnaire.

David and Chava Nachmias Research Methods in the Social Sciences (1981). A good textbook covering the main data collection methods and analytical procedures.

4. Quantitative versus Qualitative Issues

Charles Reichardt and Thomas Cook Beyond Qualitative Versus Quantitative Methods (1979). Presentation of the main issues involved in the debate and a demonstration that quantitative and qualitative methods complement each other much more than many people seem to believe.

ANNEX C: MANAGERS GUIDE TO THE DESIGN AND IMPLEMENTATION OF A MONITORING AND EVALUATION SYSTEM

This section provides a guide for the managers of project executing agencies on the main stages and decisions in the design and implementation of a monitoring and evaluation system. Seven main stages, which are approximately sequential, are defined. Most of the issues discussed in this Annex are also applicable at the level of the sectoral and national development agencies. These stages are summarized in Table C-1.

1. INITIAL DECISIONS ON THE SCOPE, ORGANIZATION AND OBJECTIVES OF THE EVALUATION

1.1 When to begin planning the evaluation:

Planning for the monitoring and evaluation should begin during the early stages of project appraisal. When the evaluation design is not defined until the project launch, problems and serious delays often arise as staff and budget have already been committed. In several cases the process of staff recruitment has taken over a year (due to the need to negotiate and create new staff positions) so it is essential to begin the process as early as possible.

1.2 Defining the objectives of the evaluation

The following are some of the main ways in which evaluation data can be used. The manager must decide the relative importance to be given to each of these objectives:

(a) To provide regular information on the progress of each element of the project, and to compare the progress with stated objectives in terms of time, volume and cost.

(b) To provide constant updates and revisions of completion dates and costs and disbursement schedules.

(c) To identify potential problems and to suggest possible solutions.

(d) To provide constant feedback on the efficiency of project implementation and to suggest improvements which could be made.

(e) To provide constant feedback on the effectiveness of the project in achieving its stated objectives.

(f) To provide estimates on project impact in areas such as income generation and employment.

Table C-1: MAIN STAGES IN THE PLANNING AND IMPLEMENTATION OF AN EVALUATION

A CHECKLIST FOR PROGRAM MANAGERS

1. INITIAL MANAGEMENT DECISIONS

* When to begin planning
* Scope and objectives
* Who should conduct the evaluation
* Position of the evaluation unit within the organization
* Duration
* Approximate estimates of staff and budget

2. DEFINITION OF ORGANIZATIONAL STRUCTURE

* Position of evaluation unit within the organization
* Coordination
* Role of Steering Committee

3. OBTAINING STAFF AND BUDGET

* Preparation of job descriptions
* Discussion and negotiation with personnel management
* Submission and negotiation of budget

4. RESEARCH DESIGN

* Definition of objectives and key issues.
* Methodology
* Geographic coverage (For impact studies)
  * Model of project impact
  * Sample size

5 INTEGRATING THE EVALUATION INTO THE PROJECT DEVELOPMENT CYCLE

* Defining studies required at each stage of cycle.

6 DEFINING MAIN USERS AND THEIR INFORMATION NEEDS

7 PLANNING AND REVIEW CYCLES

* Defining the duration of the planning cycle.
* Defining review cycles
* Defining publications
* Building-in general review of the evaluation.

(g) To assist in the planning of future projects.

(h) To assist management in obtaining additional information on problems or particular aspects of project performance when the need arises.

1.3 Defining the level of complexity and the coverage of the evaluation

Organizations vary considerably in terms of the size and complexity of their programs and information needs, as well as in terms of the professional and financial resources available for the evaluation. The program manager must reconcile information needs with available resources. If the monitoring and evaluation is too complex there is a danger of overloading the capacity of the organization to conduct and absorb the studies. When this happens one or more of the following problems are likely to arise:

* Excessive delays before the results are published, thus reducing their operational utility.

* The quality of the studies suffers.

* The organization is not able to absorb and use all of the information so that much is wasted.

* Feedback between researchers and management becomes less frequent, again reducing the practical utility of the studies.

* The evaluation becomes excessively expensive in terms of money and demands on the time of key staff.

* The cumulative effect is to create a negative image of the evaluation and hence reduce staff cooperation in conducting and reviewing studies.

When the studies begin, they should be kept as simple and economical as possible. In this way the evaluation team can gain experience with the implementation of a set of basic studies and at the same time avoid the danger of overloading the capacity of the organization to review and use the results. The volume and complexity of the studies can be increased at a later point if there is a need.

In this respect it is sometimes better not to include complex longitudinal impact studies in the evaluation of the first project. For a first project it is usually more important to develop an effective system of performance and process monitoring as the main concern of management is to

know how well the basic operating systems are functioning. Impact studies become more important for a second project. By this time the basic operational model has been tested and planners wish to know the potential development impacts of the projects in areas such as employment and income generation. The manager should try to define what are the minimum information requirements for the evaluation and to ensure that no unnecessary studies are included.

1.4 Who should conduct the evaluation

Monitoring should normally be conducted "in-house" although it is possible to use consultants for special tasks. On the other hand, impact evaluation studies will often be subcontracted. The manager should decide:

(a) Can the complete monitoring and evaluation program be conducted in-house?
(b) If this is not possible, which parts should be subcontracted and who should do them?
(c) What use, if any, should be made of consultants or expatriate advisers? (See Chapter 5 Section B)

1.5 Position of the evaluation in the organizational structure

The main options for the organization of the Monitoring and Evaluation at the level of the proejct executing agency are given in Chapter 5 Section B. Some of the questions to be answered are:

(a) To whom should the evaluation report?

(b) Should a special evaluation unit be established?

1.6 Duration of the monitoring and evaluation program

The studies normally continue throughout the physical implementation of the project, but will often end when the loan disbursements are completed and the infrastructure is installed. Monitoring and evaluation is often financed under the project loan and consequently comes to an end when the project is administratively completed. Many of the project impacts, however, can only be measured when participants have had access to the new houses or services for several years. The issue of duration of the evaluation is therefore important. The following are some of the questions to answer:

(a) Are the studies expected to continue throughout the implementation of the project?

(b) Is there interest in measuring project impact? If so, can the impact evaluation continue beyond the termination date of the project?

(c) Will the evaluation unit become permanent and continue to work on other projects, or will its life end with the present project?

1.7 Estimates of staff and budgets

Chapter 5 Section F indicates typical staffing levels for different types of monitoring and evaluation exercises. In estimating budgets it is necessary to distinguish between: (i) money available for hiring consultants; (ii) discretionary funds available within the department (travel, etc.); (iii) authorization for use of certain resources (computer, vehicles, travel allowances, etc.); and (iv) budget line items for hiring or assigning staff to the evaluation.

The following questions must be answered:

(a) How many staff positions are to be created at each professional level?
(b) What temporary staff authorizations are required (interviewers, data coders and processors etc)?
(c) How much money will be requested for consulting services and how can these funds be used (can they pay for interviewing or be used to pay for the research services of other government departments, for example)?
(d) How much discretionary funds will be requested for use within the executing agency and for what purpose?
(e) What authorizations are requested for items such as computing and travel?
(f) Will funds be requested for scholarships, travel abroad and training?
(g) What budget line items are being requested for hiring permanent staff?

2. DEFINITION OF THE ORGANIZATIONAL STRUCTURE (See Chapter 5 Section B-2)

Three main issues have to be resolved by the project manager:

2.1 The position of the monitoring and evaluation unit, or units, in the organization.

2.2 Coordination between the monitoring and evaluation unit(s) and other divisions.

2.3 Role of the Steering Committee.

3. DEFINING AND MOBILIZING FINANCIAL AND HUMAN RESOURCES

The negotiating and administrative procedures required to obtain staff and budget should be started as early as possible as they can involve considerable time and effort on the part of program management. It may be necessary to lobby the Personnel Management and Planning authorities to ensure they understand the program and the justification for the numbers and qualifications of the required staff. The following are typical steps which must be taken:

3.1 Preparation of job descriptions.

3.2 Discussion and negotiation with Personnel Management.

This process and negotiation can be very time consuming, as it will often be found that the types of staff required do not fit easily into civil service categories. Consequently, it is necessary to discuss and negotiate special ways of defining jobs and hiring the right types of staff.

3.3 Submission and negotiation of budget.

4. RESEARCH DESIGN

Before the evaluation begins, it is essential to have a clearly defined research design. Although much of the design is a technical matter for the research team, it is essential to have guidance from management on certain key issues, of which the following are some of the most important.

4.1 Definition of research objectives and key issues.

4.2 Methodology - the contribution of management is to ensure resources are being used in the right way. In particular, it is important for management to ensure that scarce resources are not being used on unimportant questions or to achieve unnecessary levels of precision.

4.3 Geographic coverage - if a project covers many sites or several cities, it will be necessary to establish research priorities. Should all project sites be covered or should the research concentrate on just a few sites? This is an important issue, because if the coverage is too wide the analysis will be much more superficial. On the other hand, if some projects

are not covered, the use of the evaluation as a management tool will be greatly reduced. Often the best solution is to achieve a minimum coverage of all project sites and to select a few sites for in-depth studies.

For impact studies, two further issues may arise:

4.4 Definition of a model of the project implementation process, of expected impacts and of the assumptions underlying the expected outcomes. It is essential to have a clearly defined model of the processes of change which are being evaluated (see Annex A) as without this conceptual framework the evaluation will be unable to explain why the expected impacts have not been produced. The manager must ensure that the research design will be able to measure all important impacts expected from the project.

4.5 Determination of sample size - (Annex D).

Although the estimation of sample size must be the responsibility of sampling specialists, management must provide guidance on the levels of precision which are required. This is determined by the types of decisions which will be made on the results of the studies. If guidance is not received from management, the samples may either be too small and not answer the important questions, or too large and waste money.

5. INTEGRATING THE EVALUATION INTO THE PROJECT DEVELOPMENT CYCLE

Table C-2 presents a typical project development cycle. Five main stages are shown:

(a) Planning
(b) Design
(c) Implementation
(d) Cost recovery and maintenance
(e) Planning new projects

At each of these stages, management requires different types of monitoring and evaluation information. A successful monitoring and evaluation program should respond to the different information needs at each stage of the project cycle. Six main types of studies can be identified, each relating to one of these stages:

Table C-2: INFORMATION REQUIREMENTS AT EACH STAGE OF PROJECT DEVELOPMENT

PROJECT PLANNING		---------- I M P L E M E N T A T I O N ----------			PLANNING NEW PROJECT
PLANNING	DESIGN	INFRASTRUCTURE CONTRACTOR CONSTRUCTION OF SHELTER UNIT SELECTION OF PARTICIPANTS PROGRESSIVE DEVELOPMENT CONSTRUCTION COMMUNITY ORGANIZATION	REVIEW AND MODIFICATION OF DESIGN AND IMPLEMENTATION POLICY	COST RECOVERY MAINTE-NANCE	PLANNING NEW PROJECTS

Type of Study	Planning	Design	Control of Progress	Accountability of Donors	Quality Control	Project Effect-iveness	Planning Studies
	Demand Affordability Types of service Location Employment creation	Materials Cost Form of con-struction Service layout	Physical progress Bottlenecks Problems Disburse-ments	Comparison with objectives and implementation schedule Costs Financial control	Efficiency of imple-mentation Quality of houses and ser-vices	Impact on participants Impact on low-income families Impact on the city Impact on housing policy Accessibility to target popula-tion	Demand Afforda-bility Location Employment components

5.1 Planning Studies

At the point when a project is being planned, management requires information on factors such as affordability, characteristics of artesans, access to health services, factors determining project location, etc. Much of this information may not be available when the first project is being planned. However, the studies produced during the first project can contribute to the design of future projects.

5.2 Design studies

These are similar to planning studies, but tend to be more specific and related to matters such as project layout, choice of materials, amount of artesan credit etc.

5.3 Performance Monitoring

When project implementation begins, regular reports will be required on the progress of physical implementation, financial status and the causes and possible solutions of delays or problems in the implementation process. This information can be supplied through the monthly and quarterly project progress reports.

5.4 Accountability to lending agencies

Most lending agencies, whether they be national or international, require regular reporting on the progress of the projects and particularly on their financial status. Some of the required information can be obtained from the Quarterly Progress Reports.

5.5 Quality control

Management will require regular information on the efficiency with which the project is implemented and the quality of work. This can be provided through the quarterly progress reports and through special studies.

5.6 Project effectiveness and impact

Project management and government planning agencies will also require information on how well the project is achieving its more general objectives and its impacts on participants, the city and national housing policies. This

information can be provided through impact studies and through interim and final evaluation reports.

6. DEFINING THE MAIN USERS OF THE STUDIES AND THEIR INFORMATION REQUIREMENTS

6.1 The main types of information required from the evaluation

It is helpful to think of 4 main types of information which can be produced by evaluation studies (See Table C-3). The collection of each type of information has a cost and it is the responsibility of the manager to define which types of information are worth paying for.

(a) General indicators of project progress - Often these will be simple numerical indicators of the status of implementation of each of the main project components. These can include financial, physical and socio-economic aspects of the program.

(b) Indicators of project effectiveness - These are indicators of the extent to which the project is achieving its goals, both specific (such as the construction of a number of shelter units) and more general goals (such as impact on national housing policies). These studies vary in duration from a few weeks to several years.

(c) Indicators of project efficiency - These indicators are comprised of assessments of overall project operation, and of its individual components such as material supply stores, cooperatives and selection of participants.

(d) General planning information - The evaluation studies, particularly the more sophisticated surveys, generate statistical data which can be of value to a large number of planning agencies. Although these studies are not strictly evaluation, their preparation may become an important function of the evaluation unit.

6.2 Main users of evaluation outputs and their information needs

To design an operationally useful evaluation system, it is necessary to identify the main consumers of the studies and the information needs of each group. There are usually four main audiences:

(a) Project implementers such as the site manager, the director of a cooperative program, and the director of the municipal tax collection department. Implementers are mainly interested in short term feedback on project progress, and in the evaluation of the efficiency of the components for which they are responsible.

Table C-3: THE MAIN TYPES OF INFORMATION AND ISSUES WHICH EVALUATION RESEARCH CAN COVER

1. INDICATORS OF PROJECT PROGRESS

   selection progress
   construction
   occupancy
   house consolidation
   drop-outs
   maintenance
   cost recovery

2. INDICATORS OF PROJECT EFFECTIVENESS

   accessibility/affordability
   increasing housing stock & access to urban services
   impact on target population
   effect on urban housing market & development policy

3. PROJECT EFFICIENCY

   Efficiency of Individual Project Components

   project planning & design
   selection procedures
   construction methods
   material loans
   maintenance
   cost recovery
   community participation
   plot occupation

   General Project Efficiency

   efficiency in terms of project goals

      design
      finance
      implementation
      maintenance
      cost recovery

   comparison with alternative shelter programs

      cost comparison
      quality comparison
      replicability

4. GENERAL PLANNING INFORMATION

   income and employment
   expenditure and consumption
   housing quality and access to services
   health
   community organization

(b) Project managers are responsible for the general management and execution of a component or project such as the general manager of a specialized low-cost housing program, or the city director of the Ministry of Health. Managers are interested in a wide range of information, but due to time pressure and the wide areas of responsiblity, he or she will require brief summary reports.

(c) Central planning and finance ministries, are responsible for overseeing all development projects, particularly those receiving outside funding. Often their priority concern relates to financial information on disbursements, cost overruns and completion dates. However, in some countries there is a national agency with responsibility for a general overview of a wide range of projects. This agency or ministry will require information on project progress (particularly the financial aspects), and depending on its mandate, issues related to efficiency and effectiveness. The general planning information will also be of interest.

(d) Donor and lending agencies are particularly interested in indicators of project progress. They may also be interested (depending on the intensity of supervision) in the analysis of project efficiency and effectiveness. Agencies interested in the planning of future projects may also be interested in the general planning information.

7. DEFINING THE PLANNING AND REVIEW CYCLES

For monitoring and evaluation to be operationally useful, reports must be produced in time to assist management with planning and control decisions. This means that the planning cycles of the evaluation must correspond to the project's planning cycles. The following are some of the key decisions which must be taken to ensure this:

7.1 Every project has its monthly, quarterly or yearly planning cycles. Management normally meets at the end of each cycle to review progress and to plan for the next cycle. It is essential that the production of evaluation reports corresponds to this cycle so that the findings are available in time to assist management in reviewing progress and making future decisions.

7.2 Defining review cycles - Procedures must be developed to ensure that all evaluation reports are reviewed and that feedback is provided to the research team on the strengths and weaknesses of the reports.

7.3 Defining publications - Dissemination within and outside the organization is an important way to ensure the results of the evaluation are known and acted upon. A publication plan should be developed.

7.4 Building in a general review of the evaluation - It is recommended that approximately once a year there should be an independent outside review of the progress of the evaluation. This can either be done by a consultant or by technical assistance through one of the international donors or lending organizations. Arrangements for this review should be built into the evaluation program.

ANNEX D: THE BASIC CONCEPTS OF SAMPLE DESIGN

Many of the evaluation methods described in this handbook require the selection of one or more samples of individuals, households or communities from whom information will be obtained. This annex briefly describes some of the basic concepts involved in sample design. All of these issues are discussed in more detail in "A methodology for impact evaluation in urban development projects." (Bamberger, 1984.)

1. Why are samples used?

A sample is a number of units (people, households, communities, etc.) which have been selected in a systematic way so as to permit estimates to be made about the characteristics of the population from which these units were drawn. For many purposes a well chosen sample will be just as useful as a census in which all households (persons, etc.) have been interviewed. Samples are often quite small compared to the population they come from, and consequently the information can be obtained much more cheaply and quickly.

A well designed sample will ensure that sufficient interviews are conducted to guarantee a required degree of precision of the estimates of population characteristics. The use of these procedures for estimating sample size can avoid the use of larger than necessary samples whilst at the same time ensuring that sufficient interviews are conducted to provide the required degree of precision.

The following example illustrates the differences in purpose of a sample and a census. Assume that the possibility of starting a literacy program is being considered, and that it is decided that the program would only be justified if at least 25% of the adult population was illiterate. It is a relatively simple task to design a sample survey which will estimate, with an acceptable level of confidence, the proportion of the population which is illiterate. The sample could provide accurate information on the proportion of illiterates, but it could not be used to identify each particular household who would like to take the literacy program. For this latter purpose it would be necessary to conduct a census in which every family in the community was interviewed.

2. Sample precision and confidence intervals

Suppose we wish to estimate the average income of a community in which approximately 5,000 families live, and that a sample survey of 100 families provides the following information:

Mean monthly income = 125 PESOS ± 25

(0.05 level of confidence)

This signifies that the mean monthly income was 125 pesos for the sample of 100 households. What does this mean? First, it does not mean that all households have this income, but rather that this was the average. Some households may have had incomes as low as 25 or 30 and others may have had income as high as 500 (or even 5,000) pesos.

Second, it does not imply that this is the true mean for all households in the community, but only for those included by chance in the sample. If another sample of 100 families was randomly selected it is possible that the mean of this second sample might be 105 pesos or perhaps 143 pesos. So what good is it to know the mean for this particular group of families if the result could have been different for another group?

The answer is that it is possible to obtain from this sample an estimate of the probable range within which the true mean of the whole population lies. In the above example the figure of ± 25 pesos is the standard deviation of the mean. We know that there is a 95% probability that the true population mean lies within two standard deviations of the sample mean. In the present case the "95% confidence limits" for the estimate of the mean are 125 ± (2x25) so that the lower confidence limit is 75 and the upper limit is 175. This confidence range (also called the 0.05 range in some texts) means that there is a 95% chance that the true population mean lies within this range. It is important to understand that sampling theory is always based on probabilities and confidence limits and never on certainties. In calculating the sample size the researcher must decide what is an acceptable level of risk of being wrong. For most evaluation purposes it is conventional to use the 95% or even the 90% limit.

