PD-AAA-906
41560

EVALUATION OF THE PROVINCIAL WATER PROJECT IN THE PHILIPPINES

BASELINE SURVEY RESULTS

PREPARED BY:

RESEARCH AND EVALUATION DIVISION
LOCAL WATER UTILITIES ADMINISTRATION
QUEZON CITY, PHILIPPINES

INTERNATIONAL STATISTICAL PROGRAMS CENTER
UNITED STATES BUREAU OF THE CENSUS
WASHINGTON, D.C.

ISSUED JUNE 1982

TABLE OF CONTENTS

		Page
	Acknowledgements	v
I.	Executive Summary	1
II.	Evaluation Plan	5
	A. Background Information Concerning the Program and the Evaluation	5
	B. Purpose of the Evaluation Study	6
	C. Study Design	7
	1. Research Design	7
	2. Selection of the Cities To Be Studied	10
	3. Timing of the Evaluation Study	11
	4. Variables Included in the Study	12
	5. Analytical Framework	13
	a. Plans for Analysis	13
	b. Methodological Issues	18
III.	Specific Objectives of the Baseline Study	25
IV.	Research Methodology	27
	A. Research Instruments	27
	B. Sample Design and Selection	27
	C. Data Processing	29
	D. Analysis	29
V.	Baseline Survey Results in the Study Areas	31
	A. Population Distribution in the Study Areas	31
	B. Socioeconomic Status of the Population in the Study Areas	32
	1. Income	34
	2. Education Level of Household Head	35
	3. Number of Household Items Owned	36
	4. Classification of House Type	37
	5. Ownership of House and Lot	39
	6. Source of Lighting	41
	C. Water Sources in the Study Areas	42
	D. Households' Opinions on Their Main Water Source	45
	E. Water Quality	46

Page

F. Water Availability and Consumption 49
G. Cost of Water 55
H. Health and Nutrition 56

1. Bivariate Analysis - Diarrhea 59
2. Bivariate Analysis - Nutrition 63
3. Multivariate Analysis - Diarrhea 66
4. Multivariate Analysis - Nutrition 69

I. Sanitary Practices 76
J. Economic Uses of Water 82

VI. Reliability of the Data 85

Appendix - Characteristics With Nonresponse Rates Above 10 Percent . 87

Figures

1 Conceptual Linkages Among Variables for the Provincial Water Project Evaluation 15
2 Connection Status at the Baseline and Follow-up Surveys By Experimental Stratum 19
3 Optional Approaches for Selection of Cohorts of Children for Assessment of Health and Nutritional Impact of Project . . 21
4 Average Annual Household Income 34
5 Household Heads With College Degrees 36
6 Number of Household Items Owned 37
7 Ownership of House 39
8 Ownership of Lot 40
9 Presence of Electric Lights 41
10 Households Using More Than One Water Source 44
11 Households With Unsafe Water 46
12 Water Quality By Boiling of Water 48
13 Households Storing Water 50
14 Households Fetching Water 51
15 Average Distance From House to Main Water Source. 52
16 Households Buying Water 53
17 Children Aged 0-4 With Diarrhea 56
18 Children Aged 0-4 With Diarrhea By Age 59
19 Households With Toilets Inside the House. 80
20 Average Distance From House to Toilet 81
21 Households Engaged in Water-related Businesses 82

Tables | Page

1 Number of Households and Median Household Size By City and Experimental Stratum, 1977-78 32
2 Socioeconomic Profile By City and Experimental Stratum, 1977-78 . 33
3 Percent of Households: Household Income By Stratum, 1977-78 . . 35
4 Percent of Households: Classification of House By City and Experimental Stratum, 1977-78 38
5 Number of Service Connections, 1977-78 42
6 Percent of Households: Main Water Source By City and Experimental Stratum, 1977-78 43
7 Percent of Households: Level of Satisfaction With Service of Their Present Water Source By City and Experimental Stratum, 1977-78 45
8 Percent of Households By Water Quality Results and By Main Water Source, 1977-78 47
9 Water Availability and Consumption By City and Experimental Stratum, 1977-78 49
10 Average Number of Gallons of Water Bought Per Household Member, 1977-78 54
11 Average Centavos Paid Per Gallon of Water By City and Experimental Stratum, 1977-78 55
12 Percent of Children 0-4: Nutritional Level By City and Experimental Stratum, 1977-78 57
13 All Children: Diarrhea in the Past 24 Hours By Method of Excreta Disposal, 1977-78 60
14 All Children 0-4: Diarrhea in the Past 24 Hours By Nutritional Status, 1977-78 60
15 All Children 0-4: Diarrhea in the Past 24 Hours By Whether Bathroom Is Fully Enclosed Or Not, 1977-78 61
16 All Children 0-4: Diarrhea In the Past 24 Hours By Distance Between House and Toilet, 1977-78 61
17 All Children 0-4: Diarrhea In the Past 24 Hours By Frequency of Washing Childrens' Hands After Playing Outside, 1977-78 62
18 All Children 0-4: Diarrhea in the Past 24 Hours By Frequency of Washing Hands After Using the Toilet, 1977-78 62
19 All Children 0-4: Diarrhea In the Past 24 Hours By Type of Garbage Disposal, 1977-78 63
20 Percent of All Children 0-4: Nutritional Distribution By Age of Child, 1977-78 64
21 Percent of All Children 0-4: Nutritional Distribution By Income, 1977-78 64
22 Percent of All Children 0-4: Nutritional Distribution By Current Breastfeeding Status, 1977-78 65
23 Percent of All Children 0-4: Nutrition By Method of Human Waste Disposal, 1977-78 65
24 Percent of All Children 0-4: Nutritional Distribution By Age of Child and City, 1977-78 70

Tables | Page

25 Beta Parameters of the Logistic Regression of Nutrition By City, Method of Excreta Disposal, and Income 73
26 Beta Parameters of the Logistic Regression of Nutrition On City, Method of Excreta Disposal, Income, and Type of Lighting . . . 75
27 Percent of Households By Frequency of Washing Hands After Playing Outside By City and Experimental Stratum, 1977-78. . . 76
28 Percent of Households: Frequency of Handwashing After Toilet Use, 1977-78 77
29 Percent of Households: Method of Excreta Disposal By City and Experimental Stratum, 1977-78 78
30 Percent of Households: Method of Garbage Disposal By City and Experimental Stratum, 1977-78 79
31 Percent of Households: Type of Bathroom By City and Experimental Stratum, 1977-78 79
32 Number of Households Engaged in and Average Income Earned in Previous Week (in Pesos) for Each Water-related Business: Type of Business By City 83
33 Percent of Households: Percent of Total Income Earned From Water-related Business By Experimental Stratum, 1977-78 . . . 84

Acknowledgements

This evaluation effort is supported by the Evaluation Division, Office of Development Planning, Bureau for Asia, U.S. Agency for International Development (USAID) and the Office of Capital Development, USAID/Manila, headed by William F. McDonald.

This evaluation project is directed by the Research and Evaluation Division of the Local Water Utilities Administration (LWUA) of the Philippines with technical assistance from the U.S. Bureau of the Census. The LWUA is headed by General Manager Carlos Leano. Maria Petronila Conti and Maria Theresa Galicia of that division were instrumental in developing and analyzing the data for this report under the general direction of Mr. Hermilo Balucan, Chief, Research and Evaluation Division.

This report was written by Steven C. Tourkin, International Statistical Programs Center, Bureau of the Census, under the general direction of Michael J. Hartz, Chief, Evaluative Studies Branch. Mr. Tourkin was assisted by Robert Fay, Robert Magnani (Bureau of the Census) and Ms. Conti and Ms. Galicia of LWUA. Additional support and direction was provided by William Stuart, Karen O'Conor and Rita Petroni (Bureau of the Census) and Hermilo Balucan (LWUA).

Questions regarding this report may be directed to Chief, Evaluative Studies Branch, International Statistical Programs Center, Bureau of the Census, Washington, D.C. 20233. Phone: (301) 763-1380.