Once the precision of the required estimates is known (the confidence interval and the confidence level) it is possible to estimate the number of interviews which must be conducted. The main determinant of the precision of the estimates is the sample size. Normally it is necessary to make quite a substantial increase in the sample size in order to significantly reduce the

confidence interval. For example, with a sample size of 100 the confidence interval is ± 50 pesos. In order to reduce the interval to ± 25 pesos, it would be necessary to increase the sample size to 400. Similarly, a confidence interval of ± 10 pesos would require a sample size of 2,500. It is clearly important to define the required level of precision before designing the sample.

3. Methods of sample selection

There are 3 main ways to select a sample: Simple random sample. With this method each unit of the population (household, person, etc.) is given an equal chance of being selected. This is the simplest type of sample to design and is often quite adequate where the population to be studied is relatively small and concentrated. The selection and application of a simple random sample can be very expensive and complex where the population units are difficult to identify and are widely scattered. In many cities no list or map exists on which all households are located. Even if such a list did exist, the costs of interviewing would increase very considerably if the sample was scattered throughout a city of the size of Sao Paulo or Calcutta. The interviewer would have to spend a great amount of time travelling and his or her interviewing rate would drop considerably (not to mention the cost of transport).

Cluster sample. With this method, the population is divided into clusters, with interviews only being conducted in a relatively small number of clusters. If it is necessary to prepare maps, a limited number of relatively small sectors of the city will be selected and maps only prepared for these few areas. Even where maps do exist it is common to cluster the interviews so as to reduce travel time and costs. From the theoretical point of view the estimates obtained from a cluster sample are less precise than those obtained from a simple random sample, but this is offset by the very significant reductions in cost per interview. In general the greater the number of clusters which are used, the more precise will be the estimates.

Stratified random sample. A frequent sampling problem is that some of the groups of interest to the survey only represent a small proportion of the total population. For example, if recent migrants only represent 5% of the population, a simple random sample of 1,000 households would probably only include about 50 of this group. If it was decided (in terms of precision

estimates) that a minimum of 100 recent migrants and 100 established residents should be included, this would mean that an additional 1,000 randomly selected interviews would have to be conducted in order to locate the additional 50 recent migrants. This is obviously an extremely expensive and inefficient way to proceed.

This problem can usually be resolved through the use of a stratified sample. The population is divided into strata - in the present example recent migrants and established residents. The required number of interviews for each stratum are then selected from among all households in that stratum. By stratification it would be possible to achieve the required precision (100 households from each group) with a sample of 200; instead of the 2,000 which would be required with a simple random sample.

A stratified sample can usually provide the most precise estimates. Unfortunately there may often be considerable costs involved in the construction of the strata. In the present example, how do we locate recent migrants? The decision whether or not to use a stratified sample will often depend on a comparison of the costs of constructing the strata and the expected benefits to be derived from the use of stratification.

4. Related (panel) and independent samples

Although most textbooks assume that the evaluation will be able to use a panel design in which the same subjects are interviewed before and after the project, in practice this is frequently not possible. In many low-income communities there is a high population turnover rate so that it would be quite common to find that a quarter or more of the households had moved in the 2 or 3 years between the first and second interviews. This means that in practice the evaluation researcher must choose between: panel or related samples in which the same subjects are reinterviewed; independent samples in which a new sample is selected in T(2) and a mixed sample which combines some of the elements of the previous two approaches. The following are some of the sampling issues involved in the use of each of these designs:

Panel sample design: In this design the same households or subjects will be reinterviewed in the second survey. In order to use this design it is important to prepare maps of the precise location of houses, or to use other similar techniques to ensure that the original households/subjects can be relocated. In many cities the process of identifying the same households two

years later can be very difficult. New houses are built, street names change and even the numbers and directions of streets can alter. Another problem is that household composition and the name of the head can also change. For example, in the first survey a women may declare that she is the household head. By the time of the second interview she may have a male companion who is now declared to be the hosuehold head. This can be very confusing to the interviewer who is trying to establish whether it is still the same family.

Another factor to be taken into account is the estimated drop-out rate and its impact on sample size. Assume for example that it is estimated that a sample of 200 households is required for valid estimates to be made. Assume also that it is expected that 25% of households may move before the time of the second survey. This means that the sample size in T(1) must be increased to 250 households, so that the final sample size in T(2) after 25% of households have moved, would be around 200. If a high drop-out rate is anticipated it is also necessary to consider the effects which this will have on the representativity of the final sample. It is advisable at the end of the second survey to conduct an analysis of the T(1) data to compare the characteristics of households who have moved and who have remained in the community to determine the ways in which they differ. If, for example, the families who have left are richer than those who remained, this must be taken into account in interpreting the findings of the analysis which is only conducted on households who remained in the community.

Independent samples: With this design a new random sample is selected for the second survey in T(2). No particular sampling problems exist.

Mixed sample design: This is the most complicated design to administer. The same procedures are used as in the panel study except that replacements will be found for orginal families or subjects who cannot be reinterviewed. The simplest option is to replace the household with the new family living in the same structure. This has several potential biases. Firstly it means that families living in new structures built since the time of the previous survey will be excluded. This produces a bias against new households. Secondly, it is difficult to define the population to which new occupants belong. Although most of them are probably new to the community, it is possible that some may have moved from other structures in the community.

A better approach is to select a new sample for the replacements. The sample should be selected from all households or subjects who have moved to

the community since the time of the previous survey. This can be somewhat cumbersome to select as a large number of screening interviews may be required in order to identify new families. In practice it is relatively simple to identify new households in sites and services shelter projects as management will usually have records of new arrivals, but the process of selection can be very cumbersome for the control areas.

The use of a mixed sample also raises issues related to the estimation of sample size. In order to make full use of the analytical potential of this sample, it is useful to be able to conduct separate analysis of original households (panel) and new arrivals. In order to do this it is necessary to make separate estimates of the required sample size for both groups.

5. Sample Designs for the Evaluation of Sites and Services Projects

It is easier to design samples for the evaluation of a sites and services project than it is for an upgrading project. In the case of sites and services, the affected population is clearly defined and relatively small. At the same time all participants receive the same package of services or one of a limited number of options. The sample of participants can either be a simple random sample, or if there are different options, a stratified sample which includes a sample of each option can be used.

Ideally the control group should be selected from all low-income households in the city. In practice, this tends to be too expensive and normally some type of cluster sampling is used with the selection of a relatively small number of areas which are considered to have similar characteristics to the participants. As participants usually represent a relatively small proportion of the total population, it is usually possible to find control groups with relatively similar characteristics

6. Sample Designs for the Evaluation of Upgrading Projects

The sample design for the evaluation of an upgrading project tends to be more complicated. First, it is more difficult to define who has been affected by the project. In a large upgrading project a number of different government agencies (water, education, transport, health, technical training, housing credit, etc.) all provide services. Not all households will receive the same package of services, and many households may not directly benefit at all. This makes it difficult to define the limits of the target population.

Second, the wide spread of the programs means that the expected impact at the level of individual households may be quite low. Consequently a randomly selected group of households may reveal a relatively low average impact, even though the project has affected large numbers of households. In some cases the impact per household will be so low as to be statistically insignificant. Third, the fact that many upgrading projects are intended to cover most of the low-income population means that it will often be impossible to find a control group of similar low-income households who have not been affected.

For all of these reasons it will often be necessary to use a different type of evaluation design which includes the following characteristics:

* Instead of dividing the population into an experimental group which has been affected, and a control group which has not, each family will be classified in terms of its degree of access to project benefits. Indicators will be used of degree of access to water, schools, health programs, credit, etc.

* Often it will not be possible to have an external control group. In some cases the control group will be formed of families in the project areas who have not yet received project benefits. In other cases the control group will be developed statistically with households who have not received a particular service being used as a control group for families who have received this particular service. The statistical procedures for the analysis of this type of control are given in the discussion of multiple regression in Annex F.

Table D-1: FEATURES OF SITES AND SERVICES AND UPGRADING PROJECTS WHICH AFFECT THE DESIGN OF AN IMPACT EVALUATION

	Sites/Services	Upgrading	
Selection of participants	All participants must satisfy criteria in terms of income, family size etc. This will often eliminate up to 25% of the poorest families in the city.	All families in upgraded areas are automatically included.	In neither cases are participants randomly selected so there are problems of locating an equivalent control group. In the case of sites and services the problem is to control for variables such as motivation. With upgrading the problem is to find similar socio-economic areas when the project may include all squatter areas.
Speed of Project	May be a gap of 3 years or more between time when first and last families occupy their project houses.	In large areas the project progresses in stages, with several years between first and last stages receiving the services.	In both cases the implementation delays complicate the design design. Easier to control in upgrading as there is a phased sequence whereas in sites and services the occupation is unpredictable. In upgrading the later areas can be used as control group for earlier areas.
Proportion of low income families	Normally less than 10%	Can cover up to 75% of the urban poor	Much more difficult to select a comparable control group for upgrading
Standardization of the package of services	Families receive same package, or one of a small number of clearly defined options	Wide differences in services actually received Difficult to measure what family received	Difficult to define and standardize the "treatment" for upgrading.
Uniqueness of project areas or participants	Many similar families	Projects often cover almost all poor, or only affect a very special group not easily matched	Difficult to find equivalent control group for upgrading.
Definition of project area and partici - pants	Area consists of new housing and is clearly defined	Limits of project effective area may be ambiguous	More difficult to define experimental group for upgrading.

FURTHER READING

Dennis Casley and Dennis Lury "Monitoring and Evaluation of Agricultural and Rural Development Projects."(1982) Explanation of how samples are designed for the evaluation of agricultural and rural projects.

Lesley Kish "Survey Sampling" (1965) One of the standard textbooks on sample design.

Michael Bamberger "A Methodology for Impact Evaluation in Urban Shelter Programs" (1984). More detailed discussion of the issues presented in this annex.

ANNEX E: EXPERIMENTAL AND QUASI-EXPERIMENTAL DESIGNS IN URBAN IMPACT EVALUATION

Managers and development planners frequently require a quantitative evaluation of the impact of their projects. This is necessary both for comparison with alternative investment strategies, and for deciding whether the project is producing a satisfactory return on the investment. To make this type of quantitative evaluation, it is necessary not only to determine whether changes have taken place in the project population, but also to determine whether the changes are due to the project or to other unrelated factors. Cities are in a constant state of change so that project participants are subject to many other factors in addition to the project. The separation of the project impact from that of other factors requires a research design which can control for the effect of these external factors.

The urban researcher is never able to achieve the closely controlled experiments conducted in the natural sciences or in animal psychology. In order to provide the best possible estimates under the difficult circumstances in which social science research is conducted, a number of "quasi-experimental" research designs have been developed which try to approximate the true experimental design.

1. The logic of the true experimental design

In order to understand the logic and some of the limitations of these "quasi-experimental" designs, it will be useful to begin with a brief discussion of the true experimental design. Let us assume that an experiment is being designed to evaluate the impact of a drug on the speed with which rats can find their way through a maze. The time taken by the rat to find its way to the end of the maze is called the "dependent variable", and the amount and type of drug is called the "experimental" or"independent variable". The action of the drug may be affected by characteristics of the rats such as age and weight (intervening variables). The effect of these intervening variables is controlled by randomly assigning some rats to the experimental group and assigning others to a control group which will not receive the drug. If the groups are reasonably large, it can be assumed that these intervening variables will be randomly distributed between the control and experimental groups and will, consequently, not influence the outcome of the experiment.

The experiment in its simplest form is designed as follows:

	T(1)		T(2)
Experimental group	E(1) _____	X _____	E(2)
Control group	C(1) _____		C(2)

T(1) and T(2) represent the time periods before and after the experiment. E(1) and E(2) indicate the average time taken by rats in the experimental group to complete the maze in T(1) and T(2). C(1) and C(2) are the corresponding values for rats in the control group. X indicates the point at which the drug ("treatment") was applied.

The time of two groups to complete the maze in T(1) is measured. The two groups are then kept in identical conditions except that the drug is administered to one group and not to the other. The two groups will then again run the maze and their average times will be measured. The average times of the two groups in T(2) are then compared by using the following, or a similar equation:

$$I \text{ (Impact)} = \frac{E(2) - E(1)}{C(2) - C(1)}$$

If the drug had no effect at all, the value of I would be close to 0 (due to chance variations it is almost never exactly 0). If, on the other hand, the drug affected the time taken to complete the maze, I would assume a value which is either negative (if the drug slowed completion time) or positive (if the drug speeds completion time). A statistical test - such as the T-Test - is then used to determine the likelihood that the observed difference has occurred by chance. If there is less than 1 chance in 20 (probability is less than 0.05) of a score as large as I occurring by chance, it is conventional to say that the result is statistically significant. A stricter test would be to use the .01 level which indicates there is only one chance in 100 of the result having occurred by chance.

If the difference is statistically significant then it will be concluded that the drug appears to have had an impact on the time taken to run the maze. A careful researcher will repeat the study a number of times with different groups of rats, and under slightly different experimental conditions, so as to eliminate any chance factors which might have influenced

the outcome. For our present discussion it is important to emphasize the following aspects of this design:

(a) A relatively large number of rats are used.

(b) The rats are assigned randomly to the two groups.

(c) The rats are kept in identical conditions during the experiment, with the only difference being that one group receives the drug and the other does not.

(d) Ideally the experiment is repeated a number of times with different groups of rats.

(e) Only one carefully controlled treatment (in this case the drug) is used, and all experimental subjects receive exactly the same dose and at the same time.

(f) The experiment will often be repeated with carefully controlled changes in the dosage and in the time period over which change is measured.

2. Threats to validity in the interpretation of the analysis

The immediate purpose of conducting an impact evaluation is to estimate the impacts which a project or set of project interventions have had on a particular population. However, this analysis is usually only the first step towards estimating the potential impact which the project could have on some larger group. For example, the project being studied may have provided piped water for several hundred families in one low-income community. Assuming the project appears to have produced some beneficial health impacts on this community, the question of interest to policy makers is to estimate the likely impacts if the same type of project was extended to cover all low-income families in the city. It is dangerous to assume that the observed results in the pilot project would necessarily occur in the same way if the project was replicated on a larger scale. There are four main sets of factors which may affect the validity of the estimates about how the project could be expected to perform on a larger scale. These factors, which have been called threats to validity are summarized below. A more detailed explanation of how they apply in urban projects is given in Bamberger (1984).

(a) Statistical conclusion validity

Problems of statistical conclusion validity arise when the sample is too small, or where it has not been properly designed. Under these circumstances it may appear that the project has had no impact, when in fact

the sample was too small for it to have been possible to obtain statistically significant differences between the experimental and control groups. Another common example is when project impacts only occur in a certain sub-group (such as houseowners or owners of large fishing boats) which has been undersampled so that the impact is not detected. A related problem is caused when there is a high population turnover so that the original households in the sample cannot be re-interviewed. This can have the effect of reducing the size of a panel sample (in which the same households are reinterviewed at several points in time) by 30 or 40 percent. In this case the originally adequate sample size may have been reduced so much that it becomes much more difficult to obtain statistically significant results.

Another statistical problem can occur if the wrong analytical model is used. Annex F shows the problems which can arise when a test such as the T-Test is used to estimate differences between an experimental and control group when the two groups were not equivalent at the start of the study. An example is given where the T-Test shows a statistically significant difference between the two groups, but when the two groups are matched through the use of regression analysis, it is found that there is no longer a difference between them.

(b) Internal validity

Once it is established that a statistical association exists between the project and the dependent variable being evaluated, the next task is to determine whether this association is evidence of a causal relationship. It is important not to confuse a statistical relationship with causality.

The fact that the income of participants in a housing project increased faster that the income of a control group does not prove that the project caused the increase. It might be that families who enter the project are those most likely to increase their income (such as small businessmen looking for new markets, or generally more ambitious people), or it might be that increased income is due to a third factor, such as the opening of a new factory near the project. A number of factors which might be responsible for a spurious causal relationship are the following (the list is taken from Cook and Campbell (1979).

(i) History: an event occurred in one area and not the other after the project had begun.

(ii) Maturation: certain changes occur spontaneously after a certain period of time. For example, housing investment is related to time living in a community.

(iii) Testing: the effect of repeating the interview may cause certain changes to come about.

(iv) Instrumentation: a slightly different questionnaire or method of measurement might have been used at different points in time, and this might have produced different results.

c. Construct validity of causal relations

Different researchers may interpret the same findings in different ways. The example was given earlier of different interpretations which might be put on the fact that more houses exhibit middle class symbols such as expensive iron grilles on the windows, ornate plants or patios. One person may interpret this as meaning there has been "middle class encroachment" and poor families have been forced to leave. Another person may argue that this shows that poor families have remained in the community but have begun to adopt middle class status symbols. The two researchers would draw quite different conclusions from the results of the study.

d. External validity

This refers to the problems of generalizing to a wider population. It often happens that a first project attracts the most dynamic participants. In this case it is possible that the results of the first project will be greater than those of subsequent projects due to the differences in the level of ability, motivation or resources of participants. Under these circumstances it would be dangerous to make inferences from the first project about likely outcomes of future projects.

3. Problems in applying the experimental design to the evaluation of urban projects

We will now consider what happens when we attempt to apply this experimental approach in a typical urban setting.

PROBLEM NO. 1 - Participants in urban projects are almost never selected randomly. In the case of a small business credit program participants are selected from among those businessmen who took the initiative to apply and, consequently, who have different motivational patterns from those who did not apply. Frequently the project will also try to select those persons or groups

most likely to be successful in the project (for example, people willing to participate in self-help house construction or transport cooperatives whose vehicles are best maintained) so that the selected group becomes even less representative of all low-income families. In the case of upgrading, all families in a given area are automatically included, and consequently it is extremely difficult to select a control group of households with similar characteristics.

PROBLEM NO. 2 - The high population turnover rates in many urban areas mean that it is often not possible to use the panel design in which the same subjects are interviewed before and after the project. The implications of this are discussed in Chapter 4 and in Annexes D and F.

PROBLEM NO. 3 - Many projects provide a number of different inputs (infrastructure, credit, technical assistance, land tenure, community services, etc.). In addition to the large number of different inputs (treatments), a further complication is that different families will receive different combinations and amounts of these inputs. This again makes the design much more complex than the neat laboratory situation in which all rats receive exactly the same amount of the same drug.

PROBLEM NO. 4 - In urban projects it is difficult both to select and maintain a control group. An objective of many projects is to cover as great a proportion of the low-income population as possible or to include all families living in certain areas of the city. In both cases it will be extremely difficult to find a control group of families with similar characteristics. A further problem is that many of the selected control areas are potentially unstable and likely to be erradicated by urban renewal, destroyed by flooding or upgraded. Even when the area itself remains, it is common to find that between 25% and 50% of the families interviewed in the baseline study will have moved during the two to three year period before the survey is repeated. The instability of the control areas complicates considerably the analysis.

PROBLEM NO. 5 - Projects are often subject to long and unpredictable delays. It is quite common for a project to start eighteen months later than scheduled and in one project studied there was a delay of almost five years. Such delays make it extremely difficult to plan the evaluation and may mean that some treatments still cannot be evaluated even at the end of a three or five year study.

PROBLEM NO. 6 - Another serious problem is the fact that the changes produced by most projects tend to be continuous rather than to occur at one point in time, as in the case of the rats. The pattern of change is also not linear but likely to fluctuate.

Tables E-1 and E-2 give two typical examples. Table E-1 shows possible levels of housing investment in a sites and services project. When families first occupy the project the level of investment many be relatively low during the first 6 to 9 months as many households have not yet occupied their houses or are still paying off the down-payment on the house. Investment may then begin to increase as families complete the house according to their needs and resources. These high levels of investment may continue for several years, but at some point investment will begin to decline once most families have completed the basic structure. If monthly housing investment before the project was compared with the level 6-9 months after houses were assigned, it might be found that very little change had taken place. However, if the study were conducted one year later the level of investment would probably have increased very considerably. If, however, the study was repeated 3 years later, it might be found that investment per month was again not very much higher than the pre-project level. This emphasizes the point that levels of housing investment should be measured at various points in time as a comparison between only two points is likely to be very misleading.