I. EXECUTIVE SUMMARY

This is the first of three reports in the evaluation of the Philippine Provincial Water Project. There will also be a final report focusing on project impacts and an interim report focusing on shorter-term effects. The project will be evaluated by comparing data on various water and health variables collected from two study areas before and after project implementation. The evaluation was designed and implemented and the data are being tabulated and analyzed by a team composed of members of the Philippine Local Water Utilities Administration's Evaluation Division and staff of the U.S. Bureau of the Census, International Statistical Programs Center. All data collection and processing was done under the supervision of this team. Funding for this evaluation was provided by the Agency for International Development, Bureau for Asia, and administered by the Office of Development Planning, Evaluation Division. This report presents the results of the baseline (before) surveys of the Bacolod and Cagayan de Oro study areas. These two cities are believed to be representative of the five cities included in the project.

The evaluation was designed to address important questions concerning the effects of providing improved quality and availability of water to households in the two provincial towns. In particular, it will examine the actual changes, if any, in the quality and availability of water at the household level, the uses made of this water, the effects on health, nutrition and sanitary practices, and other impacts of the project. The evaluation will also investigate statistical relationships among many of the variables included in the study and the health and nutrition indicators in order to attempt to "explain" the effects of these variables on health and nutrition.

The Baseline Report provides a description of existing conditions in the two study cities before the project made a change in the water supply. The final report will compare these conditions to those existing after the project has been in place approximately 5 years, providing improved water service. Pre and post-project conditions in areas of the cities not to be served by the project (nonproject areas) will also be compared to allow the evaluation to isolate the changes which may have been caused by extraneous factors.

Major Findings of the Baseline Study

Prior to the project, 53 percent of Cagayan de Oro households, and 32 percent of Bacolod households were already connected to the existing city water systems, and for purposes of the study, comprise Experimental Stratum I. Average annual income for households already connected was P18,866 and P10,901 in Cagayan and Bacolod respectively. On a per capita basis, the figures are P2,774 ($383), and P1,703 ($234) annually. These current subscribers should receive improved quality and quantity of water when the project improves the existing systems.

The project will also extend the existing water systems into new areas of the cities. Households in these areas comprise Experimental Stratum II. Our surveys show that these areas score lower than the currently connected

households in many of the socioeconomic indicators we have measured including average annual household income (P15,570 vs. P18,866 in Cagayan de Oro and P9,900 vs. P10,900 in Bacolod) and percentage of households with electric lighting (57 percent vs. 74 percent and 47 percent vs. 75 percent) although other indicators such as education (lower in Bacolod, but not different in Cagayan) or quality of house construction (lower in Bacolod) show no consistent differences.

Similarly, the areas that will not be served by the project water system, Experimental Stratum III, scored lower in selected indicators, such as education of household head and percent of electrified households, than those currently connected or those to which service will be extended. The quality of house construction is significantly lower for the nonproject areas in Cagayan and household income is lower for the nonproject areas in Bacolod. Otherwise, the indicators show no consistent differences between nonproject and project area households.

It appears, therefore, that those households currently connected to the pre-project existing water systems (stratum I) have the highest standard of living and that those to which the project will extend water service (stratum II) are less well off. It also appears that those households in areas to which water service will not be extended (stratum III) by the project score somewhat lower in selected socioeconomic indicators than those that will be served, but there does not appear to be an overwhelming difference in the standard of living between those to be served and those not to be served. (See Tables 2-4 and Figures 4-6 for a presentation of the relevant survey results.)

Forty-one percent of all households in both cities utilize water from more than one source for household use. Even among stratum I households many use water from other sources for household uses.

Most of the households in both cities fetch water from some outside source. Even among stratum I households, over one-half fetch water. The distance traveled to fetch water is usually greater in stratum II areas although the average distance traveled in Cagayan even for stratum I households is 84 meters, compared to 96 for nonconnected households. In Bacolod, distances traveled are shorter -- 21 meters for stratum I households, 36 meters in stratum II areas and 51 meters in stratum III areas.

The baseline survey shows that approximately 10 percent of all households in the cities were engaged in "water-related" businesses. These ranged from water selling, food processing, and operation of restaurants to pig and poultry raising. Most of the households with businesses (53 percent - Cagayan, 81 percent - Bacolod) felt that existing quality of water supplied was adequate.

A major finding of the baseline study was that 82 percent of the water samples taken from the houses in both cities failed to meet the Philippine

national standards for safe drinking water. Even in the households served by the pre-project system, two-thirds or more of the samples failed to meet the standard.

Two major variables examined in this evaluation are health and nutritional levels. The primary health indicator in the survey was "percent of children aged 0-4 having diarrhea during the past 24 hours". The primary nutritional indicator was "nutritional status" as derived by relating weight to age for children 0-4 years of age. The survey results show that 18 percent and 9 percent of all children aged 0-4, in Cagayan and Bacolod respectively, had diarrhea during the 24 hours prior to the interview. According to the nutritional level standard, most children in the studied cities were malnourished. Only 2 out of 5 children were considered normal, and approximately 30 to 40 percent were first degree malnourished. The rates vary between cities and experimental strata.

The relationships between the health and nutrition indicators and various variables is explored thoroughly in Section V, Part H. The initial analysis showed that the incidence of diarrhea among children was affected by several variables, including age and nutritional status of the child, the household's method of excreta disposal, the distance from the house to the toilet, the frequency of washing children's hands after using the toilet or playing outside, the method of garbage disposal and, to a lesser extent, whether the bathroom was fully enclosed. The level of nutrition was affected by the age of the child, the income of the household, whether the child was being breastfed, and the method of excreta disposal. Further analysis revealed that these associations all appear to be a consequence of a single underlying relationship between the method of excreta disposal and the health indicators. That is, children living in households with automatic flush toilets were less likely to have diarrhea or be malnourished than children using other toilets.

It is interesting to note that there was no relationship between the households' water quality indicators and the health indicators. This could be due, in part, to the small number of households that were observed to have "safe" water. In order to properly test this relationship, a larger sample of households with safe water may be necessary. Also, water quality at the household level may vary from day to day, and a measurement at the household at one point in time may or may not be representative of the "usual" quality. This relationship will be investigated again after a follow-up survey is completed.

In conclusion, there can be no doubt that both of these cities need improved water service. Most of the water samples examined were determined to be unsafe for drinking. Many of the households reported that they fetch and/or store water, indicating that the quantity or reliability of water from available sources might not have been sufficient. The baseline data also indicate that the project will extend water service into areas that have a somewhat lower standard of living than those served by the system existing before the project. The data also imply that the health benefits from the improved water systems can be increased if more households use automatic flush toilets.

The post-project surveys are now being designed and are scheduled to be carried out in late 1982 and early 1983. Conclusions on the actual measurable effects of the project on the inhabitants of the study areas will be included in the final report to be issued after this data is analyzed and compared to baseline measures. The interim report will analyze results of smaller-scale surveys conducted between the baseline and follow-up surveys.

The data on which this summary is based are contained in Section V of this report. The information in this Baseline Report and the findings that will be included in the final report are based on studies using scientifically sound design and statistical principles. The findings come from data collected in surveys of households in the project cities. These surveys were done using a stratified random sample of all households in the areas. The data collected were systematically edited and electronically processed and tabulated and were analyzed scientifically using methods outlined in Section VI of this report.

5

II. EVALUATION PLAN

A. Background Information Concerning the Project and the Evaluation

The Provincial Water Project is jointly sponsored by the Government of the Philippines and the United States Agency for International Development (USAID). The project was originally formulated with two goals: (1) the development of water systems in selected urban areas which would provide safe water on a reliable and economic basis to a significant proportion of the population, and (2) the development of the technical and institutional capability to extend such service to other cities in the Philippines.[1] These goals would be attained by replacing old, unreliable water systems with modern systems that would include new wells, treatment plants, storage reservoirs, pumping and distribution systems and administrative and support facilities. The urban areas selected for the project were Cagayan de Oro (Mindinao), Davao City (Mindinao), San Pablo (Luzon), Tacloban City (Leyete) and Bacolod City (Negros Occidental). Originally, Baguio (Luzon) was also included but was deleted later.

The project implementing agency is the Local Water Utilities Administration (LWUA). The Loan Agreement with USAID was signed on May 23, 1974. By the beginning of 1979, the new water systems in the five selected cities were functioning.

The immediate beneficiaries of the project were the residents of the urban communities concerned. In 1978, the existing water systems served about 30 percent of the households. The USAID predicted that approximately half of the households would be served by 1982 with water from the new systems, which are expected to provide safer water (less contaminants) and more reliable service (less service interruptions) than the older systems.

One of the conditions precedent to the USAID loan was that an evaluation be conducted to measure performance levels and the impact of the project on health and economic development. At the request of AID/Washington, a 5-year evaluation study of three of the five cities[2] was developed. The evaluation plan required data from household surveys conducted at three points in time in both project and nonproject areas to measure performance and environmental sanitation effects as well as ultimate health, economic and social impacts of the project.

In 1976, the LWUA contracted the Institute for Philippine Culture (IPC) to conduct the evaluation in two of the cities -- Bacolod and Cagayan de Oro. The IPC was responsible for designing the study and managing

---

[1] USAID, Capital Assistance Paper, Philippines Provincial Water, AID.DLC/P-2020, March 13, 1974.

[2] Silver, Keeve and Warner, Proposals for an Impact Study of the Philippines Provincial Cities Water Supply Project, January 17, 1975.

all evaluation related activities. The Census Bureau was requested by the USAID to provide assistance to the IPC. In June 1978, a research and evaluation unit was organized in the LWUA to participate more actively in the evaluation. Soon after, the Census Bureau's assistance was intensified in order to develop the LWUA's in-house evaluation capabilities. In July 1979, the LWUA terminated their contract with the IPC and the LWUA evaluation unit assumed responsibility for managing the study.

B. Purpose of the Evaluation Study

The evaluation of this project was originally requested by AID for the following purposes:

1. to provide information on the health impacts of improvements in water systems;
2. to assess the performance of the LWUA in constructing and maintaining the water systems; and
3. to establish a framework for the evaluation of future water projects.

One of the main concerns of the evaluation team in 1975 was that a fundamentally sound approach to impact evaluation of water projects be initiated. It was recommended that the basic study design be expandable to accommodate larger, more comprehensive studies, reducible to some smaller form for special purpose studies, and applicable to projects in other developing countries.

The major questions that this study was designed to investigate include:

1. How are the water systems performing?
2. Who is using the water from the new systems?
3. What are the short-term benefits to the users of the new systems?
4. What are the long-term health, economic and social benefits to the users of the water systems?
5. Are there causal relationships between the improved water systems, water use and the observed short and long-term benefits?
6. What are the preferred health and socioeconomic measurement techniques for this type of evaluative study?

C. Study Design

1. Research Design

In order to obtain an accurate measurement of health and economic benefits from the improved water systems, the study follows a quasi-experimental design using three study groups. Two of these groups are experimental groups in areas served by the project and the third is a control group not in the project area.

Experimental stratum I is the group of households using piped water from the existing city water system as their major source.

Experimental stratum II is the group of households in the project area whose major water source is not the existing city water system.

Experimental stratum III is a similar group of households in non-project areas.

Data will be provided by ten household surveys, conducted over a 4-year period. These will include a baseline survey of the study areas (approximately 2,500 interviews), eight rounds of smaller-scale quarterly interim monitoring surveys (approximately 250 interviews each round), and a follow-up survey (approximately 2,000 interviews).

The main purpose of the baseline survey is to obtain pre-project measurements of the health and economic status in the study cities, Bacolod and Cagayan de Oro. An additional purpose of the baseline study was to establish the similarity or dissimilarity of the various experimental groups before project intervention with respect to key variables. Key variables include health and economic variables as well as those factors which might affect those variables or affect or interact with the project intervention. The objectives of the baseline analysis are, specifically:

a. to compare households in the three experimental strata with respect to key variables,
b. to describe the social and economic characteristics of the population in all three strata.

The results of this analysis will pinpoint variables which should be considered and may have to be controlled in analyzing trend and before/after comparison data.

The purposes of the eight interim monitoring surveys are:

a. to provide data for a time-series analysis of health effects on the population,
b. to examine seasonal variations in health indicators,
c. to provide data (combining the eight surveys) that will be compared to the baseline and follow-up survey data to measure

project impacts and test causal relationships between the water system and the observed impacts, and

d. to determine when the follow-up survey should be conducted.

The follow-up survey will obtain data that will be compared to the baseline and interim monitoring data to measure changes in the key variables over the 4-year period.

The research design is illustrated in the following diagram where X's represent interventions and O's represent measurements.

Bacolod	Baseline Survey		Interim Monitoring Survey		Follow-up Survey
Stratum 1	O_1	X_1	O_2	O_3 .....O_9	O_{10}
Stratum 2	O_1	X_2	O_2 X_2	O_3 ..X_2..O_9	O_{10}
Stratum 3	O_1		O_2	O_3 .....O_9	O_{10}

Cagayan de Oro					
Stratum 1	O_1	X_1	O_2	O_3 .....O_9	O_{10}
Stratum 2	O_1	X_2	O_2 X_2	O_3 ..X_2..O_9	O_{10}
Stratum 3	O_1		O_2	O_3 .....O_9	O_{10}

O_1 represents the baseline survey conducted before the project implementation, O_2 to O_9 represent the eight quarterly monitoring surveys, and O_{10} represents the follow-up.

X_1 represents the intervention of improved water service to households previously connected to the old city water system, while X_2 represents the accessibility to the new/improved water system for those households in the project areas whose major water source was not part of the old system. The X_2 intervention will occur over time as more households choose to connect to the system. Thus, it is also shown after O_2 and O_3, O_4, etc.

The design lacks the equal number of observations both before and after project intervention that is characteristic of time-series designs. However, the use of a comparison group allows comparison of trends in effect and impact indicators over a 4-year period after the treatments are introduced, in addition to comparisons between the baseline and subsequent observations.

2. Selection of the Cities to be Studied[3]

There are USAID-assisted comprehensive water projects[4] in Bacolod, Cagayan de Oro, Davao, San Pablo and Tacloban Cities. From these, Bacolod and Cagayan de Oro Cities were selected as study areas. The existence of a significant low-income sector was the main criterion in the selection of study areas for this evaluation study. A substantial portion of the population in both Bacolod and Cagayan de Oro Cities fall into the low-income bracket clustering in slum and squatter areas throughout the poblacion and/or areas that were targeted for the improved water systems. Both cities have agriculture and fishing as basic economies.

Davao City has similar characteristics, but was not considered due to its large and rapidly increasing population of approximately 512,000. In addition, relocation of squatter colonies in Davao had begun prior to the implementation of the evaluation study.

Tacloban City was also eliminated as a possible study area because there were plans for the city to become the target of several government development programs, making it difficult to isolate the effects and impacts of improved water systems, as well as the fact that the city's water improvement project was scheduled to be completed later than in Bacolod or Cagayan de Oro. For purposes of this study it was considered necessary that the schedule of the water supply improvement occur simultaneously in the two study areas.

San Pablo was not considered because its water supply improvement began in January 1976 before the evaluation plan was finalized. Thus, a baseline survey over all cities affected was not possible.

---

[3] IPC Research Proposal, The Health and Economic Impact of Improved Water Systems on Provincial Cities, 1976-79, Ateneo de Manila University, Institute of Philippine Culture, Q.C., 1976.

[4] Comprehensive projects refer to the provision of a water supply system that will provide water needs for the next 10-20 years.

3. Timing of the Evaluation Study

The time schedule for data collection on this evaluation is shown below for Bacolod and Cagayan de Oro.