Table E-2 illustrates the typical patterns of fluctuations in household income. In many areas there are strong seasonal fluctuations as income may be higher during the harvest, the tourist season or the months in which most construction takes place. In other cases, as for example a fishing community, the fluctuations may take place over shorter periods. These fluctuations mean that household income in any given year will tend to be significantly lower in (say) June than in (say) December. On the basis of this comparison the conclusion might be incorrectly drawn that the project had had a negative effect on income. Even if income is compared during the same month in two different years, there are considerable fluctuations from one year to another. Thus although there is a gradual increase in income during the years following the project, the income in years 2 and 5 is lower at the peak than in the year in which the project began. Again a comparison between only two points in time could be very misleading. The conclusion from both of these

TABLE E-1: TYPICAL FLUCTUATIONS IN HOUSING INVESTMENT DURING THE FIRST YEARS OF A SITES AND SERVICES PROJECT

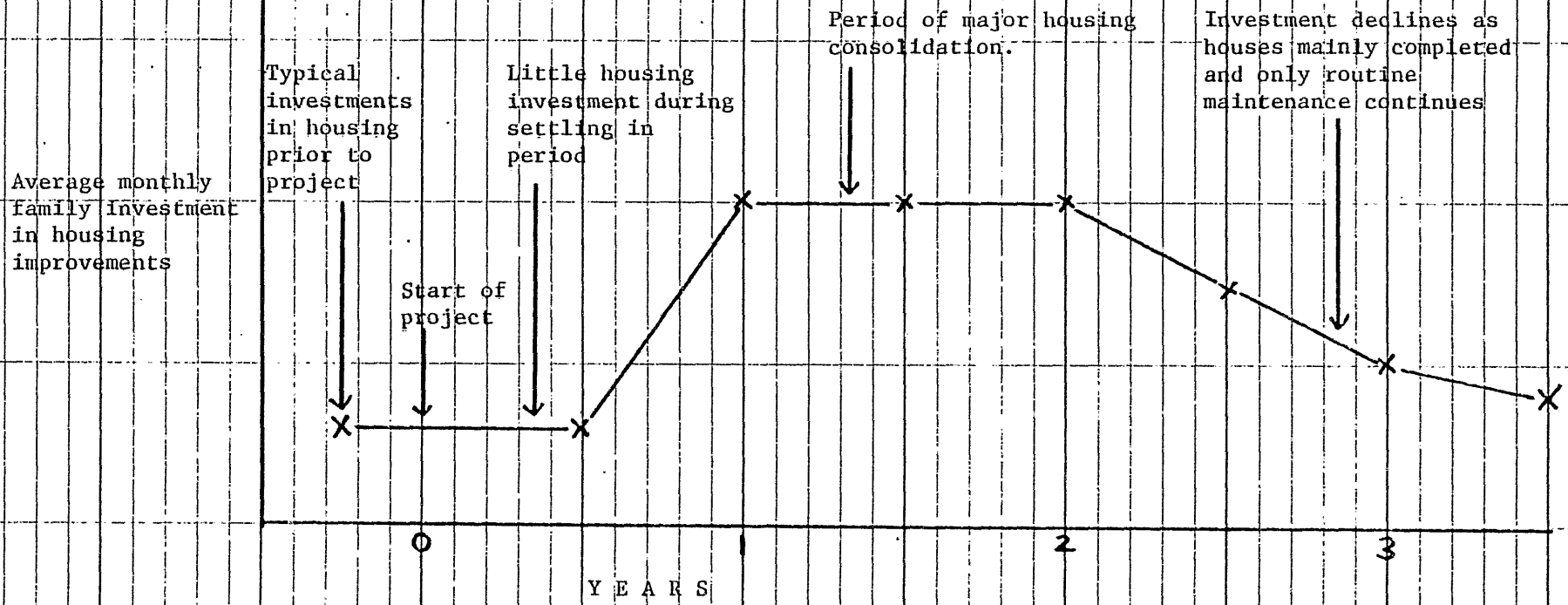
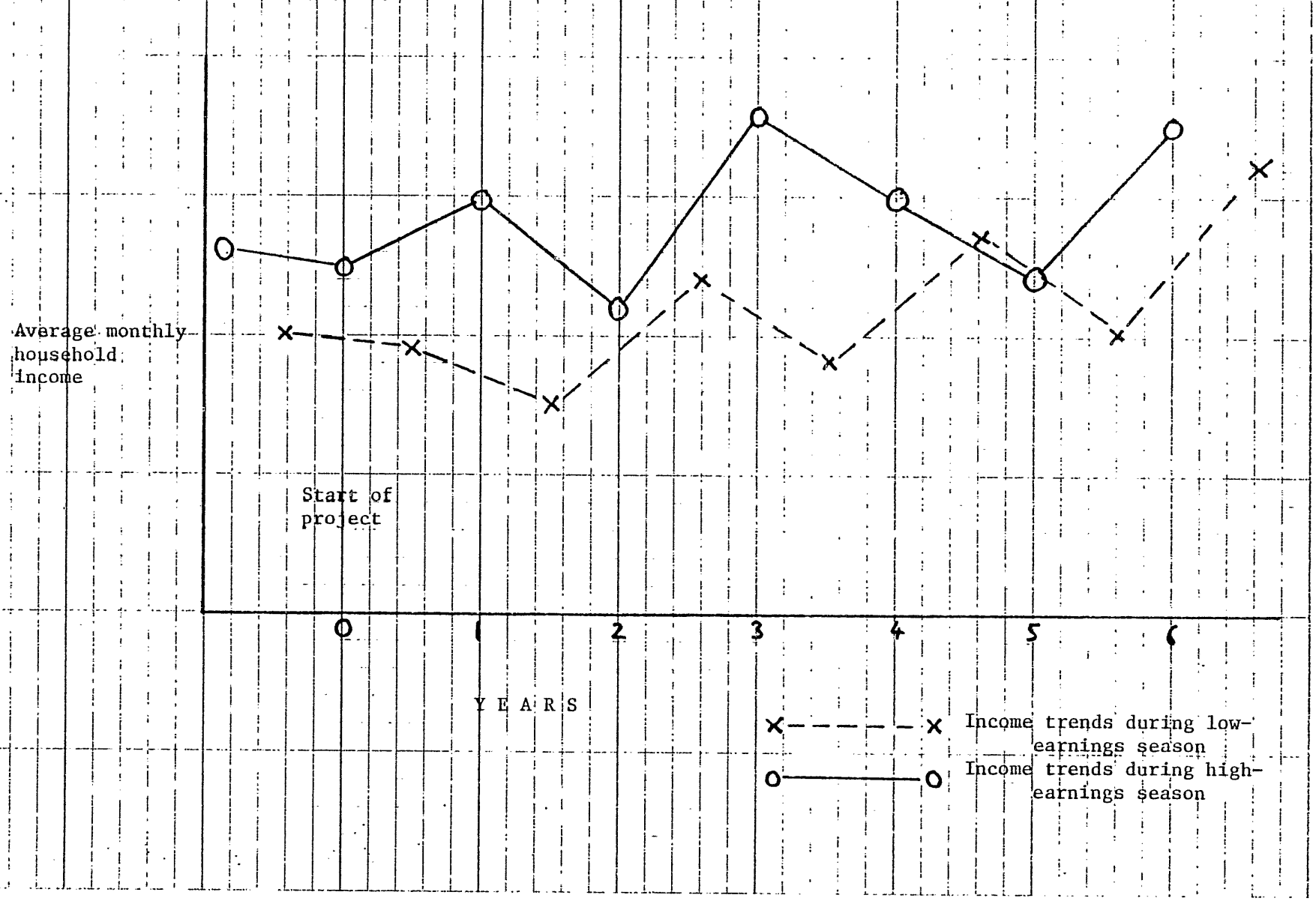


TABLE E-2: EXAMPLE OF LONG-TERM TRENDS AND SHORT TERM AND SEASONAL FLUCTUATIONS IN HOUSEHOLD INCOME IN THE YEARS FOLLOWING THE START OF A SHELTER PROJECT



examples is that it is necessary to study trends over time and that reliance should never be placed on a static comparison between two points.

An important related problem is that when a project is expected to produce several different impacts, each impact may occur over a different time period. Thus it may be possible to detect some economic changes after only a year, whereas educational or health benefits may not be measurable for 5 or even 10 years. This complicates even further the selection of the time period over which measurements will be made.

PROBLEM NO. 7 - Perhaps the most serious shortcoming of the experimental design from the policy and operational point of view is that the simple design does not provide any insights into why expected impacts did not occur. Let us return to the model presented in Table C-1. The experimental design would simply measure the amount produced by the business or household income before and after the provision of credit and technical assistance. If no statistically significant change was observed, the analysis would not be able to explain why there was no change nor would it be able to tell us much about the potential utility of the program.

It is recommended that a statistical impact evaluation should always be complemented with at least a descriptive analysis of the operation of each stage of the project model.

4. Alternative quasi-experimental designs

The characteristics of a project, and the context within which it is developed, determine to a large extent the type of evaluation design which can be used. Some of the most common designs are presented below, together with a brief discussion of some of their advantages and drawbacks.

(a) The Pre-Test Post-Test Control Group Design

In this design the experimental and control groups are compared before and after the project. The design can be represented as follows:

	T(1)	T(2)
Experimental	E(1)--- X ---	E(2)
Control	C(1)----------	C(2)

In the simplest form of the analysis it is assumed that all participants receive the same package of services. With this assumption it is possible to conduct the analysis using multiple regression with a dummy variable for

Participant Status (Participant=1; Control= 0). The questionnaire should include information on the socio-economic characteristics of the subjects so that the two groups can be matched statistically to control for the effect of these characteristics.

In those cases where the project offers a number of different services or components, the questionnaire should measure exposure to these services. In this case the multiple regression equations will include variables reflecting exposure to project components. The coefficients of these variables indicate their contribution to the impacts being studied. It is also possible to include interaction terms to measure the covariation between sets of components.

(b) Extending the pre-test post-test design to include additional points in time. This model is similar to the previous one except that one or more additional observations are included. The design can be represented as follows:

	T(1)	T(2)	T(3)	T(n)
Experimental	E(1)--- X ---	E(2) ----	E(3) ----	E(n)
Control	C(1) --------	C(2) ----	C(3) ----	C(n)

The points at which the observations will be made are determined by the nature of the project. In some cases observations may be made every year or two years. Another option is to define T(2) as the point when the project is administratively completed. T(3) would then be several years later.

This model can test whether the trends are linear or whether they fluctuate or even reverse. The more observation points which are available the less the danger of making wrong inferences about trends based on only two or three observations.

The analysis of this design can be conducted in various ways. One way is to conduct the analysis described in the previous section for T(1) and T(2); T(1) and T(3) etc and then to compare the results. Another is to select one time period for the basic analysis and to use information from the other periods as a consistency check.

(c) Combination of pre-test post-test with a continuous panel study

In this design the pre-test post-test design is complemented by the selection of a small panel of subjects who are continously observed throughout the study. The simplest form of the design is represented as follows:

	T(1)						T(2)
Experimental	E(1)	p1	p2	p3	p4	pn	E(2)
Control	C(1)	p1	p2	p3	p4	pn	C(2)

A small panel of subjects are selected and contacted at each of the periods indicated as "p". The panel study may consist of structured or unstructured interviews or even of direct observation. The logic of this design is to understand the processes involved and to identify the factors which are producing the observed changes. The number of subjects used in the panel is usually much smaller than in the survey so that this can be a more economical way of measuring trends.

(d) Pre-test post-test without control group

	T(1)		T(2)
Experimental	E(1)	--- X ---	E(2)

This design is used when it is not possible (for technical or financial reasons) to include a control group. The participants are studied before and after the project and the changes are calculated. A common situation in which this design is used is the evaluation of a project which is intended to cover most of the target population, and where it is not possible to identify an equivalent control group.

The design is very weak as it is not possible to control for other factors which may have produced the observed changes. Annex F shows how multiple regression can be used to strengthen the design, particularly in multi-component projects where households receive different combinations of services.

(e) Ex-post comparison of experimental and control groups

	T(1)		T(2)
Experimental		X	E(2)
Control			C(2)

In this design the project and a control group are compared after the project has been implemented. Again this is a very weak design as no direct information is available on the conditions of the two groups before the project began. In some designs a number of retrospective questions are included asking subjects to recall what their conditions(income, business sales etc) were before the project began. Recall information of this kind is frequently subject to a considerable margin of error and must be used with caution.

f. Interrupted time series

This design is used when a time series is available(For example, monthly sales of a cooperative or average income from rent). Ideally observations should be available for at least 10 periods before the project began, and for a further 10 periods after the project has started. In the simplest type of analysis the mean scores are plotted for each month and visual inspection is used to determine if there is any observable difference after the project has begun.

A more sophisticated approach is to estimate a regression for the period before the project and again for the period after the project. The slopes are then compared to determine whether there has been any change.

5. Application of quasi-experimental designs in upgrading and sites and services projects

Sites and services projects have a number of features making them easier to evaluate than upgrading project (see Annex D Table D-1). First, it is easy to define which families are affected by the project. The houses are constructed on previously empty land and there is no ambiguity as to whether a family is a participant. Second, families tend either to receive the same package of basic services, or to be divided into distinct groups, each of which receives a clearly distinct package of services. Third, sites and

services usually affect a relatively small proportion of a city's population, making it possible to select a control group of families who have characteristics fairly similar to those of the participants.

Upgrading projects are more difficult to evaluate. First, there are wide variations in the package of services received by different families. At one extreme, a family may gain direct access to a paved road, be located close to a public water supply, or be able to pay for the installation of a private water connection. At the other extreme, there may be families living in sectors where roads have not been paved or drainage installed, and who cannot afford a private water connection. Second, in projects involving a number of government agencies, a large diversity of services may be provided over widely scattered areas. One family may live near a day-care center, but too far from a job-training program to be able to enroll. Another family may have access to a health clinic, but not to a community center. Such circumstances make it very difficult to use a simple random sample, as the number of families who receive each service will be very low.

Third, as mentioned earlier, the scale of many upgrading projects and the high proportion of low-income families covered makes it extremely difficult to identify a control group which has similar characteristics, and who will remain unaffected by the project.

Each of the problems discussed above presents difficulties in the design and interpretation of the evaluation. When the evaluation does not satisfy the conditions of the experimental design it is difficult to state with any degree of confidence that observed differences between the experimental and control groups have been "caused" by the project.

A framework for understanding the precise nature of these analytical problems has been developed by Cook and Campbell (1979) in their discussion of the main "Threats to Validity" in different quasi-experimental designs. The way in which these threats to validity affect the design of urban evaluations has been discussed by Bamberger (1981). Annex F presents some of the statistical techniques used to minimize the effects of these threats to validity.

FURTHER READING

Thomas Cook and Donald Campbell "Quasi-experimentation. Design and Analysis Issues for Field Settings" (1979). Detailed explanation of a wide range of evaluation designs and analytical procedures.

Michael Bamberger "A Methodology for Impact Evaluation in Urban Shelter Projects" (1984).

ANNEX F: METHODS OF DATA ANALYSIS

1. Statistical Analysis of the True Experimental Design

Although it is rarely possible to achieve a true experimental design in urban research, it is useful to mention briefly the types of analysis which are used with this design in order to illustrate the statistical adjustments which must be used for the more common quasi-experimental designs. With a true experimental design, in which subjects are randomly assigned to the experimental and control groups, the impact of the experimental treatment can be defined as:

$$\text{Impact (I)} = \frac{E(2) - E(1)}{C(2) - C(1)}$$

The statistical procedure which will be used for evaluating there is a significant difference between E(2) - E(1) and C(2) - C(1) will depend on the type of variable and the "level of measurement".

(a) When the variable is binary (member or non-member of the cooperative) change will be measured as the difference between proportions. The appropriate statistical test will be the T-Test for the difference of proportions.

(b) When the variable is nominal (variables such as type of house, occupation, etc) where one category cannot be defined as being greater than or less than another) the appropriate test will be one of the non-parametric tests such as Chi-Square.

(c) When the variable is ordinal (so that categories can be ranked from greater to lesser but where the intervals between the categories are not equal) the appropriate test will be one of the rank correlation tests such as the Wilcoxon Matched Pairs Signed Ranks Test.

(d) When the variable is interval (variables such as income, age, value of the house, etc. where the intervals between each number are equal) the most appropriate test will probably be the T-Test.

A further complication in the selection of the appropriate test is to determine whether the samples are related or independent. Many of the statistical tests assume that the samples being compared are independent of each other, and when a panel design is used in which the same subjects are reinterviewed it is usually necessary to use different statistical tests. (See Bamberger, 1982, for details.)

The purpose of the significance tests is to determine whether there is a statistically significant difference in the amount or rate of change which has been observed between the experimental and the control groups. For example, it was found in El Salvador that monthly household expenditure on food in the control group increased from 155.2 pesos in 1977 to 241.5 pesos in 1979 (an increase of 86.3 pesos). During the same period, food expenditures by project participants increased from 177.4 pesos to 240.6 pesos (an increase of 63.2). The T-Test showed that a difference as large as this had only a 0.0005 possibility (5 chances in 10,000) of occuring by chance. Consequently, it could be concluded that the increase in food expenditures by control households was statistically greater than for project households (see Bamberger, 1982, pages 179 onwards for more details).

2. Tests of Association and Difference for use with some of the simpler evaluation designs.

The simplest type of statistical analysis is to compare two or more groups and to determine whether or not there is evidence that the group affected by the project is different from the control group. Examples include:

(a) Comparison of project participants before the project has begun and after it is completed or has been underway for a certain time, to determine whether there is a difference. A test such as T-Test can determine whether there is a significant statistical difference between the two time periods. The form of the test will depend upon whether related or independent samples are used (Bamberger, 1982 Chapter 6). This design is very weak as without a control group it is not possible to determine whether the changes are due to the project or to other factors.

(b) Comparison of the scores of participants and a control group to determine whether there is any difference after the project. Again tests such as T-Test will determine whether there is a statistical difference, but as there is no measurement before the project began, it is not possible to assess how much the two groups differed before the project began and consequently how much of the change might be due to the project.

(c) Comparison of the project group with a control group before and after the project. It is again possible to use a test such as T-Test to compare the difference in the difference of means (i.e has the mean of the project group changed more or less than the control group). This method is more powerful but

still faces the problem that where there was no random assignment to the control and experimental groups it is likely that differences existed between the two groups at the start of the project and that these initial differences explain part of the apparent project effect.

3. The Use of Multiple Regression in the Analysis of Quasi-Experimental Designs

One of the major statistical problems in evaluating the impact of urban projects is that the types of control groups which are available tend to differ in some important ways from the project participant (experimental) group. In some cases the average household income will be different; in other cases there will be differences in the number of years the small businesses have been operating or in the number of vehicles owned by each bus company.

Table F-1 shows the differences which were found between the experimental and control groups at the start of the evaluation of the sites and services project in El Salvador. The household heads in control areas were older and had significantly less education; control families were also smaller and had lower incomes. When the T-Test was used to compare incomes before and after the project, it was found that there had been a statistically greater increase in the income of participant than of control families. This would seem to indicate that the project had affected income. However, the question arises as to whether the more rapid increase in participant incomes might not be due to some of the initial differences between the two groups with respect to income, family size, age of head, etc., rather than to the project.

This question can be partially answered through the use of Multiple Regression Analysis. Table F-2 shows the type of information which is provided on the example we have been discussing. The column headed B (regression coefficient) indicates how much change in income is produced by a unit change in the indicated variable when all other variables remain constant. The final column indicates whether the coefficient is statisticlly significant (if the probability is equal to or less than 0.05 it is assumed that the difference is significant). For example, the coefficient for "Education of head" is 17.8. This means that for every additional year of education of the head, household income after the project, defined as T(2) would increase by 17.8 pesos. Similarly, every additional household member would increase income by 21.9 pesos. The table indicates that even without the presence of the project the existence of differences in initial income,

Table F-1: The problem of non-equivalent control groups, Initial differences between the control and experimental groups in Sonsonate, El Salvador at the start of the evaluation (1977)

Variable	Participants	Control	Difference	Test	Score	Probability
Years of education	4.52	3.05	+1.47	T	3.96	0.0001*
Weeks worked last month	3.9	3.78	+.12	T	1.39	0.168
Months in present job	112	126	-14	T	-1.01	0.339
Family size	5.67	4.82	+.85	T	2.89	0.004*
Household income last month	3.85	301	+84	T	3.52	0.001*
Age of head	37.2	43.5	-6.3	T	-3.7	0.0001*
Sex of head (% male)	60.3	66.7	-6.4	T	.725	0.39
Branch of economic activity				X2	4.04	0.67

Note: For all calculations N = 238

* = statistical difference between two groups at 0.05 level or beyond

Table F-2 The use of multiple regression analysis to control for initial differences between the experimental and control groups. The example of income changes in Sonsonate, El Salvador.