Activity	Bacolod	Cagayan de Oro
Baseline survey	Jan.-Mar. 1978	Oct.-Dec. 1977 May 1978
Interim Monitoring Surveys	(1) Dec. 1979	Dec. 1979
	(2) Mar. 1980	Mar. 1980
	(3) June 1980	June 1980
	(4) Sept. 1980	Sept. 1980
	(5) Dec. 1980	Dec. 1980
	(6) Mar. 1981	Mar. 1981
	(7) June 1981	June 1981
	(8) Sept. 1981	Sept. 1981
Follow-up Survey	Jan. 1983	Nov. 1982
Final Report	August 1983	August 1983

Each data collection effort is followed by a data processing phase. The results are then analyzed and reported.

There have been several delays in the implementation of this evaluation schedule, beginning with the baseline surveys. The contractor responsible for the design, data collection, data processing and analysis of the surveys had difficulty processing the baseline data after it was collected. The Census Bureau agreed to edit the data and produce tabulations after the fact, which proved to be time consuming.

Beginning in mid-1979 the LWUA assumed primary responsibility for conducting the evaluation and the Census Bureau became more directly involved as advisors on the study. Since the LWUA staff was inexperienced in all survey related activities, the timeliness of data processing and analysis was compromised in favor of the development of a staff within LWUA capable of performing these tasks.

4. Variables Included in the Study

The following is a list of key variables to be used as indicators in this study.

Health Impact Variables

Incidence of diarrhea among children
Nutritional level of children

Economic Impact Variables

Household income
Income from water related activities
Days absent from work
Medical expenses
Fire damages
Time spent fetching water

Characteristics of Water Supply

Water quality
Type of water source
If supply is enough

Factors Affecting Water Quality

Storage of water
Boiling of water
Treatment of water with chemicals
Filtering water

Factors Affecting Susceptibility to Disease

Breastfeeding of children
Nutritional status of children
Washing hands, dishes, fruit
Garbage disposal
Waste water disposal
Location of bathroom
Sharing of bathroom
Type of toilet
Location of toilet
Sharing of toilet
Disposal of toilet water

Consumption of Water

Amount of water fetched
Water bill (connected households)

Factors Affecting Consumption of Water

Cost of water
Location of water source
Distance to water source
Household income
Household size
Availability of water 24 hours per day

Characteristics of Household Members

Demographic variables
Level of living indicators

5. Analytical Framework

a. Plans for Analysis

In specifying an analytic framework for an evaluation as complex as the one at hand, it is important to begin with a clear understanding of the phenomena to be studied and the linkages among the variables on which information is to be collected and analyzed. A useful starting point in clarifying the nature of the problem would be to distinguish between the several types or classes of variables that will play a role in the project evaluation.

In experimental or quasi-experimental design situations, an initial distinction may be made between the "treatment" and the "response". The treatment consists of the "stimulus" (in the present case, the provision of a safe, reliable water supply) which is to be administered to one (or some) group(s) and not to the other(s). Accordingly, the impact or effect of the treatment is properly determined by means of pre- and post-treatment comparisons among the various experimental strata in the study population.

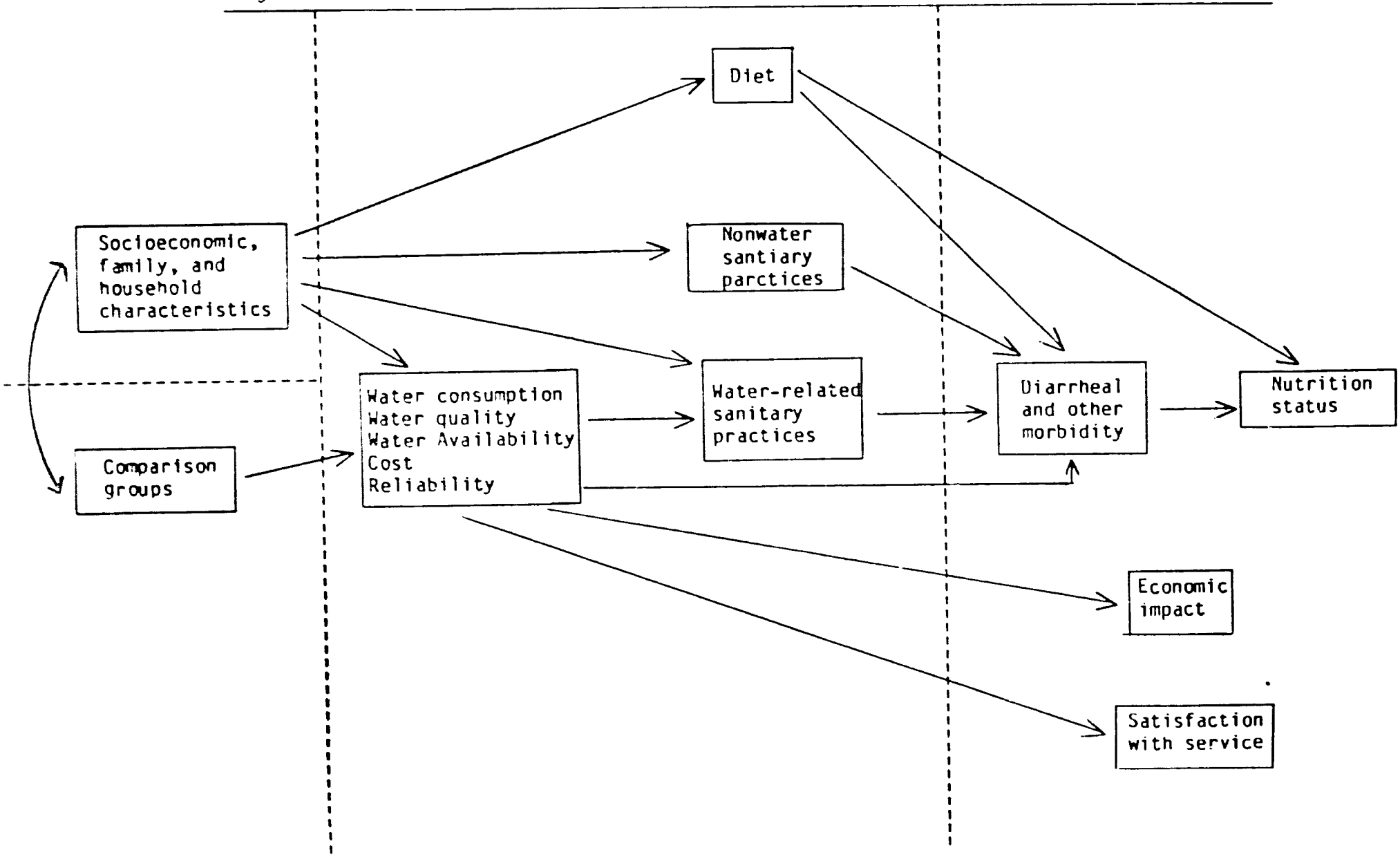
In this study, experimental strata I and II, both part of the water project area, can be affected by the new water system while stratum III cannot be (see definitions of the experimental strata on page 7). Thus, for analytic purposes, "membership" in stratum I or II is interpreted as representing the "treatment." The response, on the other hand, consists of those characteristics, attributes, rates, etc. which the treatment is intended to influence and which thus provide the basis for quantifying the impact or effect of the treatment or project intervention in the study population. In the present case, the response consists of changes in health status indicators (diarrheal and other morbidity rates, nutritional status), project-related economic indicators (water-related income), and various indicators of the efficiency/reliability/quality of the water systems.

Another class of variables, socioeconomic/family/household background characteristics, reflect the environment in which the project takes place and are a critical feature in quasi-experimental design studies in which: (1) the comparison groups are neither randomly assigned nor otherwise "matched" initially with respect to characterisitics which may affect or interact with the project intervention, and (2) the study population is not "controllable" during the course of the study from the standpoint of changing composition. Both conditions exist in the cities studied in this project. Such characteristics need be controlled statistically in the analysis of quasi-experimental design data in order to minimize the influences of initial differences in experimental group composition and changes in composition during the course of the study on observed differences in the response variables.

In the present evaluation, background characteristics such as family income, family size, number and spacing of children, household crowding, quality of houseing facilities, etc. are relevant factors which may determine (at least partially) the effects of the project on health and economic impact variables and on which the experimental groups may differ in terms of composition. In other words, the inclusion of relevant background characteristics permits measurement of the impact of the project intervention "net" of the influences of other factors.

The study also includes variables which indicate the specific mechanisms or behavior through which both the project intervention and background characteristics influence the response or impact variables. For example, the fact that a given household happens to be located in an area where piped water is available is not likely to affect any of the health impact variables unless the household members take advantage of the availability of the clean and reliable water supply in terms of such factors as increased consumption of clean water, improved sanitary practices, or related behavior. Similarly, higher household income may be translated into better health only through such factors as more adequate nutrition, better sanitary conditions, and the purchase of greater quantities of clean water. Thus, variables in this class are "intermediate" variables in the sense that they relate the treatment and background factors to the response or impact variables through specific characteristics or behavior. Such variables are extremely useful for analytic purposes in that they permit the analyst to pinpoint the specific factors responsible for the success of a project, or conversely, for its lack of success.

Figure 1. CONCEPTUAL LINKAGES AMONG VARIABLES FOR THE PROVINCIAL WATER PROJECT EVALUATION



The basic analytic framework and linkages among the important categories of variables for the Provincial Water Project Evaluation are indicated in Figure 1.[5] The types or classes of variables described above are separated in Figure 1 by the dashed lines. The variables to the far right constitute the primary response or impact variables of interest, the variables in the center section of the figure represent the several categories intermediate or specification variables, while the two categories of variables to the far left represent the relevant background characteristics and comparison groups. The double-headed curved line connecting these two classes of variables is indicative of the fact that they are to be treated as jointly exogenous for analytic purposes. The single-headed arrows connecting the other categories of variables indicate the anticipated paths of influence.

It should be noted that not all of the response variables which are relevant for one aspect of the project evaluation, namely the assessment of the performance of the LWUA in constructing and maintaining the water system, are included with the other response variables in Figure 1. This was done for the purpose of simplicity, since many of the response measures of performance for this aspect of the evaluation (e.g., water quality, availability, and cost) are intermediate or specification variables for the other aspects of the evaluation (i.e., health and economic impacts). It is to be understood, however, that such variables are to represent the response criteria for this aspect of the evaluation.

Figure 1 suggests something of the complexity of the problem at hand. It should be evident that, particularly with respect to the evaluation of project impacts on the health status indicators (diarrheal and other morbidity, nutrition status), some form of analysis which considers several variables simultaneously (i.e., multivariate analysis) will be required in order to isolate the effects of the project intervention.

In order to impose some order on the analysis and at the same time keep costs in line, a two-stage analytic plan is proposed. The first stage of analysis will be bivariate in nature and focus attention on the changes in the various intermediate and impact variables respectively that had occurred in the comparison groups during the study period. The second stage of analysis would be multivariate in nature and would be directed to quantifying the net effect of the project intervention and the decomposition of the total effect into components due to various intermediate variables.

---

[5]Figure 1 is a simplified version of the conceptual model presented in the Baseline report. It was modified for the purposes of making an analytic approach more explicit.

This two-stage strategy will be advantageous in several respects. First, the examination of the data on a bivariate basis initially would provide the users of the data with an opportunity to understand the dynamics of the project impact by being able to observe changes in each variable individually. Second, the initial examination of bivariate relationships is likely to lead to greater efficiency in the multivariate analysis to the extent that some of the intermediate variables can be eliminated as factors responsible for observed changes in the impact variables. Finally, it is entirely likely that the examination of bivariate relationships could lead to the investigation of more complex multivariate relationships among variables that had not been considered a priori.

The analysis that will be accomplished after the follow-up survey data are collected will be organized as follows:

(1) Assessment of Changes in Health and Economic Impact Variables

- Incidence of Diarrheal, (other morbidity)
- Nutritional Status (anthropometric measures)
- Economic Benefits (water-related income, medical expenses, work absenteeism)

(2) Assessment of Changes in Water-related Behavior

- User Status (source of water)
- Consumption (household and productive)
- Quality
- Cost (monetary and time)
- Water-related Sanitary Practices
- Availability (satisfaction)

(3) Socioeconomic Background Characteristics and Nonwater-related Behavior

- Income (per capita income)
- Family Size (number and spacing of children)
- Housing (number of rooms, persons per room, facilities)
- Breastfeeding Practices
- Nonwater-related Sanitary Practices

(4) Multivariate Analysis

In the preceding outline, the first three sections constitute the bivariate portion of the analysis. These are further divided as follows: the first section will examine changes in the respective impact variables that had occurred during the study period in the respective comparison groups (baseline data vs. follow-up data), the second will examine changes in water-related behavior, and the third will focus attention on group differences in socioeconomic background and non-water-related intermediate

variables among the comparison groups and any changes that had occurred during the course of the study (baseline data vs. follow-up data). The multivariate analyses will be performed on the follow-up data set and will be contained in the fourth section of the final report.[6] It is not advantageous to attempt to specify at this time the format of this stage of the analysis since this will be contingent upon the results of the bivariate analyses. As noted previously, however, the purpose of this portion of the analysis will be to measure the effect of the project intervention net of other factors and weight the various project relevant intermediate variables in terms of their contribution to the total effect or impact on the response variables.

b. Methodological Issues

There are several methodological issues which will play an important role in shaping the final form of the analysis. These are discussed below along with proposals for dealing with them.

(1) Treatment of Interim Monitoring Phase (IMP) Data

As indicated earlier, eight small-scale quarterly surveys were conducted during the period December 1979 to September 1981 in addition to the Baseline and Follow-up Surveys. While these data provide a basis for a time series study, it would be desireable for the proposed analysis to combine the data gathered from the IMP surveys into a single set of estimates with a time reference at the mid-point of the IMP (i.e., October 1980) for the purpose of measuring the impact of the project. The primary reason for this is that the sample sizes used for the IMP surveys (250 households in each of the cities per quarter) are likely to be far too small to permit firm conclusion as to the projects' impact if analyzed as eight independent surveys. A time series analysis of the IMP survey data would, however, be useful as a supplementary analytic exercise and is scheduled to be performed.

(2) Appropriateness of Comparison Groups

A more serious issue concerns the adequacy of the comparison groups (experimental strata) for the purpose of measuring the impact of the project. In particular, the analytic treatment of households in the experimental stratum II, namely households not utilizing city water systems sources prior to the project but which were located in areas which would be connected to the water system during the course of the project, needs to be assessed.

---

[6] A multivariate analysis using the baseline data is contained in this report.

Under a conventional quasiexperimental design, all households receiving the "treatment" would be classified as the experimental group, while those not receiving the treatment would constitute the control group. In the present case, it was necessary to further divide the treatment group into two subgroups, depending upon whether they were using the city water system prior to the project, since these are to receive somewhat different "treatments". A problem arises, however, in that while the project benefits are being administered on an "areal" basis, the second experimental stratum is likely not to be homogeneous with respect to the project intervention on a "household" basis. In other words, the fact that a given household is located in a project area (i.e., an area where water connections are to be made available) does not mean that all households in the area will avail themselves of the project benefits. Accordingly, if the project impacts are to be measured on an areal basis, that is, according to whether or not sample households are located in a project area, some of the sample households in that area are likely not to have been affected by the project. This would dilute, to some extent, the measurement of the magnitude of the provision of improved water. Thus, the second water service experimental stratum will consist of households that had and had not been exposed to the treatment. (See Figure 2.)

FIGURE 2. CONNECTION STATUS AT THE BASELINE AND FOLLOW-UP SURVEYS BY EXPERIMENTAL STRATUM

Experimental Stratum	Household Connection Status - Baseline Survey	Household Connection Status- Follow-up Survey
1	Connected	Connected
2	Not connected	Connected Not connected
3	Not connected	Not connected

For the purpose of analysis, it would seem essential that some attention be directed to differences among the two subgroups within the second experimental group distinguished above. One approach might be to proceed initially along the lines of a conventional quasi-experimental design analysis in which the relevant comparisons would be between the two water-service strata on the one hand and the control group on the other. A second set of comparisons might also be made between the two water service strata in order to assess the extent to which the impact of replacement of an existing, albeit inferior, water system differs from the provision of a new service. A third set of comparisons within the second water service stratum which assess differences between households which received city water connections during the study period and those which did not also seems essential.