Dependent variable = Income in T(2)

Independent variables	BETA	B	F	Probability
Income in T(1)	0.49	0.83	67.9	0.01
Family size	0.14	21.9	5.9	0.05
Education of head	0.17	17.8	6.9	0.01
Age of head	0.13	3.58	4.1	0.05
Participant status	0.01	68.2	3.31	0.10

Note: Other non-significant variables not incldued in the table. These include: weeks worked last month, months in job, sex of head, type of residence.

family size, education and sex of household head between the project and control groups would tend to produce differences in household income between the two groups in T(2). The final row shows that when we control for the initial differences in these 4 variables there is no longer a significant difference in post project income between participants and control group (defined as participant status). Thus, the apparent impact of the project was in fact due to the initial differences between project and control households.

The form of the analysis will depend upon whether a panel sample design is used, with the same respondents being interviewed twice, or whether separate samples were selected in T(1) and T(2). In the former case, the analytical model can be specified as follows:

$$Y_2 = a + b_1Y_1+b_2C_1+b_3C_2+\ldots.b_nPART$$

where: Y_2 = value of the dependent variable(income) in T(2)

Y_1 = value of the dependent variable(income) in T(1)

C_1, C_2, C_n = household characteristics (age, education etc)

PART = Dummy Variable with Participant=1 and Control=0.

The purpose of the analysis is to determine project impact on a dependent variable (say income). The dependent variable is income in T(2), and this is regressed on income in T(1), on a set of socio-economic characteristics of the household, and on a dummy variable indicating whether the family was a participant or member of the control group. If the coefficient of PART is statistically significant this indicates that there is a difference in the value of income between the two groups after controlling for household characteristics and income in T(1). In other words, project participation is associated with changes in income.

The analysis is more complicated in those cases where separate samples are selected in T(1) and T(2). In this case it is necessary to conduct 2 separate cross-sectional regression analyses. The first is conducted on the pre-test survey and the second is conducted on the post-test survey. The form of the pre-test analysis is as follows:

$$Y_1 = a + b_1C_1+b_2C_2+b_3C_3+\ldots b_n PART$$

where: Y_2 = value of the dependent variable(income) in T(1)

C_1, C_2, C_n = household characteristics (age, education etc)

PART = Dummy Variable with Participant=1 and Control=0.

The post-test analysis will have the same form except that data from the second survey will be used (and the dependent variable will of course be Y_2). In its simplest form the analysis of project impact consists in comparing the coefficients for PART before and after the project. If the project has had an effect on income, then there should be a statistical difference between the PART coefficients.

The same approach could be used in the analysis of small businesses, public transport companies or health programs.

4. <u>The use of multiple regression when not all participants receive the same package of services or when there is no control group</u>

Two characteristics of many projects make it virtually impossible to use the conventional analytical approaches to the evaluation of project impact. First, many projects such as upgrading, transport and small business development are intended to affect almost all of the low income populations so it is impossible to find a comparable control group of families who will not be affected. Second, projects frequently comprise a number of different components which beneficiaries may receive in different combinations. One small business may only receive credit, whilst another receives both credit and technical assistance. There are also differences in the order in which services are received. Consequently, it is not possible to divide the population into an experimental group affected by the project and a control group not affected.

The most effective approach to both of these problems is the use of the multiple regression techniques discussed in the previous section. Instead of referring to an experimental and a control group, the analysis is conducted in terms of the degree of exposure to each project component. Table F-3

Table F-3 Exposure of three families to the services provided by a hypothetical upgrading project

Variable Name	Indicators of exposure to project services			
	Service	Family A	Family B	Family C
WATER	Water	Indoor tap	Indoor tap	No water
LOAN	Building material loan	No loan	500 pesos	1000 pesos
MARKET	Location	4 blocks from market	20 blocks from market	25 blocks from market
TENURE	Land tenure	Received tenure	Received tenure	Did not receive tenure
TRANS	Transport	3 blocks from public transport	20 blocks from public transport	10 blocks from public transport
LOT	Lot size	95 square meters	50 square meters	70 square meters

illustrates the way in which degree of exposure of 3 households can be measured with respect to:

Water:	Indoor tap (Score=1) v No indoor tap (Score=0)
Building materials loan	Score = value of the loan
Access to market:	Score = number of blocks from the market
Tenure:	Received tenure (Score = 1)/ Not received tenure (Score = 0)
Access to transport:	Score = blocks from bus stop

With indicators such as these it is possible to measure in a questionnaire the degree of access of each household to all of the services provided by the project. Similar measures could be used for access to small business services or health programs.

Table F-4 shows how these measures of exposure are treated in the multiple regression analysis. In this example a number of important simplifying assumptions have been made. The most important is the assumption that there are no interactions between the project inputs (which permits the use of a simpler additive model). This assumes, for example, that the impact of water on housing investment is not affected by whether the family has received a material loan, has land tenure or the distance they live from a paved road. If these types of interactions do exist, the specification of the equation would become more complex as it would be necessary to include interaction terms to describe the way in which the level of one variable affects the impact of another. A second simplifying assumption is to express all variables in their simplest linear form. In more complex forms some of the variables might be expressed as powers or roots. Although the researcher must be aware that these simplifying assumptions have been made, they are justified for the first phase of the analysis. In most studies a simple linear additive form will be used for the first runs, and then more complex forms may be tested. In this example the objective is to evaluate the impact of each project component on the stimulation of housing investment. As in the previous example, the column "B" indicates the amount of change in housing investment produced by a unit increase in each of the project variables given in the first column. Similarly the final column "probability" indicates

Table F-4 Simplified example of multiple regression analysis. The impact of water, building material loans, technical training, tenure and paved roads on increases in housing investment

Variable		B	F	Probability
X_1	Water	0.2	4.1	.05
X_2	Material loan	2	7.2	.01
X_3	Technical training	3	2.8	Greater than .05
X_4	Land tenure	200	5.9	.025
X_5	Paved roads	0.1	4.9	.05
Constant				

Multiple correlation coefficient (R^2) = .59

$F(4,120) = 3.8$ Probability is less than .01

Standard Error of R^2 = 3.8

whether the coefficient is statistically significant. It can be seen, for example, that when all other variables are held constant, a 1 peso increase in the materials loan will produce a 2 peso increase in housing investment. Similarly, acquisition of land tenure produces an average increase of 200 pesos in housing investment.

How can we interpret these statistics? Table F-5 indicates the maximum change in housing investment which can be produced by each component. For example, let us assume it is known from project records (or from the survey itself) that the maximum daily water consumption of a household is 250 liters. The coefficient for water is 0.2 which means that the maximum impact for a family consuming 250 liters will be:

250 x 0.2 = 50 pesos

Similarly, the maximum possible building materials loan is 1,000 pesos. Combining this with the coefficient of 2, we know that the maximum possible impact is:

1,000 x 2 = 2,000 pesos

The table shows that the greatest potential contribution to housing investment is the building materials loan (potential impact = 2,000), followed by the acquisition of land tenure (potential impact = 200). Using the simplifying assumption that the regression equation is additive, the maximum possible project impact can be estimated by adding the maximum values for each individual component. Thus if a family used 250 liters of water per day, received a 1000 peso building materials loan, took a technical training course, received land tenure and lived close to the road, the equation predicts that a total of 2238 pesos would be invested in housing improvements. This is only a very approximate estimate, but if used with caution, it is a helpful way to evaluate the relative contribution of different project inputs.

The important conclusion for the policy maker is that water, technical training, and access to roads have a negligible impact on housing investment. The most effective single intervention is a building materials loan program, as this contributes over 80% of the maximum potential project impact.

The same analysis can be repeated to evaluate project impact on households income, housing value, employment, etc. The only requirement is that it must be possible to define and measure precisely the dependent variable.

Table F-5 Estimating maximum potential project impact and the contribution of each project component to this impact. The example of increased housing value.

Project component	Maximum value	Coefficient(b)	Increase in Y for maximum value
Water	250 liters per capita per day	0.2	50
Building material loan	1000 pesos	2.0	2000
Technical training	1 (binary)	3	3
Land tenure	1 (binary)	200	200
Nearness to road	500	0.1	50
Constant	25		
MAXIMUM IMPACT			2328 pesos

Note: An additive linear model is assumed.

Although this is a very powerful analytical tool for policy makers, care must be taken in the interpretation as the findings are based on the peculiar conditions in the communities studied. Water is a good example. One might have expected that access to water would provide a strong incentive to invest in house improvements. However, the figures show water as having a very small impact. The reason for this may be that many of the richer families, with the greatest potential for housing investment, have already paid to have water installed. Consequently, during the period of the survey, most of the households who installed water were much poorer and were able to mobilize fewer resources for investment. Examples such as this suggest that great care should be taken when interpreting findings from only a small number of communities.

How does this technique resolve the problem of an absence of a control group? The logic of the approach is to consider all households who have not received a particular service as the control group for that particular component. For example, if 240 households have not received land tenure and 260 have received it, the 240 would be used as the control group for analyzing the impact of this component. What the model permits us to say is: "When we compare households with the same level of access to all other project components, having access to land tenure will, on average, produce an increase of X pesos in housing investment". It is also possible to introduce socio-economic characteristics of the households into the analysis so that the impact of land tenure can be compared for households of different income levels, with different numbers of members, etc. Using these techniques it is possible to statistically produce a control group with similar characteristics to the experimental group. From the statistical point of view this approach is weaker than the use of an external control group, but if used carefully it permits us to conduct an impact evaluation for many types of projects which conventional analytical procedures could not handle

The above discussion is based on the simple case in which a panel design is used. In those cases where the analysis is based on independent samples, the same procedures will be used as were discussed in the previous section.

5. Path Analysis

Although regression analysis can provide estimates of total project impact, it has the limitation of not helping the policymaker to understand the

dynamics of the project implementation process. A useful analytical technique for this purpose is Path Analysis. Table F-6 illustrates the use of Path Analysis to evaluate the impacts of a small business development program in Colombia (DESAP). The DESAP program offered loans, technical assistance and training courses. Five indicators of project impact were developed (increase in number of workers, production, sales, profits and fixed capital). These indicators were combined into a summary index of project impact (PROGRESS). In addition, increases in knowledge (KNOWLEDGE) was defined as an intermediate project impact.

The Table indicates the paths which connect characteristics of the company, project inputs and project impacts. The numbers on each path are the path coefficients which are the standardized regression coefficients. Thus the path coefficient between loans and PROGRESS is 0.14, and the coefficient between loans and KNOWLEDGE is 0.11.

It is possible to derive the residual path coefficients which indicate the proportion of the variance not explained by the variables in the model. This is defined as:

$$r_i = 1-R_i^2$$

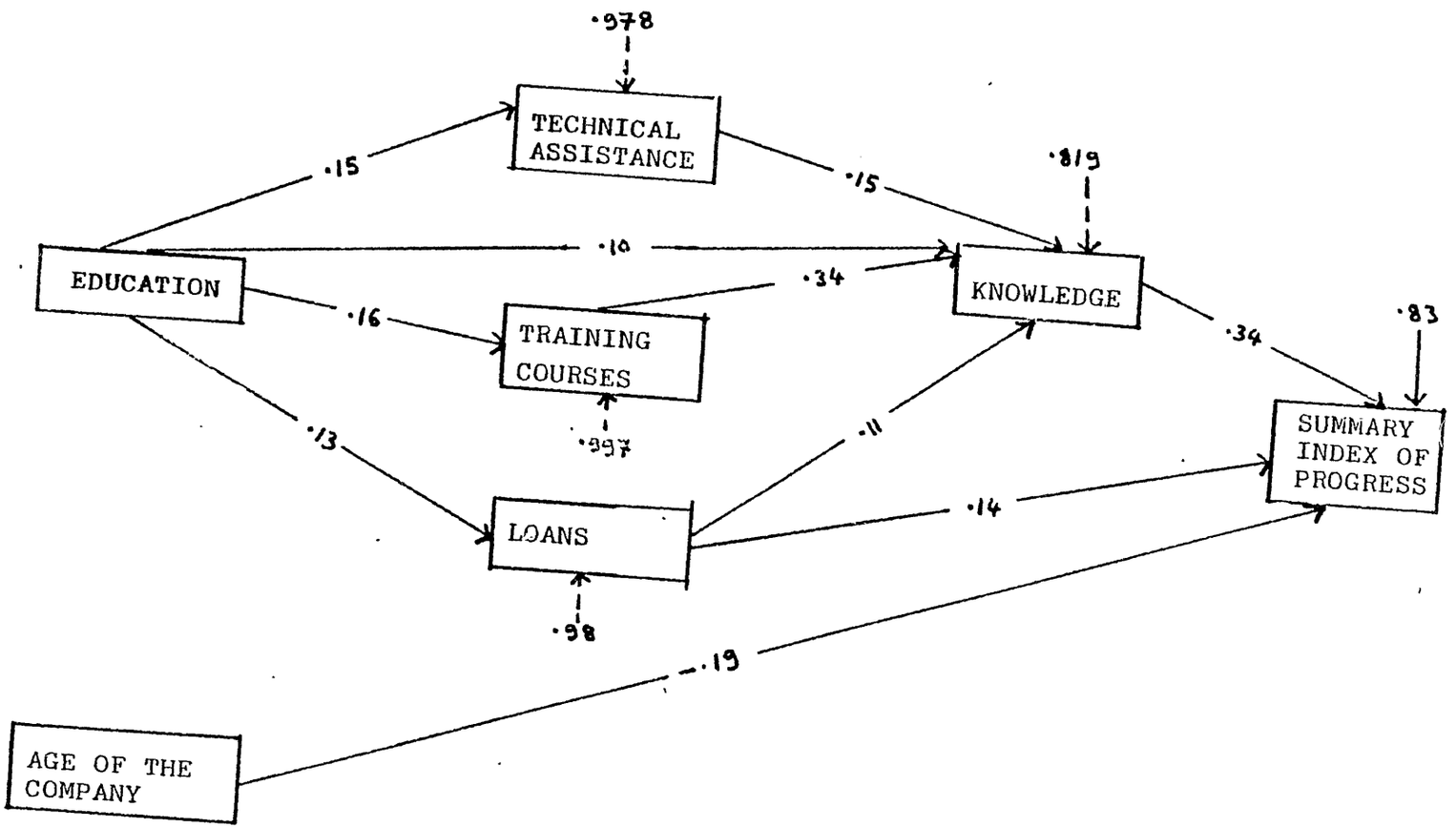
where: r_i = the residual path coefficient for variable i

R_i^2 = the multiple correlation coefficient for variable i.

The diagram shows that the unexplained proportion of the variance of PROGRESS is 0.83 (83%) which means that between them KNOWLEDGE, loans and age of the company explain 17% of the variance (100% - 83%). This means that although the factors we have studied in the model can explain a significant part of the variance in PROGRESS there may be other important factors which are not included in the model.

One of the advantages of the path diagram is that it helps the project manager to understand the relative influence of the various factors being studied and the ways in which they interact. A one unit increase in KNOWLEDGE will produce a 0.34 unit increase in PROGRESS; whereas a one unit increase in loans will produce a 0.14 unit increase in PROGRESS.

TABLE F-6: Example of the use of Path Analysis to evaluate the impacts of a small business credit program in Colombia (DESAP)



The diagram also helps management to understand the direct and indirect influences of each factor and hence to evaluate the most effective ways to increase project impact. For example, it can be seen that loans have both a direct effect on PROGRESS and also an indirect effect through their influence on KNOWLEDGE.

It is very simple to construct a path analysis diagram as all of the required information can be obtained from a set of standard regression equations.

6. Cost Benefit Analysis

Although an internal evaluation of a project can give us some idea of how efficiently it was organized and what types of impact it produced on beneficiaries, it cannot tell us how this project compared in terms of costs, efficiency and accessibility with other options available in the housing market. One way to make this type of comparison is through cost benefit analysis.

A cost benefit analysis of housing projects will usually comprise the following steps:

(a) A set of housing options are selected. Ideally these should include all types of formal and informal housing to which the low-income population could have access. In the case of the El Salvador study (Fernandez Palacios and Bamberger, 1984) the following housing options were included in the analysis:

(i) Unregulated squatter settlements
(ii) Public squatter upgrading programs
(iii) Extra-legal subdivisions
(iv) Tenement housing
(v) and (vi) 2 types of sites and services
(vii) Public multi-family units
(viii) Public single family units
(ix) Private single family units

(b) For each option, information is obtained on all costs related to the project. These costs typically include:

(i) Land
(ii) Site preparation
(iii) Installation of services
(iv) Contractors labor costs
(v) Contractors materials costs
(vi) Administrative costs
(vii) Additional labor costs for owner
(viii) Additional material costs for owner

(c) The costs can be evaluated from the point of view of the nation (public costs) or of the individual household (private costs). The calculation of public costs requires the estimation of the opportunity cost of the project in terms of the alternative uses of the resources. For example, the public cost of labor must take into account the unemployment level. If unemployment levels are very high, then the opportunity cost of the labor is much lower (i.e., there is little alternative demand). On the other hand, the private cost of labor is the foregone income. If a storekeeper could have earned 250 pesos in the time he spent working on his house, the cost to him is 250 pesos.

In the evaluation of public costs, it is necessary to take into account a series of explicit or disguised subsidies. Many of the administrative costs of public housing agencies are not charged to the project. Similarly, land value may be provided free by a local authority and its cost not charged to the project. One of the most time consuming parts of the cost benefit analysis is the identification of all these costs and the estimation of their social cost.

(d) Identification and measurement of all project benefits. Whereas objective measures of costs can usually be obtained, benefits must be imputed. The usual approach is to assume that the benefits obtained from a house are reflected in the market rent for this type of unit. Where a well functioning housing market is in existence, it may be possible to observe rents. However, in many low-cost housing projects rents must be imputed by asking households how much they think the house would rent for. This procedure is likely to produce a greater margin of error than occurs with the estimation of costs. There is also a question of whether rents reflect benefits such as improved health of which families may not yet be aware.

(e) The flow of costs and of benefits is computed by estimating the year in which each cost and benefit will occur. For example, if a family must pay a fixed purchase cost over a period of 15 years, this cost will occur during each of these years. On the other hand, building materials are usually paid for over a relatively short period of time.

(f) Costs or benefits which occur in the future have a lower "present value" than similar costs or benefits which occur at the present time. If a family has to pay 100 pesos in 2 years time, this has a lower <u>present value</u> than 100 pesos which must be paid today. The reason is that if a person owes

100 pesos but does not have to make the payment for 2 years, he could invest the money and earn interest during this period. If he had to make the payment today he would lose this interest. Thus, future costs or benefits are discounted by an appropriate rate of interest.

For each year, costs are subtracted from benefits to obtain the net benefits for that year. During the early years the net benefits will usually be negative as costs are higher than benefits, but during the later years the benefit streams will be positive. The net benefits are discounted by the appropriate rate of interest to obtain the present value of the net benefits for each year. The present values for each year are added to obtain the Net Present Value (NPV) of the project. If the Net Present Value is positive, this means that investment in the project yields a higher rate of return than could have been obtained from investing the money at the current rate of interest. If NPV is negative, the project yields a lower rate of return than could have been obtained from an alternative investment.

(g) An alternative analytical approach is to estimate the "Internal Rate of Return" (IRR). This is achieved by discounting the net benefit flows by different interest rates until a Net Present Value of 0 is obtained. For example, if an interest rate of 10.5% produces NPV=0, this means that investment in the project produces a rate of return of 10.5%. IRR produces results consistent with NPV in that:

(i) When IRR is less than the discount interest rate, NPV will be negative.
(ii) When IRR is equal to the discount interest rate, NPV will be 0.
(iii) When IRR is greater than the discount rate, NPV will be positive.

In order to compare different projects, NPV must be divided by the NPV of the cost stream to take into account the fact that different quantities of resources have been invested in each project. For example, the NPV of the IVU rehabilitation project is significantly lower than that of the meson (1078 compared to 1674), but the NPV/Cost of the former is nearly 20 times as high as the latter (0.2640 compared to 0.0141). It is advisable to conduct both IRR and NPV analyses as each can tell us slightly different things about the project.

Table F-7 presents the results of a cost benefit analysis in which El Salvador housing options were compared. In this case a discount rate of 12% (the current market interest rate for housing loans) was used. The IRR varies from a maximum of 33% for the FSDVM Basic Sites and Services Unit, to a

Table F-7: Comparison of housing options in terms of economic rate of return, net present value and net present value/total cost. San Salvador. 1978

Housing Option	Rate of Return	Net Present Value (Colones)	NPV/Cost	Ranking on 3 Indicators (1 = highest)
Upgrading and Sites and Services				
FSDVM Basic Unit	33	4065	1.2016	1
FSDVM Serviced Lot	28	2329	0.7269	2
IVU Rehabilitation	18	1078	0.2640	4
Traditional Housing				
IVU Multi-Family Unit	9	-1828	-0.1304	9
IVU Single-Family 2 Bedroom Unit	11	- 606	-0.0720	8
FSV Single-Family Unit	13	452	0.0641	5
Informal Market				
Colonia Ilegal	22	1788	0.3509	3
Meson	12	1674	0.0141	7
Tugurio	20	373	0.2972	6

Source: Marisa Fernandez-Palacios and Michael Bamberger, "An Economic Analysis of Low-Cost Housing Options in El Salvador." DEIRB (draft). WUDD Discussion Paper No. 55 1984

minimum of 9% for the IVU multi-family unit. NPV is presented in the second column. For projects with an IRR above 12% (the discount rate), NPV is positive; where IRR is less than 12%, NPV is negative. In the third column, NPV is divided by total cost so as to control for the substantial differences in the magnitude of the cost of each project. It can be seen that the orders of magnitude of this ratio have the same rank orders as the other two indicators.

An interesting extension of the above analysis is to compare the costs and benefits to the nation and to the household. For projects which include significant subsidies, the relative rankings in terms of their public and private cost-benefit ratios can be quite different. This is important information for the policymaker as it suggests that price distortions (subsidies) are encouraging households to invest in shelter options which are not the most efficient from the point of view of the nation.

7. Hedonic price analysis

Housing is not a single product of which one purchases a certain number of units, but a package which includes varying numbers and sizes of rooms, level of services, quality of construction, location in the city, and neighborhood characteristics. The purpose of hedonic price analysis is to estimate the amount which households are willing to pay for each component of the housing package. This is achieved through the use of multiple regression analysis where housing attributes are regressed on rent. The coefficients of each component can be interpreted as the additional amount households are prepared to pay to obtain an extra unit of this component. Table F-8 presents the findings of an hedonic analysis conducted on low-income housing in El Salvador (Quigley, 1980). The coefficient for number of rooms is given as 10.77. This can be interpreted as indicating that when all other components of the housing package are the same, a family would, on average, be prepared to pay an additional 10.77 pesos of rent per month for each extra room. Similarly they would pay 0.66 pesos for each extra meter of space and 2.76 pesos for access to piped water. Hedonic prices can be used at the planning stage to estimate potential demand and willingness to pay for different housing packages. The technique can also be used in an ex-post evaluation This latter application is extremely useful in those cases in which there is not a well functioning housing market and where, consequently, it is not possible to obtain directly

Table F-8: Example of different functional forms of a Hedonic Regression based on housing data from El Salvador

253 Observations--1976 Data

Coefficients	Linear Form 1	2	Semi Log 3	4	Log-Log 5
Number of Rooms	10.770	10.750	0.373	0.573	0.574
	(8.38)	(8.41)	(7.72)	(7.71)	(7.61)
Living Area (meters^2 x 10)	0.661	0.653	0.028	0.028	0.027
	(2.01)	(2.00)	(2.22)	(2.19)	(2.16)
Lot Size (meters^2 x 10)	0.528	0.477	0.034	0.033	0.412
	(0.70)	(0.64)	(1.20)	(1.15)	(1.59)
Electricity (1 = available)	2.904	2.823	0.159	0.153	0.154
	(1.34)	(1.34)	(1.95)	(1.91)	(1.87)
Piped Water (1 = available)	2.759	2.669	0.184	0.178	0.173
	(1.14)	(1.12)	(2.02)	(1.97)	(1.84)
Sanitary Quality	3.465	3.501	0.093	0.095	0.178
	(3.52)	(3.59)	(2.51)	(2.58)	(2.50)
Floor Conditiion	1.365		0.077		0.067
	(0.52)		(0.78)		(0.47)
Wall Condition	0.753		0.075		0.142
	(0.39)		(1.04)		(1.04)
Roof Condition	3.875		0.205		0.258
	(0.75)		(1.05)		(0.715)
Aggregate Condition		1.233		0.088	
		(0.89)		(1.68)	
Intercept	-17.270	-12.550	1.224	1.419	2.070
	(1.43)	(1.51)	(2.68)	(4.52)	(7.44)
R^2	0.363	0.363	0.358	0.357	0.345
R^2 (In Original Space)	0.363	0.363	0.384	0.384	0.353
SEE/mean	0.438	0.436	0.124	0.124	0.125

Source: John Quigley "The distributional consequence of stylized Housing Programs. Theory and Empirical Analysis." Urban and Regional Economics Division. The World Bank 1980. The table has been simplified.

information on the amount of rent paid by the shelter package offered by the project being evaluated. In the previously mentioned study in El Salvador, the coefficients given in Table F-8 were obtained from a study of the types of tenements in which participants lived before moving to the project. As is often done, various different specifications of the equation were compared to determine which form explained the highest proportion of the variance. The R^2 was found to be highest for the semi-logarithmic form. The coefficients were then applied to the services provided by the project (lot size, water supply, roof conditions, etc.) to impute the amount which households would have been prepared to pay for a shelter unit with this level of services. It was found that this hedonically imputed rent was higher than the monthly charges participants were required to pay. The difference between the imputed rent and the actual payments was interpreted as a consumer surplus obtained by participants.

FURTHER READING

Marisa Fernandez-Palacios and Michael Bamberger "An Economic Analysis of Low Cost Housing Options in El Salvador. (1984). Example of the use of cost-benefit analysis to compare housing options.

Michael Bamberger "Statistical procedures for the evaluation of project impact (1982). Explanation of the methods of statistical analysis which can be used with each evaluation design.

Stephen Malpezzi "Analyzing an urban housing survey: economic models and statistical techniques." (1984). Explanation of how to analyze the results of a housing survey. Includes a simple explanation of the logic of regression analysis.

Emmanuel Jimenez "The value of squatter dwellings in developing countries" (1982). Example of the application of hedonic price analysis to the estimation of the benefits obtained from each component of a squatter upgrading project.

Hubert Blalock "Social Statistics"(1972). One of the best textbooks on statistical analysis.

S. Siegel "Non-parametric statistics for the behavioural sciences" (1956). Easy to follow explanation of the statistical procedures to use with ordinal and nominal variables.

N.H.Nie and others "Statistical Package for the Social Sciences"(1975). Presentation of the most widely used statistical computer package which contains all of the procedures referred to in this chapter. Also an easy to follow textbook on statistics.

ANNEX G SAMPLE OUTLINE OF A QUARTERLY PROGRESS REPORT

1. General

The example which is presented in this Annex refers to an urban shelter project. However, most of the principals discussed can be applied equally well to other types of project, although the specific content of the report will obviously vary.

The purpose of the progress report is to collate key information from the monitoring system for management review. To this end, the information contained in the reports should cover at least the following aspects of the project.

(a) comparison of actual progress with the original estimated schedule of implementation;

(b) actual or contemplated major deviations from original plans or schedules and reasons for these deviations. Changes which would require prior consultation with management should be reported immediately and included in a subsequent report;

(c) conditions which could significantly affect construction and procurement schedules and/or the cost of the project;

(d) the latest cost estimates; commitments and expenditures and the availability of funds to meet the cost of the project;

(e) progress made in community development activities.

2. Contents of the report

(a) Summary

A brief summary of the problem features of implementation should be included.

(b) Physical Implementation

This section should describe issues, events and changes concerning construction of civil works on each subproject, site by site. Where actual or expected problems become apparent, information should be given on measures taken or planned to correct them and the probable effects on scheduling and costs. The following tables should be used to summarize the progress:

Table 1: Implementation Schedule (comparing actual versus scheduled targets)

Table 2: Illustrates how the implementation schedule could be used to describe the progress of a project to renovate the existing housing stock. The first bar indicates the planned starting and completion dates of each component or activity. The second bar indicates progress to date as a percentage of total work. The third bar indicates the new estimated starting and completion dates.

Table 3: Construction Progress Sheet

Table 4: Contract Data Sheet. This is used to summarize and record events up to the point of award of contract.

Table 5: Summary Data on Contractors' Payments. This is used to monitor payments to contractors.

3. Project Financial Status

Table 6 - Project Contract Sheet should show substantial changes in the cost estimates that have become necessary as the project unfolds, linking expenditures to the physical parameters of the project.

4. Services and Programs

Key information relating to the "software" elements of the project should be discussed in this section. Events, issues, changes, difficulties or delays should be discussed along with expected effects on the implementation schedule and costs. Reporting should be assembled into at least the following subheadings:

(a) Plot Allocation 1/

The status of this phase should be discussed, including the following selection steps--advertising plots, community outreach efforts, processing applications and actual allocations. Delays should be noted along with the reasons for these delays.

(b) Construction Loans

A description of the progress of loan disbursements should be given including the income ranges reached. The status of loan collections/payments should also be discussed. Whenever defaults on loans diverge substantially from appraisal or revised estimates, further information should be submitted on possible causes for these defaults. Table 7 should provide a useful reporting format.

(c) Income and Business Support

Market upgrading, market construction, lorry parks, open-air garage workshops, and other project elements relating to income and business support should be addressed in this subsection and compared with implementation schedules.

(d) Community Development

Because this component is a relatively novel project element it must be closely monitored and reported on a regular basis. A timetable should be prepared for specific community relations activities which should facilitate monitoring. A number of complementary monitoring strategies were discussed in Chapter 2.

5. Project Management

This subcategory should reflect changes in the staffing situation, recruitment and staff training programs. Any actual or perspective changes in key personnel should be reported here.

6. Attitudes of Project Beneficiaries

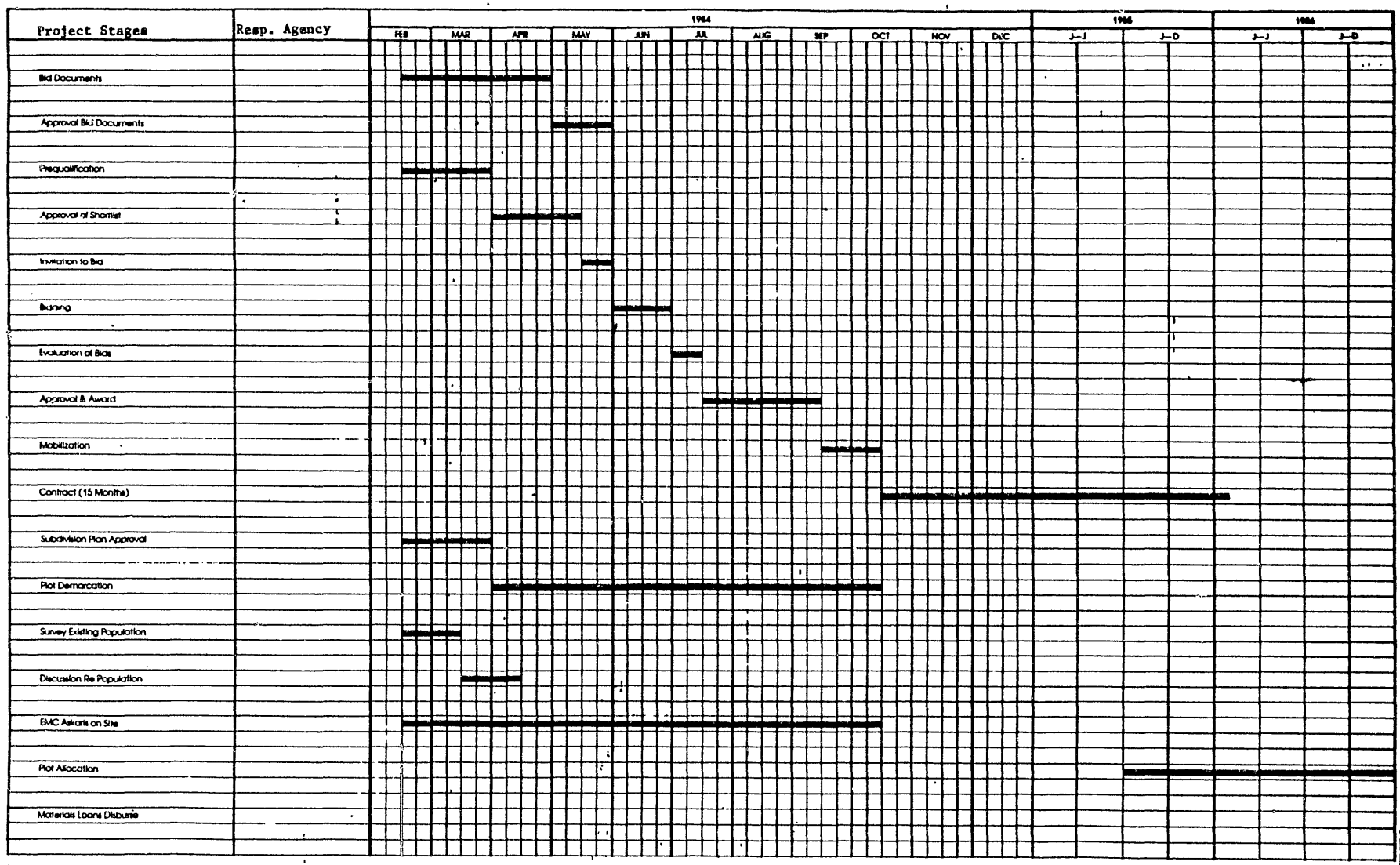
In addition to regular information on the physical and financial progress of projects, managers also need regular feedback on the attitudes of targetted

1/ See also: "Designing the sites and Services Plot Allocation Process: Lessons from Project Experience," Lauren Cooper. World Bank Technical Paper No. 3, November 1981; and "House Registration Handbook: A Model for Registering Houses and Plots in Unplanned Settlements," Saad Yahya. World Bank Technical Paper No. 4, 1982.

beneficiaries to assess the project's effectiveness in responding to their needs. Experience has shown that serious problems can arise when effective communication linkages are not maintained with beneficiaries, and for this reason the Quarterly Progress Report should regularly include a section on attitudes of project beneficiaries. Table 8 presents a completed sample survey of the attitudes of households involved in a hypothetical sites and services project.

TABLE 1

Implementation Schedule (3-bar Chart)

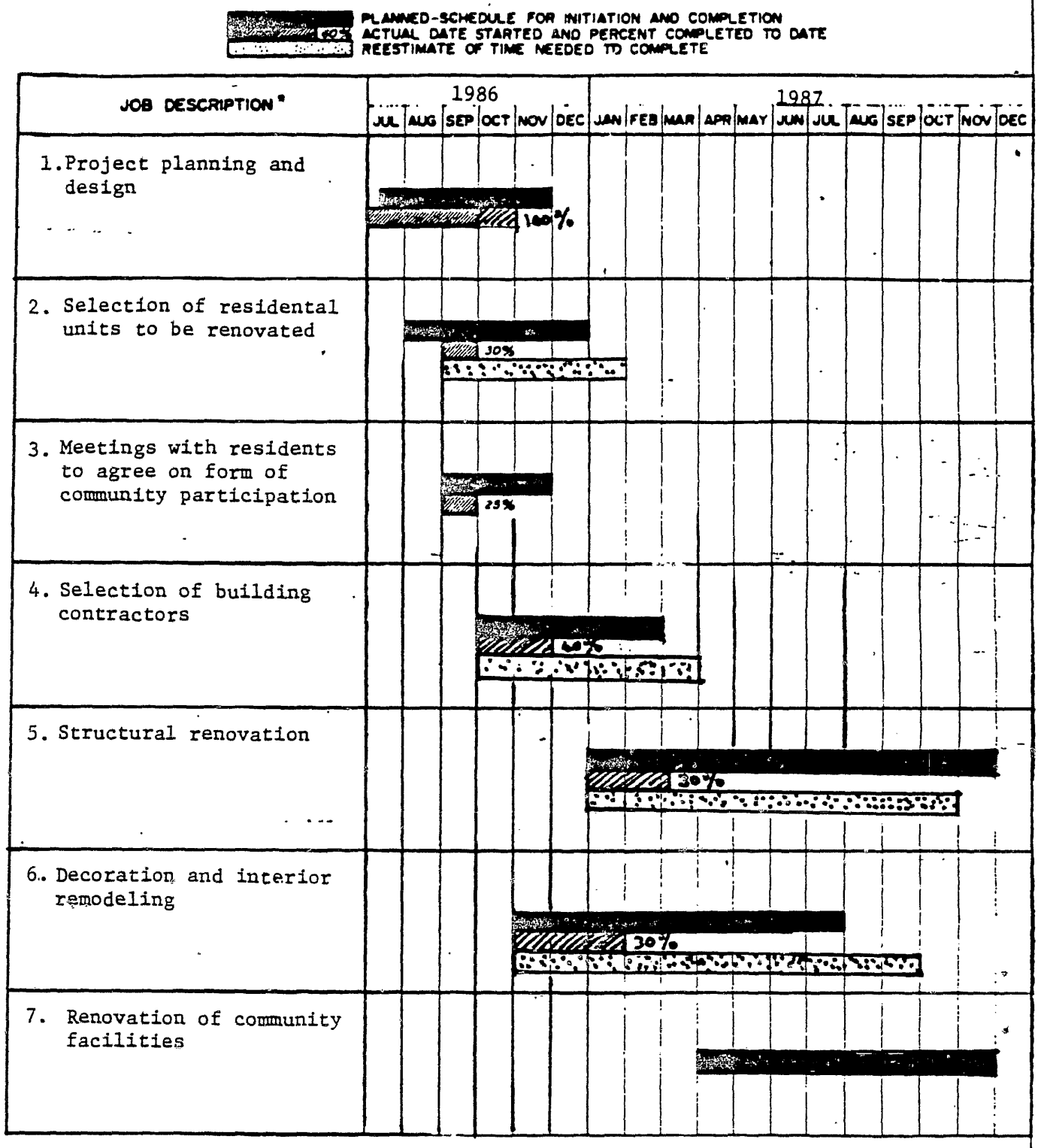


Legend
- Planned Schedule for Initiation & Completion
- Actual Date Started & Percent Completed to Date
- Reestimate of Time Needed to Complete

Note: This table provides a breakdown of actual project components by stages, taken from a recently approved Bank urban project.

Table 2: CHART FOR MONITORING THE PROGRESS OF PROJECT IMPLEMENTATION
Example: Renovating old houses

DATE JANUARY 1987



Source: Material prepared for the Monitoring and Evaluation Module of the Economic Development Institute Course on Urban Project Preparation and Implementation to be given in Shanghai in May/June 1985

TABLE 3

CONSTRUCTION PROGRESS SHEET

Location:
Type of Work:

Name of Contractor:

Original Contract Value: | Revised Contract Values:

Contract Period:

Starting Date:

Original Completion Date: | Revised Completion Date:

PROGRESS

Year																											
Calendar Month																											
Contract Month																											
Original Contract Time Elapsed																											
Revised Contract Time Elapsed																											
Financial																											
% Completion Physical (actual)																											
(expected)																											

NOTES: Include here changes (such as variation orders and wage increases) which have an impact on total contract value.