One possible outcome of this type of analysis might be that, depending upon the proportional breakdown within the second water-service stratum with respect to system utilization and the magnitude of observed differences between the subgroups, the initial set of comparisons would need to be revised by excluding the 'nonusers'. This would allow a better assessment of the impact of the water itself. (Given this possibility, it might be expeditious to explore this issue as a preliminary step in the analysis.)

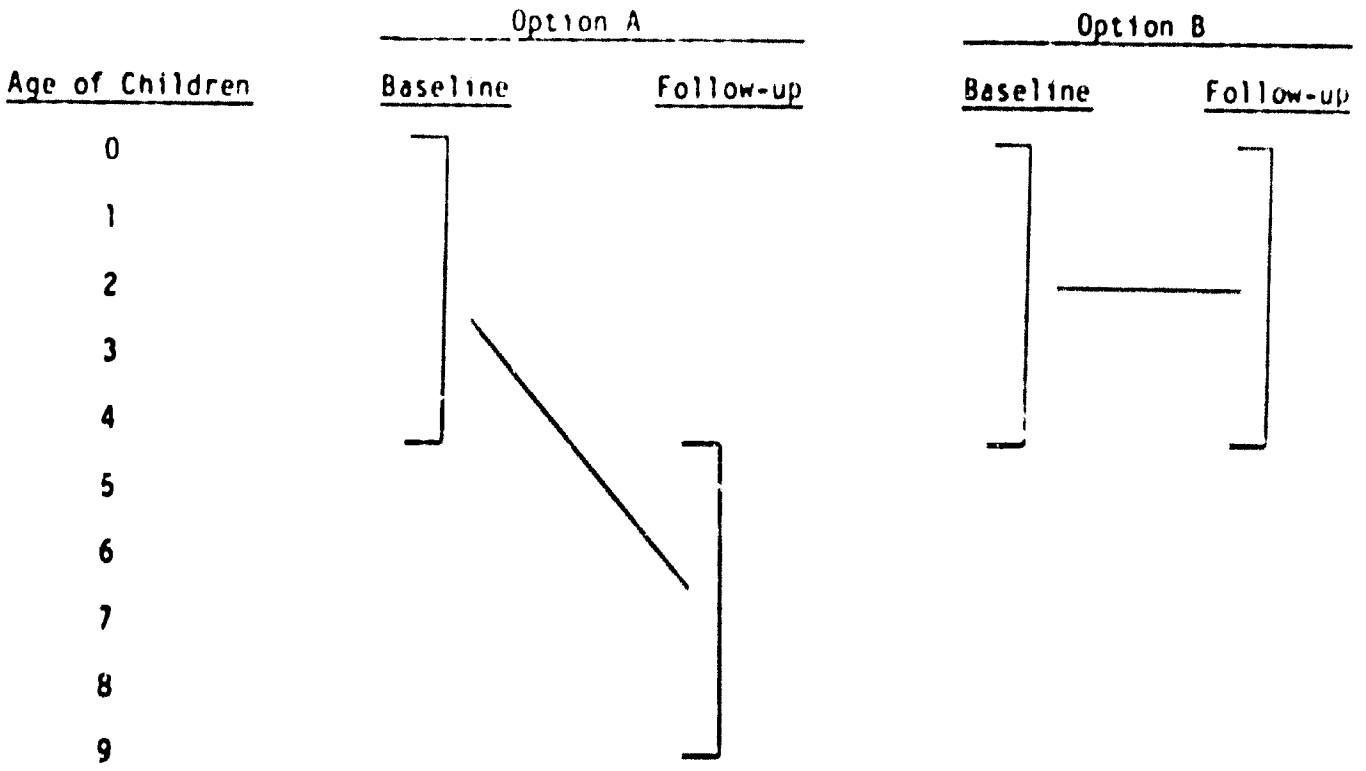
(3) Defining Appropriate Cohorts for the Assessment of Health Impacts

The primary health-status indicators to be employed in evaluating the impact of the project consist of measures of diarrheal morbidity and nutritional status among young children. Young children were chosen as the target population for this phase of the evaluation because they tend to be the population subgroup that is most susceptible to water-related maladies, and their status thus provides a sensitive indicator of the effects of the project. The impact of the project intervention is to be assessed through the comparison of health status indicators in the respective comparison groups at the beginning and the end of the study period.

An issue that arises concerns the selection of appropriate cohorts of children for conducting such before and after comparisons. Specifically, at issue is whether the comparisons should be made for the same cohort of children who are of a fixed age at each point in time. The nature of the problem is illustated in Figure 3.

Under Option A, the analysis would focus on the extent to which experimental stratum differences in the health status indicators at the time of the Baseline Survey, at which time the subpopulation of interest was children aged 0-4, had changed by the time of the Follow-up Survey, at which

FIGURE 3. OPTIONAL APPROACHES FOR SELECTION OF COHORTS OF CHILDREN FOR ASSESSMENT OF HEALTH AND NUTRITIONAL IMPACT OF PROJECT



time this cohort of children would be aged 5-9. The appropriate response variables would consist of such things as proportions of children experiencing diarrheal or other forms of morbidity in a reference period prior to the survey and proportions falling in various nutrition categories.

It should be noted, however, that some caution need be exercised in interpreting the results of such comparisons. In analyzing the data on nutritional status, it would be appropriate to infer a project effect directly from falling proportions of children in the various malnutrition categories in the project status since malnutrition is determined on the basis of anthopometric standards which are age-specific. (Strictly speaking, one would nevertheless want to make comparisons with the control group to determine the extent to which improved nutritional status might be due to factors other than the intervention.) However, falling proportions of children experiencing diarrheal morbidity should not be construed as direct evidence of a beneficial project impact since young children (ages 0-4) are considerably more prone to this form of morbidity than older children. In this case, the project impact must be assessed through comparisons among experimental groups.

The second option, Option B, would involve the comparison of morbidity rates and nutritional status for different cohorts at different points in time. The cohorts being compared will be different because the cohort aged 0-4 at the time of the Baseline Survey will be aged 5-9 at the time of the Follow-up Survey. Under Option B, comparisons would be made between the sample populations aged 0-4 in the various experimental strata at various points in time. The two options will thus yield information on the extent of health status improvement from what different perspectives.

(4) Informational Limitations

There appear to be three areas of substantive interest in the present evaluation in which the analytic plan described above will have to be modified. These modifications are necessitated by the fact that information was not collected on the Baseline Survey for one area of interest, and for two other cases, the information that was collected on the Baseline Survey is suspect.

The first case concerns the issue of nutritional status. While it is clear that water does have an impact on nutrition through nutritional malabsorption brought about by diarrheal and other water-related intestinal infections (see Figure 1), the research literature suggests that it would be useful in attempting to isolate the impact of the water on nutritional status to have at least some rudimentary

information on differences in diet. Accordingly, the collection of some basic information on food consumption is being considered for the Follow-up Survey. Such information will be available only for one point in time, thus precluding the strategy of monitoring changes in group differences. For this aspect of the analysis, a multivariate cross-sectional approach will have to be adopted in which the focus would be to isolate the effects of water-related factors from dietary and other factors. This will be in-included in the multivariate section of the report.

The second case concerns the morbidity (other than diarrheal) and water consumption data that were collected in the Baseline Survey. With respect to morbidity, there is reason to believe that the reports of incidence of "afflictions" were not limited to those types of illness which are typically included in such data, but also included accidental injuries which are beyond the potential impact of the project intervention. Thus, these data are of questionable value in evaluating the impact of the project.[7]

For water consumption, information is available only for those hosueholds having meters, and thus comparisons between experimental strata cannot be made. More appropriate data are to be collected on the Follow-up Survey, but as in the case of the data on diet, these will have to be analyzed within a multivariate, cross-sectional framework.