ACTION RECOMMENDED:

TABLE 4

CONTRACT DATA SHEET

Contract No./Package	Locations of Work	Date of			No. of Bids Received	Contractor Name	Contract Amount	Proposed	
		Bid Opening	Bid Closing	Award				Commencement Date	Completion Date

TABLE 5

SUMMARY DATA ON CONTRACTORS PAYMENT

Date: _____

Contract Number/ Package	Original Contract Amount	Change Orders & Escalations To Date	Total Contractors Payments to Date	Expected Remaining Contractors Payment	Contractors Payment Schedule							
					1984		1985		1986		1987	
					Actual	Projected	Actual	Projected	Actual	Projected	Actual	Projected

TABLE 5

Project Control Sheet

Project Budget Item	Budget as per investment agreement (1)	Variances to arrive at control budget (2)	Control budget at start of project (3)	Contracts and orders placed To date (4)	Accountable expenditures to date (5)	Reestimate of budget item as of report date (6)	Estimated percentage of completion of project items (7)
A. Infrastructure							
Street Paving							
Bridges							
Electrification							
Traffic Signs							
Sewerage							
Water							
B. Community facilities							
Sports Ground							
Social Centers							
Health Centers							
Kindergartens							
Other Buildings							
C. Equipment & Vehicle							
D. Project Management							
E. Technical Asst.							
F. Cons. Loan Program							

Explanatory Notes of the Above:

A. PROJECT BUDGET ITEM:

Items A through F are shown as examples of typical items in a project budget: every project budget should be set up with sufficient detail to give a clear overall picture as of the report date of the status of the project budget in all its principal elements.

B. COLUMNS (1) through (7)

Column (1): Summary breakdown of budget into principal items, agrees with the capital requirements outlined in the Investment Agreement.

(2): shows specific variations in the Capital Budget items due to refining of estimates or other adjustments to establish the final Project Control Budget

(3): final approved project budget established for accounting control: it does not necessarily differ from (1). Amounts in Columns (1), (2) and (3) are never changed once control system is established.

(4): total values of firm contracts, orders and financial commitments to date for the project.

(5): shows amounts disbursed as at date of report.

(6): shows re-estimates of total capital required for each line item as of the date of the report

(7): is an estimate of the physical progress toward completion and is not necessarily a reflection of the amounts spent on a particular item.

C. RELATION TO BALANCE SHEET

The toal of column 5 is to be equal to the total of the account or accounts on the project contained in the assets of the balance sheet should there be any differences, these should be explained in a separate statement.

Construction Loan Program
Residential

Date:
Town:
Site:

	Original Target at Appraisal		Loans in Process		Loans Approved		Loans Fully Disbursed				No. of Loans by Income Group		
									Cumulative Dist.				
Calendar Year	No. of Loans to be disbursed	Amount	No.	Amount	No.	Amount	No.	Amt	No.	Amt	Lowest 3rd	Mid 3rd	Highest 3rd
1984 II													
III													
IV													
1985 I													
II													
III													
IV													
1986 I													
II													
III													
IV													
Total													

TABLE 7

TABLE 8: SUMMARY OF INTERVIEWS ON PARTICIPANT SATISFACTION AND INFORMATION: PROJECT "LA ESPERANZA"

Part 1: Satisfaction with the Project	Families Already Living in The Project [20 Interviews]			Families Who Have Not Yet Moved to the Project [20 Interviews]		
	Not Satisfied	Satisfied	No Reply	Not Satisfied	Satisfied	No Reply
General satisfaction with the house	15	3	2	14	4	2
Satisfication with the toilet	12	7	1	Not Applicable		
Power supply (electricity)	8	9	3	Not Applicable		
Garbage collection	9	7	4	11	5	6
Bus service	16	2	2	15	3	2
Drainage (flooding)	15	1	4	16	2	2
Security (need for night watchman)	12	5	3	8	8	4
Opening hours of project office	4	8	8	5	7	8
Cost of building materials	2	15	3	4	13	3

Part 2 Knowledge about this Project	know about	not know about	know about	not know about
1. Plans for a cooperative	15	5	10	10
2. Organization and program of this junta	18	2	8	12
3. Sports club	10	10	4	16

ANNEX H: DESIGNING A NETWORK BASED SYSTEM TO MONITOR PROJECT IMPLEMENTATION

1. Introduction and objectives

The system presented in this Annex is based on a modified PERT/CPM model and has the following objectives:

(a) To present a simple methodology which can be used in the initial planning of the project. The System will indicate potential bottlenecks and where the sequence of project activities should be rescheduled to reduce some of these bottlenecks.

(b) Preparation of a calendar of activities for each project component.

(c) Indicate the impact of delays on overall project implementation and the need to re-estimate completion dates.

(d) Identify key aspects of project implementation on which management needs to take actions.

2. Designing the System

(a) Defining the Project Components and their main stages

The first step in designing the system is to identify all of the components of the Project and their main stages. Chart 1 presents a list of components for a typical sites and services project. Ten components are identified: Design; Land Acquisition; Tendering; Office Infrastructure; Construction of Core Unit by Contractor; Selection of Participants; Materials Loans Program; Completion of Habitable Unit; Occupation of Units; and Start of Cost Recovery.

Each component has a number of stages. In the example, the four stages of the Land Acquisition component are: (i) identification of possible sites; (ii) identification of owner(s) and verification of title; (iii) topographical and other feasibility studies; and (iv) finalization of purchase agreement. Similarly, nine stages were identified for Participant Selection and Training. There is no firm rule on how many stages should be identified, but the number should be small enough for each stage to represent significant progress but sufficiently large to permit monitoring of progress. An arbitrary number of stages was given to each of the other components for illustrative purposes.

The meaning of Nodes will become clear in the discussion of the Network Chart

(b) Estimating the time required to complete each component

A chart should be prepared for each component indicating the estimated time for completion of each stage. The chart should indicate both the likely completion time and the possible delays which might occur. Obviously there is an element of unpredictability in many types of delay, but experience can often provide a general indication.

Charts 2 and 3 give examples of the estimation of completion time for land acquisition and for the selection and training of participants. In the case of land acquisition, all stages are sequential so that one cannot be started until the previous one is completed. However, for selection of particiants a number of stages can be undertaken at the same time. Each chart indicates the preceding component which must be completed before this component can begin. There is also an indication for each stage of possible causes of delay and the amount of delay which might occur. The cumulative delay is indicated at the bottom of the chart. Finally, the component which follows that being studied is indicated.

(c) Constructing a logical network of project implementation

Once the time required for each component has been estimated, it is then possible to construct a Logical Network which analyzes the linkages between different components and estimates the time required for the completion of the total project. The proposed system is a simplified version of PERT/CPM. Chart 4 presents a typical network chart. The stages in its construction are the following:

(i) A rough sketch is prepared of the logical sequence of project components and of the linkages between them. The sequence follows approximately the order of components indicated in Chart 1.

(ii) Stage of completion of a component (as per Chart 1) is indicated on the Chart as a circle (called a Node). The left hand side of the Node indicates its number. Numbering approximately follows the logical sequence of the Nodes.

(iii) Nodes which logically follow each other are indicated by arrows. For example, Node 5 (identification of owners of land) logically follows Node 4 (identification of possible project sites) and the two are connected by an arrow.

(iv) A number is placed on each arrow to indicate the estimated number of weeks to complete the stage following the arrow. For example it is estimated that four weeks will be required to complete stage 1 of the project design (number 4 on the arrow leading to Node 1).

(v) The number in the upper right hand side of each Node indicates the shortest number of weeks from the start of the project in which this stage could be completed. This number is obtained by adding the estimated number of weeks for all stages leading to this one. Care must be taken in the estimation of this number. For example, the minimum number of weeks to complete Node 4 is eight weeks. This is because it cannot be completed until Node 2 (part of project design) has been completed, and it is estimated that eight weeks will be required to complete Node 2. Any stage whose completion is conditional upon the completion of two or more separate components, is indicated by a double circle, as this is a point where a potential bottleneck could occur.

(vi) The number in the bottom right corner of each Node indicates the maximum number of weeks which might be required for the completion of this stage if all anticipated delays actually do occur.

(vii) At some points in the chart, two or more arrows connect to a single Node at the point where two components are completed (for example where the Off-site Services and Tendering components merge at the start of the construction of core units at module 15). The arrow which represents the longest time path has a number in the square box. This indicates the number of weeks of Lag-Time available for the completion of this component without causing delays in the start-up of the following component. For example, the arrow between Nodes 3 and 8 indicates there is a Lag-Time of five weeks between the estimated completion of project Design and the start-up date for Tendering.

The Lag-Time indicators are important because the smaller the Lag-Time the greater the potential for delay in start-up of the following Component. For example, there is a Lag-Time of zero weeks between Nodes 27 and 28. This means that any delay in the completion of either Selection of Participants or the implementation of the Materials Loan Program will automatically produce a delay in the start-up of the Completion of a Habitable Unit. This point is, therefore, a potential bottleneck and will need to be monitored carefully by management.

(viii) In some cases one network loop may intersect with another in the middle of the implementation of a component. For example, Node 2 of the project design process must be completed before feasibility studies can be conducted (Node 6) in the process of land acquisition. In this case, the link between Nodes 2 and 6 has the number -2 in a box, indicating that the completion of this stage of the design is delaying by two weeks the land acquisition process.

(ix) Finally, it is possible to compute the total estimated time for the completion of the project. If all steps are completed according to their estimated time, the project will be completed in 146 weeks. if all of the possible delays were to take place, the total time would be increased to 175 weeks.

(x) On the basis of the above computations, a Diagnosis of Implementation Time and Potential Bottlenecks can be made (Chart 5). This provides management with two sets of guidelines as to ways in which project implementation time could be reduced. Firstly, it indicates the contricution to total project implementation time of each component. The first column shows the number of weeks required to complete each component while the second column indicates the number of additional weeks of elapsed time required to complete each component. For example, the process of land acquisition requires 13 weeks, but much of this can be done at the same time as project design so that only 3 weeks are added to elapsed time. It is clear from this chart that any significant decrease in project implementation time can only be achieved by reducing the time of tendering, constructi of the core unit or selection. Elapsed time can be reduced either by completing a component more quickly or by altering the scheduling. For example, it might be possible to start the selection process while the construction of the core unit is still underway.

Secondly, the potential bottlenecks are identified, together with the time period in whih they are likely to occur. This suggests the need for special attention to monitoring during these critical periods to identify any potential delays or problems at an early stage.

(d) Preparing the implementation calendar for each component

Chart 6 gives an example of how the estimates produced in the earlier charts are used to produce an implementation calendar for each proejct component. This example shows a chart for Project Design and Land

Acquisition. The two components, which have relatively few steps, have been combined into a single chart. As by this stage the date for the start of the project will have been decided, it is possible to put dates on the implementation calendars.

Project design has three steps which are sequential. The bar chart indicates the estimated number of weeks for each step and the starting and finishing dates. The connecting arrows indicate that the start-up of a step is conditional on the completion of previous one.

Land acquisition has four steps, the first of which will begin at the same time as Project Design. However, it can be seen that step 3 cannot begin until the completion of step 2 of project design. This causes a delay of two weeks in the start-up date for step 3. The spaces labled "Lag" indicate the number of weeks between the completion of a step and the start-up of the next step which is conditional on it. This indicates a margin for delay which will not hold-up the overall project schedule.

The chart indicates the preceding steps which must be completed before the components can be implemented. These are blank in the present case, as the two components are the first in the project schedule. The chart also indicates the following steps, which in the present case are Tendering and the provision of off-site services.

3. Designing the monitoring system

(a) Monitoring the implementation schedule of each component

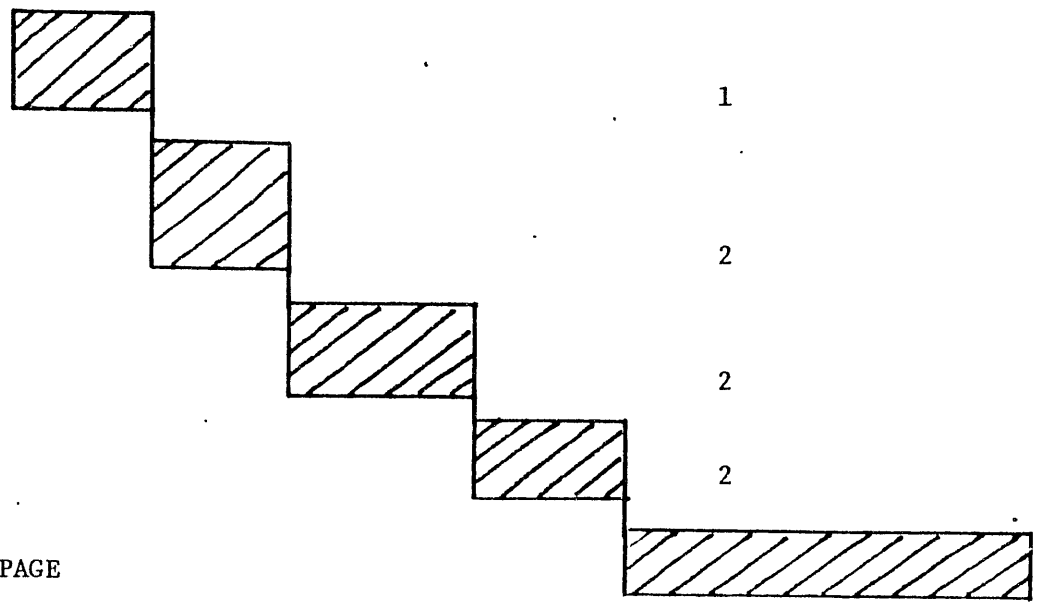
Chart 7 presents a simple system for recording the delays in project implementation of each component and their effects. A chart of this kind should be prepared for each component on which work has, or should have, begun. For each step the chart indicates the delays, if any, in start-up dates, length of implementation and completion dates. An important feature of the chart is that it also indicates the causes of delays, their impacts on the project, and possible actions which could be taken.

An important feature of the chart is to demonstrate the interactions between the different project components and the cumulative impact of delays in any particular component. This is an important tool for management as project monitoring is often compartimentalized so that the information collected on each component is often not synthesized to present an overall picture of the status of the project and the seriousness of the different delays.

CHART 1
LISTING COMPONENTS AND STEPS OF THE PROJECT

COMPONENT	STAGE		NODE
Project design	1	Definition types and nos	1
	2	General design and costs	2
	3	Detailed design	3
Land acquisition	1	Identification of sites	4
	2	Identification of owners and verification of title	5
	3	Feasibility studies	6
	4	Negotiation and acquistion	7
Tendering	1		8
	2		9
	3		10
	4		11
Off-site services	1		12
	2		13
	3		14
Construction of core unit	1		15
	2		16
	3		17
	4		18
Selection of participants	1	Definition types, numbers, and costs of units	19
	2	Selection criteria	20
	3	Preparation publicity	21
	4	Mass media publicity	22
	5	Distribution application form	23
	6	Interviews with applicants	24
	7	Analysis characteristics of applicants	25
	8	Selection applicants and reserves	26
	9	Orientation sessions	27
Completion habitable unit	1		28
	2		29
Material loans	1		30
	2		31
Occupation	1		32
Start of cost recovery	1		33
	2		34

CHART 2: IDENTIFICATION OF MAIN STAGES AND ESTIMATION OF COMPLETION TIME FOR SUB COMPONENT: LAND ACQUISITION

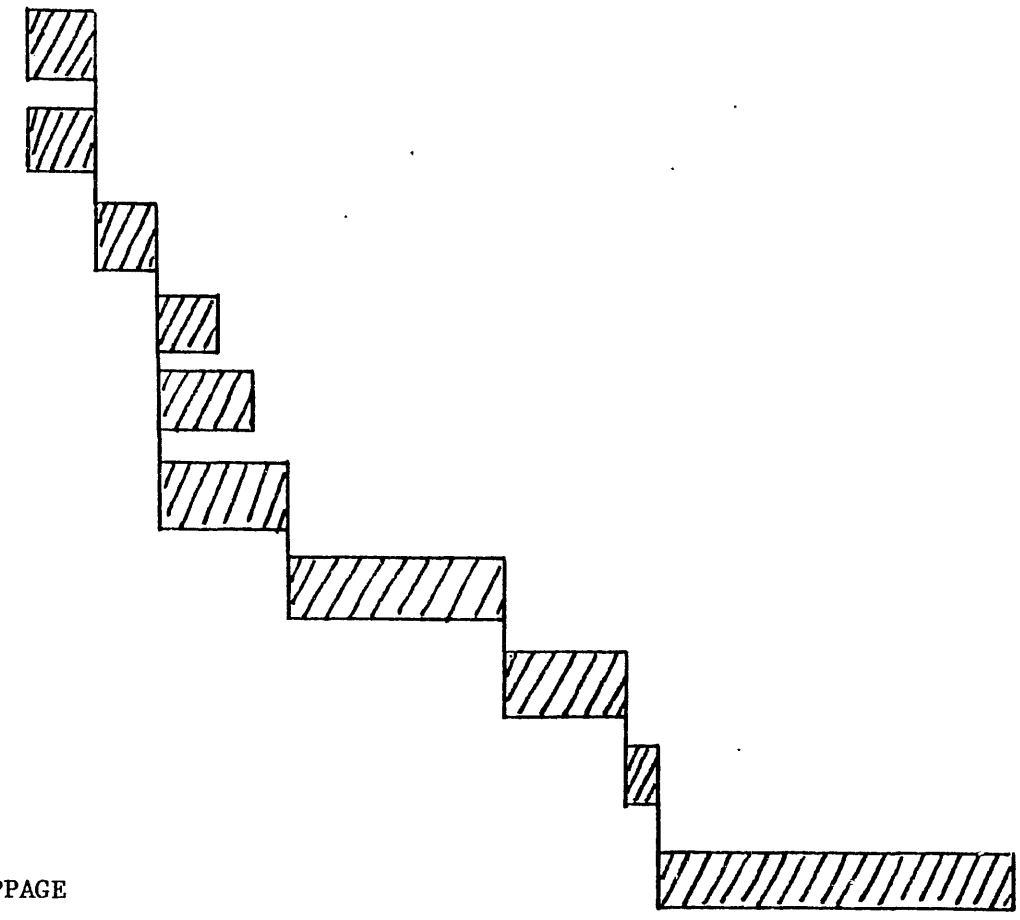


Preceding Activity	Stage	Potential Slippage (weeks)	Causes	Following activity
Start of Project	1. Identification of possible sites	1		
	2. Identification of owners and checking title	2		
	3. Site feasibility study	2		
	4. Purchase negotiations	2		Offsite services Tendering
	CUMULATIVE SLIPPAGE			

Total projected time 13 weeks
Total with slippage 20 weeks

CHART 3 IDENTIFICATION OF MAIN STAGES AND ESTIMATION OF COMPLETION TIME FOR SUB COMPONENT: PARTICIPANT SELECTION

Preceding Activity	Stage	Potential Slippage (Weeks)	Causes	Following Activity
Construct. Core unit	1. Define no./ types of unit	1		
	2. Selection criteria	1		
	3. Publicatinn material	1		
	4. Publicity	1		
	5. Distribute applications	1		
	6. Interview applicants	2		
	7. Analysis applications	2		
	8. Selection participants	1		Material loans Completion habitable unit
	9. Orientation sessions	2		
	CUMULATIVE SLIPPAGE			



Total projected time 19 weeks
Total with slippage 31 weeks

CHART 4 EXAMPLE OF A LOGICAL NETWORK CHART OF PROJECT IMPLEMENTATION FOR A SITES AND SERVICES PROJECT

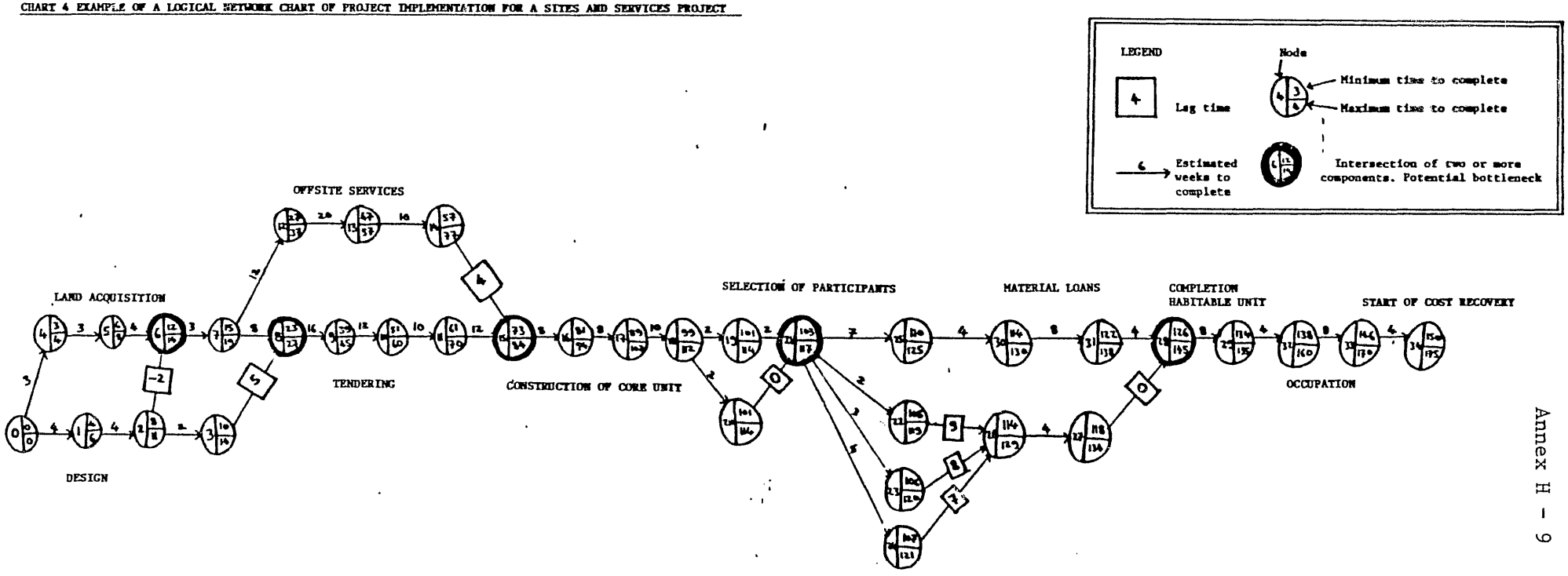


CHART 5: EXAMPLE OF TABLE DIAGNOSING ESTIMATED PROJECT DURATION AND IDENTIFYING POTENTIAL BOTTLENECKS

(Data is used from Chart 4)

COMPONENT	DURATION (Weeks)	CONTRIBUTION TO TOTAL PROJECT DURATION (Weeks)
Design	10	10
Land acquisition	13	3
Tendering	46	46
Offsite services	42	0
Construction of core unit	38	38
Participant selection	19	19
Material loans	12	4
Completion of habitable unit	12	12
Occupation	4	4
Start cost recovery	12	12
TOTAL	202	148

POTENTIAL BOTTLENECKS

	Critical weeks
Design and feasibility study for land acquisition (Nodes 2 and 6)	8-12
Completion of design and land acquisition and tendering for core units (Node 8)	15-23
Completion of tendering and offsite services with start of construction (Node 15)	37-66
Completion of various steps of selection process at same time (Node 26)	110-114
Completion of selection and setting up material loan program with completion of habitable unit (Node 28)	118-126

CHART 6: CALENDAR OF PROJECT IMPLEMENTATION: DESIGN AND LAND ACQUISITION COMPONENTS (Data from Chart 4)

Preceding Activity	Component	Stage	Node	JAN 1984 8	15	22	29	FEB 5	12	19	26	MARCH 5	12	19	26	APRIL 2	9	16	Following activity
Project starts	Design	Define types and nos. of units	1																
		General design	2																
		Detailed design	3																Tendering
	Land acquisition	Identify possible sites	4																
		Identify owners	5																
		Feasibility studies	6																
		Negotiation	7																Tendering Offsite services

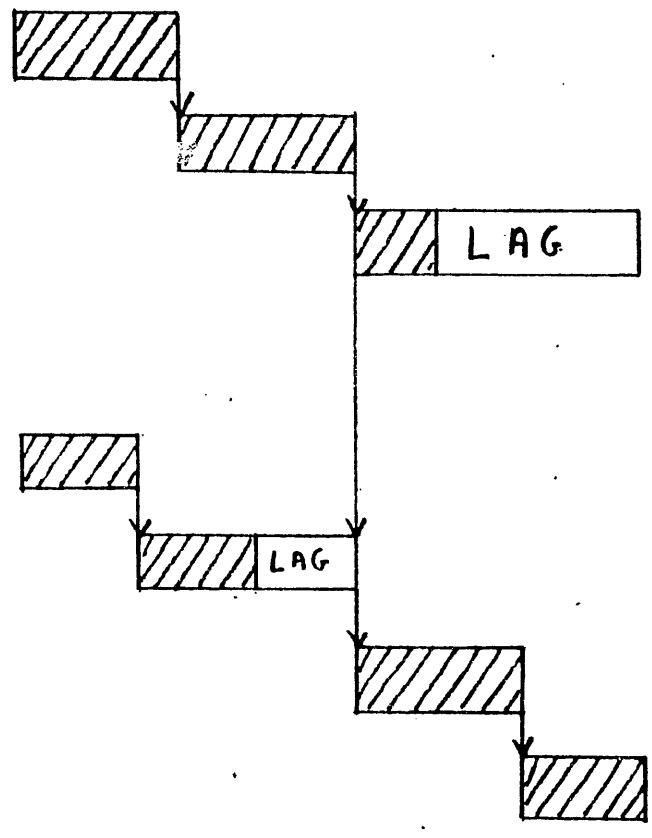


CHART 7: SUMMARY OF DELAYS IN PROJECT IMPLEMENTATION AND THEIR CAUSES AND CONSEQUENCES

COMPONENT _____

STAGE	START-UP DATE			DURATION OF IMPLEMENTATION			COMPLETION DATE			SUMMARY OF DELAYS (weeks)		
	Plan	Actual	Expected	Plan	Actual	Expected	Plan	Actual	Expected	Start-up	Implementing	Total

Original date for completion of component: _____
Expected/actual date for completion: _____
Delay (weeks) _____

Main reasons for delay: _____

Effects of delay: _____
Recommended actions: _____

ANNEX I: DESIGNING A SYSTEM FOR NETWORK-BASED FINANCIAL MONITORING

1. The elements of network-based budgeting

A network-based budgeting system (NBB) is a financial planning and monitoring system in which cost items and expenditure patterns are linked to components and steps in the implementation process. This means that changes in the process or scheduling of project implementation will automatically be reflected in estimates of costs, expenditures and disbursements.

The main elements of the NBB system which will be described are:

- Definition of cost-bearing activities.
- Estimation of expenditure schedules for each component.
- Cost variance analysis.
- Revision of cost and implementation schedules.

2. Definition of cost-bearing activities

The first step in setting up the NBB system is to identify all cost-bearing activities. Chart 1 presents an example of the definition of cost-bearing elements in the construction of a pedestrian subway. A total of 14 items are defined, each with its estimated base cost at the time the project was appraised.

For the purposes of network planning, it is important to ensure that the listing of cost-bearing items is consistent with the steps defined in the physical implementation process. This is important as it will be necessary to develop integrated charts which show both the progress of physical implementation and of expenditures.

A problem always arises in that not all costs can be directly assigned to specific project components. For example, it is usually not possible to assign administrative overheads between particular components. The usual procedure is only to include direct costs which can be directly assigned. Overheads and other indirect costs will then be distributed over time in some logical way and will be included as a separate item.

3. Estimation of expenditure schedules for each component

The next stage is to estimate the schedule of expenditures over time. The first step is to prepare a calendar of the physical implementation schedule of

the project (see, for example, Charts 2 and 3 in Annex 3). It is also necessary to obtain information on the payment schedule as this may differ significantly from the implementation schedule. In some cases, expenditures will precede physical implementation as some contractors receive an advance payment before work begins. In other cases, payments are made after work is completed. Expenditure schedules will also vary for different institutions as some funds are on-loaned or advanced from one organization to another.

A decision must also be made of how to define estimated start-up and completion dates. Early and late start-up and completion dates will often have been estimated. One common approach is to use the mid-point of the early and late estimates.

4. Monitoring cost and expenditure schedules

Chart 3 presents a simple system for monitoring project costs. This provides, for each cost-bearing activity, the following information:

(a) Expected total cost at time of appraisal (Column 2).

(b) Current expected total cost (Column 3).

(c) Percentage difference between Columns 2 and 3.

(d) Appraisal estimate of expenditure for current financial year (Column 5).

(e) Current estimate of expenditure for current financial year (Column 6).

(f) Percentage difference between Columns 5 and 6.

(g) Appraisal estimate of expenditure during reporting period (Column 8).

(h) Current estimate of expenditure during reporting period (Column 9).

(i) Percentage difference between Columns 8 and 9.

(j) Current estimate of cumulative expenditure to this quarter (Column 11).

(k) Percentage difference between Column 11 and appraisal estimate (Column 12).

(l) Current estimate of expenditure for next quarter (Column 13).

(m) Percentage difference between Column 13 and appraisal estimate (Column 14).

Chart 4 presents similar information, but where the cost status of each project is combined to summarize the overall cost status of the complete Program.

Chart 5 shows a convenient way to summarize the status of physical progress and costs in a single chart. This is often a very useful chart as it demonstrates the relationship between delays in physical implementation and the expenditure and disbursement schedule.

5. Cost Variation Analysis

An essential complement to the monitoring of costs is an analysis of the causes of observed variations in costs. Variations in costs can be caused by three main factors:

(a) Price variations. These can either occur when the project has been implemented according to its original schedule and where unexpected price increases have taken place, or they can be due to delays in project implementation which means that prices have been affected by inflation.

(b) Contractor performance. Variations in the quantity or quality of work performed by contractors. An example of a change of quantity would be when the contractor uses more or less materials than had been budgeted. A change in quality would be when the contractor changes the types of materials being used.

(c) Project design. Changes in the quantity or quality of work specified by the project design.

Chart 6 presents a system for summarizing cost variations and for identifying the factors which have produced these changes. The chart permits a percentage breakdown of cost variation for each activity between changes in Prices, Quantity and Quality of Contractor Work and Project Design.

6. Revision of cost and implementation schedules

The main purpose of monitoring costs and implementation schedules is to assist management in defining actions needed to correct problems and to make the necessary revisions in cost, disbursement and implementation schedules. The minimum frequency for updating estimates is once every financial year, but in most cases it will be necessary to do this every quarter or semester.

CHART 1

DEFINING COST-BEARING ACTIVITIES

Example of the Construction of a Pedestrian Subway

Direct Costs	ACTIVITY	BASE COST (Rupias in Lakhs)
1	Fund deposit (1)	50
2	Fund deposit (2)	100
3	Fund deposit for junction	25
4	Local fund deposit	25
5	Contractor mobilization and diversion road	23.5
6	Earth works (station)	34.7
7	Levelling course (1)	54.8
8	Tunnel and staircase (1)	212.2
9	Roadworks and traffic diversion	93
10	Earthworks (1)	34.6
11	Levelling course (2)	56.4
12	Tunnel and staircase (2)	221.2
13	Roadworks and approach slab	155.8
14	Finishing works	30
TOTAL DIRECT COSTS		
INDIRECT COSTS		

Source: This example is a simplified version of Exhibit C-1 of "Project Monitoring and Performance Evaluation: Systems Manual". Developed by Operations Research Group Baroda for the Madras Metropolitan Development Authority as part of the First Urban Development Project.

CHART 2: SCHEDULE OF EXPENDITURES FOR EACH COMPONENT

STAGE	NODE	1983					1984					FINAL
		JAN-MAR	AP-JUN	JUL-SEP	OCT-DEC	1983 TOTAL	JAN-MAR	AP-JUN	JUL-SEP	OCT-DEC	1984 TOTAL	TOTAL

CHART 3: PROJECT COST SUMMARY

Activity	EXPECTED TOTAL COST			BUDGET FOR FINANCIAL YEAR			EXPENDITURE DURING REPORTING PERIOD			CUMULATIVE EXPENDITURES			BUDGET NEXT QUARTER	COMMENTS
	Plan	Actual	% change	Plan	Actual	% change	Plan	Actual	% change	Plan	Actual	% change		

CHART 4: SUMMARY OF COSTS AND EXPENDITURES FOR TOTAL PROGRAM BROKEN DOWN BY PROJECT

Component	Implementing Agency	COST TO COMPLETION			BUDGET FOR FINANCIAL YEAR			EXPENDITURES				Comments
		Plan	Actual	% change	Plan	Actual	% change	Current Quarter	Cumulative to date			
									Plan	Actual	% change	

CHART 5: INTEGRATION OF PHYSICAL PROGRESS AND COST STATUS ANALYSIS

COMPONENT OR PROJECT _____

STAGE	Node	1983 JAN FEB MAR AP MAY JUN JUL AUG SEP OCT NOV DEC 1984 JAN FEB	Cost at completion			Expenditure			
					%	This quarter		Cumulative	
			Plan	Actual	change	Plan	Actual	Plan	Actual
		TOTAL							

KEY

Original plan

Actual dates

→ Expected completion

CHART 6: COST VARIANCE ANALYSIS

PROJECT/ COMPONENT _____

Activity	EXPECTED FINAL COST		ANALYSIS OF COST VARIATION								SUMMARY OF VARIATION			
			Price Variation		Contractor Variation				Project Design Changes		Percentage change from plan			
					Quantity of Work		Quality of work							
	Plan	Current	Plan	Current	Plan	Current	Plan	Current	Plan	Current	Price	Quantity	Quality	Total
TOTAL														

ANNEX J: A BASIC EVALUATION LIBRARY

The purpose of this annex is to recommend books which could be included in a basic evaluation reference library. Many researchers do not have easy access to libraries and research journals so it will often be necessary to acquire a basic set of books so as to be relatively self-sufficient. There are a very large number of books and journals on evaluation and the list is intended to be suggestive rather than definitive.

General textbooks

1. Joseph Wholey "Evaluation: Promise and Performance". The Urban Institute. Washington, D.C. 1979.
   This is a very clear and readable guide to the design and implementation of evaluations. It is particularly good on the policy and organizational aspects, but does not enter into as much detail on the research design as some other books. Like most other textbooks it refers primarily to the experience in the evaluation of public programs in the U.S, so that many of the examples are not directly relevant to the issues being faced in many developing countries.

2. H. Freeman, P. Rossi, and S. Wright "Evaluating Social Projects in Developing Countries." Organization for Economic Cooperation and Development. Paris. 1979.
   This draws a lot of material from an earlier book by Rossi and Williams on the evaluation of social programs in the U.S. It is an easily understandable overview with the advantage that the case studies and bibliography are drawn from developing countries rather than the United States (as is the case with most of the other texts).

3. E. Struening and M. Brewer (Editors) "Handbook of Evaluation Research" Sage Publications. California, 1983.
This is a one volume selection from a very influential handbook which provides a useful overview of many of the key issues and stages in the design, implementation and analysis of an evaluation. All of the articles are quite old, but each chapter contains an updated bibliography.

4. Dennis Casley and Dennis Lury "Monitoring and Evaluation of Agricultural and Rural Development Projects" Johns Hopkins (1982). Although this refers to rural projects, most of the concepts are equally applicable in urban areas. Clear and simple to follow.

More specialized but easily understandable references books

1. T. Cook and D. Campbell "Quasi-experimentation: Design and analysis.
This book presents a very complete discussion of the design and analysis of all of the main types of quasi-experimental designs. Several chapters are devoted to the use of multiple regression analysis to compensate for problems of non-equivalency of control groups.

2. E. Webb, D. Campbell and others "Unobtrusive Measures: Non-reactive research in the social sciences." Rand McNally, 1973.
This is the classic on the way in which a wide range of observational and other non-reactive techniques can be applied in evaluation research.

3. John Van Maanen (Editor) "Qualitative Methodology" Sage Publications, California, 1979.
A comprehensive presentation of the use of qualitative methods in the evaluation of organizations.

4. Gareth Morgan (Editor) "Beyond method: Strategies for Social Research". Sage Publications, California, 1983.
A comprehensive description of all of the main methods which can be used in organizational research. One of the purposes of the book is to demonstrate how the choice of research method is likely to influence the findings of the research and the conclusions which will be drawn. This is an extremely important issue which is hardly discussed in most textbooks.

5. Michael Patton "Qualitative Evaluation" Sage Publications, 1980.
A very readable and well argued presentation of the importance of qualitative methods by one of the leading exponents of these approaches.

6. Lisa Peattie "The View from the Barrio" University of Michigan, (1969).
A very readable book on how to understand and describe the dynamics of a small urban community and how to study the ways in which the community organizes itself to oppose the construction of a sewage outlet which would contaminate the water supply. It is also a very good description of an activist approach to participant observation.

7. Lyn Squire and Herman van der Tak "Economic Analysis of Projects". Johns Hopkins, (1975)
Comprehensive explanation of the World Bank's approach to cost-benefit analysis.

8. Michael Bamberger "A basic methodology for impact evaluation in urban shelter programs". Water and Urban Development Department. The World Bank, 1984.
This is a more detailed presentation of the techniques for impact evaluation which are described in the present handbook.

Basic statistical textbooks

1. H. Blalock "Social Statistics" McGraw Hill, 1979.
   Clear and comprehensive coverage of most of the main research procedures required in evaluation research.

2. H. Blalock and A. Blalock (Editors) "Methodology in Social Research" McGraw Hill, 1968.
   More advanced discussion of theory building and the construction of models.

3. Michael Lewis-Beck "Applied Regression: An Introduction". Sage Publications, California, 1980.
   A clear, brief and low-priced presentation of the basic concepts of regression analysis which are likely to be needed in basic evaluation research.

4. Christopher Achen "Interpreting and using regression" Sage Publications, 1982.
   A complement to the previous title.

5. N. Nie and others "Statistical Package for the Social Sciences", McGraw Hill, 1975.
   This is probably still the most widely used statistical computer package for the social sciences and includes all of the analytical procedures required for most types of evaluation. This is also one of the most comprehensive and easily understood explanations of how a wide range of statistical procedures are used and interpreted.

ANNEX K: GLOSSARY OF TERMS USED IN THE HANDBOOK

COMPONENT — A specific set of activities which form part of a project and which are conducted by one agency. For example, the construction of 5,000 serviced plots might be a component of an urban development project.

CONTROL GROUP — A group of families, organizations or communities which are used in impact evaluation. The control group is selected to match as closely as possible the characteristics of the project beneficiaries in order to estimate what would have been the conditions of participants if the project had not taken place.

COST-EFFECTIVENESS ANALYSIS — Projects are compared in terms of the costs required to produce a given output. The project with the highest output/cost ratio is considered to be the most cost-effective. This method is particularly useful when policy makers wish to select between alternative projects, but when it is not possible to measure impacts,

EVALUATION — Evaluation establishes criteria for defining success and assesses the extent to which these criteria have been achieved by the project. Evaluation can be quantitative, in which case numerical values are estimated for the net project impacts or for the cost-effectiveness ratios; or it can be qualitative, in which case the purpose is to understand and describe the way in which the project has affected, and been affected by, the population groups who have been exposed to it.

EXPERIMENTAL DESIGN — A true experimental design involves the random assignment of subjects (persons, groups etc) to either an experimental or a control group. The condition of both

groups is measured before the experiment begins. The "treatment" (drug, special reading program etc) is then administered to the experimental group. The conditions of the two groups are then compared after the treatment, and statistical tests are used to determine whether there is a statistically significant difference between the two groups. If such a difference exists, it will be taken as evidence that the treatment had an effect. The careful researcher will, however, repeat the experiment several times under slightly different conditions to check whether the same results are found.