(5) Survey Design Implications (explained in Section IV)

The design of the Provincial Water Project Evaluation is a multistage, clustered sample design in which relevant information was collected for all residents of selected households. In order to obtain unbiased estimates of the impact of the project intervention, the analytic procedures adopted (particularly those employed in the multivariate analysis) must: (1) take into account the complex (i.e., nonsimple random) sample design, and (2) minimize bias due to intraclass correlation (which, in the present case, arises due to the fact that data are collected for multiple subjects within households).

For the bivariate analyses, the complex nature of the sample design is to be taken account of by the calculation

---

[7]While the morbidity rates implied by these data are likely to overestimate morbidity in the conventional sense, they may nevertheless be used to assess stratum health differentials if the reporting of accidental injuries were random with respect to experimental strata (i.e., accidents were included in the reports of morbidity in equal proportions in each of the strata).

of variances which reflect the stratification and clustering procedures employed in collecting the data. These are to be used in conducting statistical tests of significance. For the purposes of the multivariate analyses, it will be necessary to utilize a statistical approach and computer software which meet the requirements noted above, thus precluding the use of commonly used software packages such as SPSS or SAS, which assume simple random sampling and make no allowance for intraclass correlation. There does not seem to be any reason for not using the log-linear, jackknifed chi-square approach[8] used in the analyses of the Baseline Data (in this report) for the analyses of the data for the final report, although other options (e.g., generalized weighted least squares) are available and should be considered further.

---

[8]See Robert E. Fay III "On Jackknifing Chi-Square Test Statistics, Part I," April 1981, for a description of this procedure.

III. SPECIFIC OBJECTIVES OF THE BASELINE STUDY

The baseline study aims to evaluate the health and economic conditions in Bacolod and Cagayan de Oro before the completion of the water supply improvements. It will also establish the similarity or dissimilarity of the three experimental groups before the project intervention with respect to the key variables listed in Section II, C.2., and measure the cross-sectional associations between the variables as presented in the conceptual model. The specific research questions to be addressed are shown below.

Baseline Survey Research Questions

1. How do the three experimental strata differ with respect to:

   a. socioeconomic background characteristics?
   b. kinds of water sources used?
   c. quality (safety) of drinking water?
   d. water availability and consumption?
   e. amount paid for water?
   f. health and nutritional status?
   g. sanitary practices?
   h. economic uses of water?

2. Is there an association between water quality (safety) and:

   a. water source?
   b. boiling of water?
   c. chemical treatment of water?

3. What factors influence the incidence of diarrhea among children aged 0-4 years?

4. What factors influence the nutritional status of children aged 0-4 years?

IV. RESEARCH METHODOLOGY

A. Research Instruments

Much of the data was obtained on the basis of a sample survey designed to meet conventional scientific standards. Relevant data were collected from the sample households using a household survey questionnaire and a reporting form for the bacteriological test of water samples. The height and weight of preschool children was also measured using bar scales to determine their nutritional status.

B. Sample Design and Selection

The sample for the Provincial Water Project was selected from the households in two cities -- Bacolod and Cagayan de Oro. In each city a stratified multistage sample design was implemented. First-stage units were divided into strata as follows:

Stratum 1: Units containing households in city water areas which either have piped water or are expected to get it during the study period.

Stratum 2: Units containing households which do not have piped water and are not expected to receive water during the study period.

Cagayan de Oro

The sampling universe for Cagayan de Oro consisted of approximately 78 percent of the entire household population. The remaining 22 percent of the total were excluded from the sampling universe because they were considerably beyond the coverage of the water supply improvement project and because, unlike other households in the city, they were located in extremely rural and mountainous areas.

For Cagayan de Oro the sample is a stratified two-stage design which was conducted in two phases. Interviews were conducted in 1,263 households during the first phase and 256 households during the second phase.

For the first phase the primary sampling units, blocks, were selected with unequal probabilities without replacement within each stratum with any duplicates rejected. The probabilities were based on the number of households per block from the 1975 census of population. The sample selection was carried out using a table of random numbers.

Because the first-phase selection of blocks did not allow blocks to be selected more than once, the blocks were not selected proportional to size as was originally intended. Large blocks (blocks of more than 276 households) were under-represented. In stratum 1 large blocks account for over 39 percent of the total households in 1975,

Previous Page Blank

but represented only 17 percent of the selected sample. If households were weighted by the inverse of their probabilities of selection, households from these large blocks would have weights up to six times larger than the weights on households in smaller blocks. (If a PPS systematic sample had been implemented, households would have been self-weighting.) This variation in the weights would cause a considerable loss in efficiency, roughly 40 percent. Thus a supplemental sample, the second phase, was selected to recuperate much of this loss. The blocks for the supplemental sample were selected probability proportional to size.

Within most of the blocks, the households were listed and a sample of households was selected systematically with probability inversely proportional to the blocks' measure of size (size is the estimated number of households). But in large blocks (blocks larger than 276 households), the blocks were segmented and one segment was selected probability proportional to size. For those blocks selected in both phases, two segments were selected probability proportional to size using the Durbin method.[9] The only block of over 276 households not originally sampled was added to the sample (i.e., selected with certainty).

Bacolod

Of the 60 barangays in Bacolod, 45 were classified as belonging to stratum I, 11 to stratum II, and 4 were not classifiable with respect to the stratification criteria and were excluded from consideration. Excluded barangays represented about 16 percent of all households. In each of the two strata, sampling was done in two stages. First-stage units were barangays. Second-stage units were households in selected barangays.

Barangays within strata were selected by PPS based on the number of households per barangay from the 1975 census of population using a systematic sample. Households within selected barangays were selected using systematic sampling so that 50 to 51 households would be selected. This was accomplished by modifying the original selection rates within first-stage units and led to varying probabilities of selection for households, so that the design implemented was not self-weighting. This selection procedure resulted in about 1,090 selections.

9J. Durbin, "Design of Multi-stage Surveys for the Estimation of Sampling Errors," 1967, Applied Statistics, Volume 16, pp. 152-164

C. Data Processing

A computerized data processing system was developed for the baseline survey. Data were keyed on ENTREX, which provided a limited degree of checking for valid values at point of data entry.

The data processing system consisted of:

1. a program which reformatted the original data that was keyed in fixed field format to data in full-word binary format. This made data referencing easier and less error-prone due to the extremely large number of data items (over 800).

2. a valid values edit which edited the binary questionnaire for keying errors.

3. an update and correction program which changed data values based on user input.

4. a list program which printed user selected questionnaires.

5. a consistency edit program which checked internal record consistency.

6. a program which assigned weights to questionnaires.

7. several recode programs, written to create new variables from existing ones. This was done for use in tabulations and variances.

8. several hundred tabulations which were written in CENTS (Census Tabulation System). These tables, ranging from simple cross tabulations to complicated analytical tables, were also modified and re-run to create percentage tables.

9. a variance program, which performed more than 1,000 variance calculations involving 140 variables.

D. Analysis

Statistical Statements

All statistical statements were tested using conventional "t-test" procedures. Differences that were significant at the 95 percent level of confidence are considered significant, while differences significant at the 90 percent level are referred to in the test as "marginally significant."

X^2 Test

Statements in the bivariate and multivariate analyses involving polytomous characteristics were also tested at the 95 percent level of confidence using the jackknifed chi-square test, which is an analogue of the

familiar Pearson or likelihood ratio chi-square test adjusted for the effect of the complex design and weighting. References to "marginally significant" again reflect testing at the 90 percent level of confidence.

Seasonal Adjustment

An investigation of the appropriateness of adjusting the diarrhea rates obtained from the baseline survey with seasonal indices obtained from related health data was considered. It was not possible to locate complete health records for Cagayan de Oro or Bacolod for the necessary time period, however.