IMPACT

A social or economic change which is produced as a consequence of a project. Examples include: changes in household income and employment; lower rates of certain infectious diseases; jobs created by a small business program; reduced travel time to work. Impacts differ from outputs in that the latter are directly produced by the project whereas the former occur as a consequence of the project outputs but are not directly under the control of the implementing agency. For example, the project is directly responsible for authorizing loans to small businesses, but the businessmen himself (or herself) will make the decision as to whether to hire more employees. Impacts can be positive or negative and can also be intended (planned in the project design) or unintended.

IMPACT EVALUATION

The use of an experimental or quasi-experimental design to estimate the net impact produced by the project on the affected population. This is the difference in income, height for weight ratio of a child etc., which exists between participants and the control group after the project has taken place after controlling for all other socio-economic characteristics of the two groups.

INPUTS — The resources allocated for the implementation of a project component. Money, materials and professional staff are some of the most common components.

MONITORING — Monitoring is an internal project activity concerned to assess whether project resources (inputs) are being used and administered as intended and whether they are producing the intended outputs. Monitoring can be divided into performance monitoring and process monitoring (see separate entries).

MULTI-METHOD APPROACH — The combination of two or more independent research techniques in the evaluation design so as to obtain consistency checks and to obtain a more complete understanding of the meaning of the results. It is recommended that quantitative and qualitative techniques should normally be combined in all evaluation designs.

OBJECTIVE — A precisely defined output or impact which a project intends to achieve. Objectives are the basis on which performance monitoring is conducted, and must be clearly quantified and have a timetable.

OUTPUT — A clearly definable direct product of a project component. For example, the number of serviced plots prepared for sale; the number of patients treated in a health center; the number of small business loans approved.

PARADIGM — A model or approach which is considered to be the ideal way to conduct a study or analysis. Often it is not possible to follow exactly the methods recommended by the paradigm but it serves as a reference point for evaluating the strengths and weaknesses of the actual design. The true experimental design is an example of a paradigm.

PERFORMANCE MONITORING — Assesses the extent to which inputs are being used in accordance with the approved budget and timetable and whether the intended outputs are being produced in a timely and cost-effective manner.

PROCESS MONITORING — Assesses the efficiency of the project implementation process.

PROGRAM — Long term urban development strategy which usually includes a number of separate projects

PROJECT — A set of urban development activities included in a government grant or authorization or in an agreement with an international development agency. A project will usually comprise a number of separate components (see separate entry).

QUALITATIVE EVALUATION — Approaches which seek to understand the processes through which a project is implemented and the ways in which it is perceived by and affects the intended beneficiaries. Many advocates of these approaches question the feasibility of obtaining reliable and meaningful quantitative measures of project impact.

QUANTITATIVE EVALUATION — Approaches to evaluation which seek to produce quantitative estimates of net project impact or cost-effectiveness. Many of these approaches use an experimental or quasi-experimental design.

QUASI-EXPERIMENTAL DESIGN — In most urban development projects it is not possible to use the true experimental design as it is not possible to randomly assign subjects to experimental and control

groups, or to control the environment to ensure that other factors do not interfere. A number of quasi-experimental designs have been developed to approximate as closely as possible the true experimental design. Great care must be used in interpreting the results of these designs as there are a number of threats to validity which can make it appear that the project has produced the impacts when in fact they were caused by external factors unrelated to the project (see Threats to Validity).

RANDOM SAMPLE

A method for selecting a sample of persons, families or groups which ensures that all members (persons, families etc) in the population being studied have an equal chance of being selected. If a stratified random sample is used, members of different strata will have different selection probabilities but each member of a particular stratum will have the same chance of selection.

SUB-COMPONENT

A clearly defined activity which forms part of a component. Technical assistance to families on house construction might be a sub-component of a sites and services component.

THREATS TO VALIDITY

There are 4 sets of factors which can mean that an apparent project impact (or lack of impact) is caused by a factor unrelated to the project. These are (a) Statistical conclusion validity (b) internal validity (c) construct validity of causal relations and (d) external validity. These factors are described in Chapter 4 Section C.

TRIANGULATION

The use of two or more independent methods to estimate a particular impact or value. If the estimates are consistent, the researcher can have greater confidence in the validity of the estimates. If the results are inconsistent, it is necessary to investigate further to determine the reason for the discrepancies.

REFERENCES CITED IN THE TEXT

(Many of the cited studies are unpublished or difficult to obtain. Questions concerning access to these documents should be directed to the authors of this Handbook.)

ABT, Clark (1976): "Evaluation of Social Programs." Sage, Beverly Hills, California.

ACHEN, Christopher (1982): "Interpreting and Using Regression." Quantitative Applications in the Social Sciences No. 29, Sage, Beverly Hills, California.

BAMBERGER, Michael (1979): "Problems of Sample Attrition in the Evaluation of Social and Economic Change: A Case Study from El Salvador." Paper presented to the Evaluation Research Society session on "Evaluation in Developing Countries".

BAMBERGER, Michael (1980): "Shelter Programs for the Urban Poor - A Comparative Review". Urban Edge, December 1980.

BAMBERGER, Michael (1981): "Quasi-Experimentation in an Urban Context: A Review of Experiences in the Evaluation of World Bank Shelter Programs." Urban and Regional Report No. 81-5, Urban and Regional Economics Division, The World Bank, 1981.

BAMBERGER, Michael, SANYAL, Bishwapura and VALVERDE, Nelson (1982): "Evaluation of Sites and Services Projects: The Experience from Lusaka, Zambia." World Bank Staff Working Paper No. 548, Washington, D.C.

BAMBERGER, Michael, GONZALEZ-POLIO, Edgardo and SAE-HAU, Umnuay (1982): "Evaluation of Sites and Services Projects: The Experience from El Salvador." World Bank Staff Working Paper No. 549, Washington, D.C.

BAMBERGER, Michael (1984): "A Methodology for Impact Evaluation in Urban Shelter Programs," Water Supply and Urban Development Department Discussion Paper No. 59. The World Bank, Washington, D.C.

BAMBERGER, Michael and OTTERBEIN, Julie (1982): "Designing a Questionnaire for Longitudinal Impact Studies," Urban and Regional Report No. 82-5, Urban and Regional Economics Division. The World Bank, Washington, D.C.

BAMBERGER, Michael (1982): "Statistical Evaluation of Project Impact Through Longitudinal Surveys," Urban and Regional Report No. 82-8, Urban and Regional Economics Division, The World Bank, Washington, D.C.

BLALOCK, Hubert (1979): "Social Statistics." McGraw Hill, New York.

BLALOCK, Hubert and Ann (1968): "Methodology in Social Research." McGraw Hill, New York.

BOGDAN, R. and TAYLOR, S. (1975): "Introduction to qualitative research methods." John Wiley, New York.

BOUGON, Michael (1983): "Uncovering Cognitive Maps: The Self Q" in Morgan, Gareth (1983) Editor Beyond Method: Strategies for Social Research. Sage Publications, Beverley Hills, California.

BUREAU D'EVALUATION (1976): "Second Housing Survey - The Control Group from Gran Dakar." Office D'Habitation de Loyer Modere. Dakar, Senegal.

BUREAU D'EVALUATION (1977): "The target population of the sites and services project. The effect of the selection process on some social and economic variables," Report III-6, Office D'Habitation de Loyer Modere. Dakar, Senegal.

BUREAU D'EVALUATION (1978): "Study of purchasers of serviced lots who have not begun construction of their houses." Series 3, No. 3, Office D'Habitation de Loyer Modere. Dakar, Senegal.

BUREAU D'EVALUATION (1979): "Study of house construction in the sites and services project." Series 4, No. 2, Office D'Habitation de Loyer Modere. Dakar, Senegal.

BURNS, L. and GREBLER, L. (1977): "The Housing of Nations: Analysis and Policy in a Comparative Framework." Wiley.

CAMPBELL, D. and STANLEY, J. (1966): "Experimental and Quasi-Experimental Designs for Research." Rand McNally.

CASLEY, D. and LURY, D. (1982): "Monitoring and Evaluation of Agricultural and Rural Development Projects." Johns Hopkins.

CHAMBERS, R. (1979): "Shortcut methods in information gathering for rural development projects." (Mimeo).

CONNER, R. (1978): "Selecting a control group: An analysis of the randomization process in 12 social reform projects". Evaluation Studies Annual Review, No. 3 (pp. 104-153), 1978.

COOK, T. and CAMPBELL, D. (1979): "Quasi-experimentation: Design and Analysis Issues for Field Setting." Rand McNally.

COOPER, Lauren (1981): "Designing the Sites and Services Plot Allocation Process: Lessons from Project Experience." World Bank Technical Paper No. 3, Washington, D.C.

EVALUATION STUDIES ANNUAL REVIEW (1979): Sage, Beverley Hills, California.

FAIRWEATHER, G. and TORNATSKY, L. (1977): "Experimental methods for Social Policy Research." Rand McNally.

FERNANDEZ-PALACIOS, Marisa and BAMBERGER, Michael (1984): "An economic evaluation of low-cost housing options in El Salvador." Water Supply and Urban Development Department Discussion Paper No. 55, The World Bank, Washington, D.C.

FREEMAN, Howard; ROSSI, Peter; and WRIGHT, Sonia (1979): "Evaluating Social Projects in Developing Countries." Organization for Economic Cooperation and Development, Paris.

FUNDACION SALVADORENA DE DESARROLLO Y VIVIENDA MINIMA (1977): "Analysis of the Development Process of the El Pepeto Project." San Salvador, El Salvador.

FUNDACION SALVADORENA DE DESARROLLO Y VIVIENDA MINIMA (1979): "Analysis of progressive development and self-help housing." San Salvador.

HAM-BIRF (1983): "Design for an evaluation of an artesan credit program (draft). Honorable Alcaldia Municipal. La Paz, Bolivia.

HATRY, H.; WINNIE, R. and FISK, D. (1981): "Practical program evaluation for state and local governments." The Urban Institute.

HENDRICKS, Michael (1981): "Service Delivery Assessment: Qualitative Evaluations at the Cabinet Level," in N. Smith (Ed) Federal Efforts to Develop New Evaluation Methods. Jossey-Bass, San Francisco.

HONADLE, G. (1979): "Rapid reconnaissance approaches to organizational analysis for development administration." Presented at the Conference on Rapid Rural Appraisal held at the Institute for Development Studies. Sussex, December, 1979.

INTERNATIONAL DEVELOPMENT RESEARCH CENTRE (1982):
"Low income urban shelter projects: An Annotated Bibliography of research funded by IDRC-IBRD." Ottawa.

INSTITUTO SER DE INVESTIGATION (1981):
"Modelo e evaluacion de impacto del programa IPC Linea de Base," 2 volumes, Bogota.

INSTITUTO SER DE INVESTIGACION (1984):
"Diagnostico Evaluativo del Avance del Programa IPC." IFT-072, April, 1984, Bogota.

JIMENEZ, E. (1982): "The Value of Squatter Dwellings in Developing Countries." Economic Development and Cultural Change, Vol. 30, No. 4 (1982), pp. 739-52.

KAPLAN, M. (1973): "Urban Planning in the 1960's: A Design for Irrelevancy." Praeger, New York.

KAUFMANN, Daniel and LINDAUER, David (1980):
"Basic Needs, Interhousehold Transfers and the Extended Family." Urban and Regional Report No. 80-15, Urban and Regional Economics Division. The World Bank.

KAUFMANN, Daniel (1982): "Social Interactions as a Strategy of Survival Among the Urban Poor: A Theory and Some Evidence." Doctoral Dissertation. Harvard University.

KAUFMAN, Daniel and BAMBERGER, Michael (1984):
"Income Transfers and Urban Projects: Research Findings and Policy Issues." Water Supply and Urban Development Department Discussion Paper No. 56, The World Bank, Washington, D.C.

KEARE, D. and PARRIS, S. (1982):
"Evaluation of shelter programs for the poor: Principal findings." Staff Working Paper No. 547, The World Bank.

KEARE, D. and JIMENEZ, E. (1983):
"Progressive development and affordability in the design of urban shelter projects." Staff Working Paper No. 547.

KISH, Leslie (1965): "Survey Sampling." John Wiley, New York.

LEVIN, Henry (1984): "Cost Effectiveness Analysis." Sage, Beverly Hills, California.

LEWIS-BECK, Michael (1980): "Applied Regression: An Introduction." Quantitative Applications in the Social Sciences, No. 22, Sage, Beverly Hills, California.

LEWIS, Oscar (1961): "The Children of Sanchez." Random House, New York.

LIEBOW, Elliot (1967): "Tally's Corner." Little Brown.

LINDAUER, David (1979): "Sources of Income, Selection Strategies and Project Affordability." Urban and Regional Economics Division. The World Bank.

MALPEZZI, Stephen (1984): "Analyzing an Urban Housing Survey: Economic Models and Statistical Techniques." Water Supply and Urban Development Department Discussio Paper No. 52, The World Bank, Washington, D.C.

MALPEZZI, Stephen; BAMBERGER, Michael; and MAYO, Stephen (1982): "Planning an Urban Housing Survey: Key Issues for Researchers and Program Managers in Developing Countries." Water Supply and Urban Development Department Discussion Paper No. 44, The World Bank, Washington, D.C.

MAYNARD, R. and MURNAME, R. (1981): "The Effects of Negative Income Tax on School Performance: Results of an Experiment." Evaluation Studies Annual Review, Vol No. 6.

MAYO, Stephen (1983): "Housing demand in developing countries." Water Supply and Urban Development Department, The World Bank.

McCALL, George and SIMMONS, J.L.: "Issue in Participant Observation: A Text and Reader." Addison-Wesley. Reading, Massachusetts.

McLEARY, R. et al (1979): "How a regression artefact can make any delinquency interventional program look effective." Evaluation Studies Annual Review, No. 4 (pp. 626-653).

MORGAN, Gareth (1983) Editor: "Beyond Method: Strategies for Social Research." Sage Publications, Beverley Hills, California.

NACHMIAS, David and Chava (1981): "Research Methods in the Social Sciences." St. Martins Press, New York.

NATHAN, R. et al (1981): "Public Service Employment: A Field Evaluation." Brookings Institute, Washington, D.C.

NIE et al (1975): "Statistical Package for the Social Sciences." (Second Edition). McGraw Hill, New York.

NUCLEO DE ACOMPANHAMENTO E AVALIACAO (Natal) (1983): "Desenvolvimento da Pesca Artesanal." Unidade de Administracao do Projeto. World Bank Medium Cities Project, Natal, Brazil.

OPERATIONS RESEARCH GROUP BARODA: "Monitoring and Evaluation Design for the First Madras Urban Project." Madras Metropolitan Development Authority.

PASTEUR, David (1979): "The management of squatter upgrading." Saxon House, England.

PATTON, Michael (1980): "Qualitative Evaluation." Sage, Beverly Hills, California.

PEATTIE, Lisa (1969): "The View from the Barrio." Ann Arbor, University of Michigan Press.

PEATTIE, Lisa (1981): "Marginal Settlements in Developing Countries." Annual Review of Sociology. 157-175.

QUIGLEY, John (1980): "The Distributional Consequences of Stylized Housing Projects: Theory and Empirical Analysis." Urban and Regional Report No. 80-18, Urban and Regional Economics Department, The World Bank, Washington, D.C.

REFORMA, Mila and OBUSAN, Ricci (1981): "Household networks and survival strategies among the urban poor: Monetary and non-monetary transfers among selected families in Tondo." Urban and Regional Report No. 81-22, Urban and Regional Economics Division, The World Bank, Washington, D.C.

REICHARDT, Charles and COOK, Thomas (1979): "Beyond Quantitative Versus Qualitative Methods." in Cook and Reichardt Qualitative and Quantitative Mehtods in Evaluation Research. Sage, Beverly Hill, California.

REICHARDT, R. (1979): "The statistical analysis of data from non-equivalent group designs" in Cook. T. and Campbell, D., Quasi-experimentation: Design and analysis for field settings. Rand McNally.

RESEARCH AND ANALYSIS DIVISION (1978):
"A Study of Income and Expenditure Patterns of Housholds in the Tondo Foreshore Arts." National Housing Authority, Manila, The Philippines.

RESEARCH AND ANALYSIS DIVISION (1979):
"A Study of the community participation process in Tondo." Philippine National Housing Authority. Manila.

RESEARCH AND ANALYSIS DIVISION (1979):
"A study of the impact of the project on the Physical Environment of Tondo." Philippine National Housing Authority. Manila.

RESEARCH AND ANALYSIS DIVISION (1981):
"The reblocking process in Tondo Foreshore." Philippine National Housing Authority. Manila.

ROGERS, Everett (1962): "Diffusion of Innovations." Free Press.

ROSSI, P. and WILLIAMS, W. (1972):
"Evaluating Social Programs." Academic Press. New York.

ROSSI, Peter and WRIGHT, James (1984):
"Evaluation Research: An Assessment." Annual Review of Sociology, Vol 10: 331-352.

SALMEN, Lawrence (1983): "Participant observer evaluation of urban projects in La Paz, Bolivia and Guayaquil, Ecuador." Water Supply and Urban Development Department Discussion Paper No.36, World Bank, Washington, D.C.

SIEGEL, Sidney (1956): "Non-Parametric Statistics for the Behavioral Sciences." McGraw Hill, New York.

SQUIRE, Lyn and VAN DER TAK, Herman (1975):
"The Economic Analysis of Projects." The Johns Hopkins University Press.

STRASSMAN, P. (1982): "The transformation of urban housing." Johns Hopkins.

STRUENING, Elmer and BREWER, Marilynn (1983):
"Handbook of Evaluation Research." Sage, Beverly Hills, California.

UNIDADE DE AVALIACAO (CAMPINA GRANDE) (1984):
"Estudo Rapido de Acompanhamento de Projeto de Apoio as Atividades Produtivas." Prefeitura Municipal de Campina Grande, Paraiba, Brazil.

UNITED STATES AGENCY FOR INTERNATIONAL DEVELOPMENT (1982):
"Assisting Small Business in Francophone Africa - The Entente Fund African Enterprise Program." PN-AAL-002.

UNITED STATES GENERAL ACCOUNTING OFFICE (1984):
"An Evaluation of the 1981 AFDC Changes: Initial Analysis." April 2, 1984, Washington, D.C.

URBAN AND REGIONAL ECONOMICS DIVISION (1978):
"Analysis of the level of demand for low-cost housing in Usulutan (El Salvador)." Urban and Regional Report No. ME-2, The World Bank, Washington, D.C.

URBAN AND REGIONAL ECONOMICS DIVISION (1978):
"The effectiveness of the cooperative as a means of social and economic change." Urban and Regional Report No. ME-5, The World Bank, Washington,D.C.

URBAN AND REGIONAL ECONOMICS DIVISION (1980):
"A Study of Income and Expenditure Patterns of Households in the Tondo Foreshore Area." Urban and Regional Economics Division, The World Bank, Washington, D.C. (Summary of a Report prepared by the Research and Analysis Division. National Housing Authority. Philippines.)

VALADEZ, Joseph (1982):
"Non-survey techniques in the evaluation of urban shelter programs." Urban and Regional Economics Division, The World Bank, Washington, D.C.

VAN MAANEN, John (1983) Editor:
"Qualitative Methodology." Sage Publications. Beverley Hills, California.

WEBB, E. et al (1965):
"Unobtrusive measures: Non-reactive research in the social sciences." Rand McNally.

WEISS, C. (1972):
"Evaluation Research." Prentice Hall.

WEISS, C. (1972) Editor:
"Evaluating action programs: Readings in Social Action and Education." Allyn and Bacon.

WHOLEY, J. (1979):
"Evaluation: Promise and Performance." The Urban Institute.

YAYHA, Saad (1982): "House Registration Handbook: A Model for Registering Houses and Plots in Unplanned Settlements." World Bank Technical Paper No. 4.