V. BASELINE SURVEY RESULTS

The results of the 1977-78 surveys are presented in this section. Many comparisons are made between experimental strata I, II, and III. Once again, these are defined as follows:

> Experimental stratum I is the group of households using piped water from the existing city water system as their major source.
>
> Experimental stratum II is the group of households in the project area whose major water source is not the existing city water system.
>
> Experimental stratum III is a similar group of households in non-project areas.

A. Population Distribution in the Study Areas

The provincial cities selected for the study differ somewhat in both population size and density. Cagayan de Oro had a population of 165,000 people or approximately 20,000 households according to the 1975 Census. In Bacolod the population was approximately 223,000 people in 36,000 households. The population density in Bacolod was 1,430 people per square kilometer compared to 400 per square kilometer in Cagayan de Oro. Cagayan de Oro was among the eight lowest cities in density for cities of 80,000 or more in the Philippines. Bacolod, on the other hand, was among the top eleven cities in density.

Estimates from the baseline survey indicate that there were 25,473 households in Cagayan de Oro and 33,862 in Bacolod at that time. Of these, 53 percent of the Cagayan de Oro households and 32 percent of the Bacolod households reported that their main source of water was the old City Water System.[10]

---

[10] Some of these households were not directly connected to the system, but used someone else's city water connection.

[11] If the determination of a household's major water source was not possible due to missing data, then that household could not be classified into one of the three experimental strata.

Table 1. Number of Households and Median Household Size by City and Experimental Stratum, 1977-78.

Experimental Stratum	Cagayan de Oro		Bacolod	
	Number of Households	Median Household Size	Number of Households	Median Household Size
All Strata	25,473	6.6	33,862	6.5
Stratum I	12,528	6.8	10,887	6.4
Stratum II	9,938	6.4	18,561	6.6
Stratum III	1,279	7.2	4,072	6.6
Stratum Missing[11]	1,728	6.5	342	7.0

B. Socioeconomic Status of the Population in the Study Areas

The indicators of the socioeconomic status of households collected on the baseline survey include:

- household monetary income
- highest school grade completed by household head
- number of household items owned
- classification of house type by strength of construction
- ownership of house and lot
- source of lighting used

Obtaining measurements of these indicators before the new water systems were completed provides a profile of potential beneficiaries and establishes the similarity or dissimilarity of the experimental strata. Table 2 summarizes the results obtained during the baseline survey.

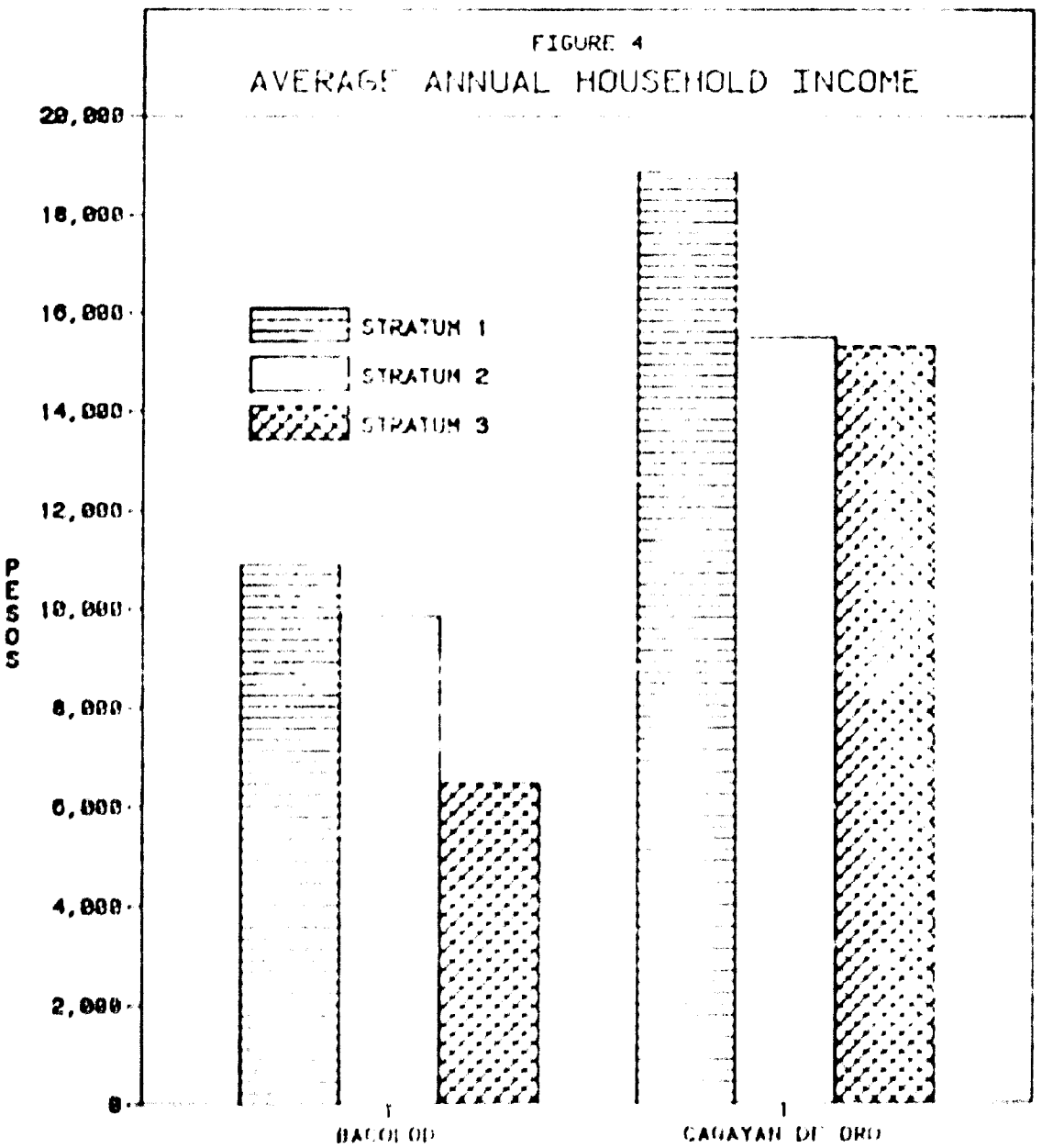
Table 2. Socioeconomic Profile by City and Experimental Stratum, 1977-78 (Standard Errors in Parentheses).

Indicators	Cagayan de Oro Experimental Stratum I	Cagayan de Oro Experimental Stratum II	Cagayan de Oro Experimental Stratum III	Bacolod Experimental Stratum I	Bacolod Experimental Stratum II	Bacolod Experimental Stratum III
Average Annual Household Income (in pesos) (in dollars)	18,865 2,595	15,510 2,133	15,308 2,105	10,901 1,499	9,859 1,356	6,475 890
Average Monthly Household Income	1,572 (153)	1,293 (121)	1,276 (275)	908 (80)	822 (74)	540 (126)
Percent of Household Heads With College Degrees	19 (1.4)	23 (2.1)	14 (3.8)	26 (3.0)	16 (2.0)	9 (4.1)
Average Number of Household Items Owned	7 (0.3)	7 (0.7)	6 (0.8)	7 (0.3)	6 (0.3)	5 (1.0)
Percent of Households Owning House	55 (1.9)	64 (2.1)	71 (3.9)	56 (3.0)	75 (3.0)	71 (9.2)
Percent of Households Owning Lot	18 (1.4)	19 (3.0)	20 (4.5)	18 (3.0)	17 (2.5)	19 (8.6)
Percent of Households Living in Concrete or Wooden Houses, in Good Condition	37 (2.2)	37 (3.4)	24 (4.0)	23 (3.0)	14 (2.1)	13 (6.3)
Percent of Households with Electric Lighting	74 (2.2)	57 (3.1)	45 (4.8)	75 (3.9)	47 (3.0)	29 (9.6)

Each of the socioeconomic indicators are discussed in detail below.

1. Income

Comparisons between strata I, II, and III regarding the average annual monetary household income were made for both cities. The only significant difference was that the stratum I households had relatively higher average incomes than the stratum III households in Bacolod. (See Figure 4.)



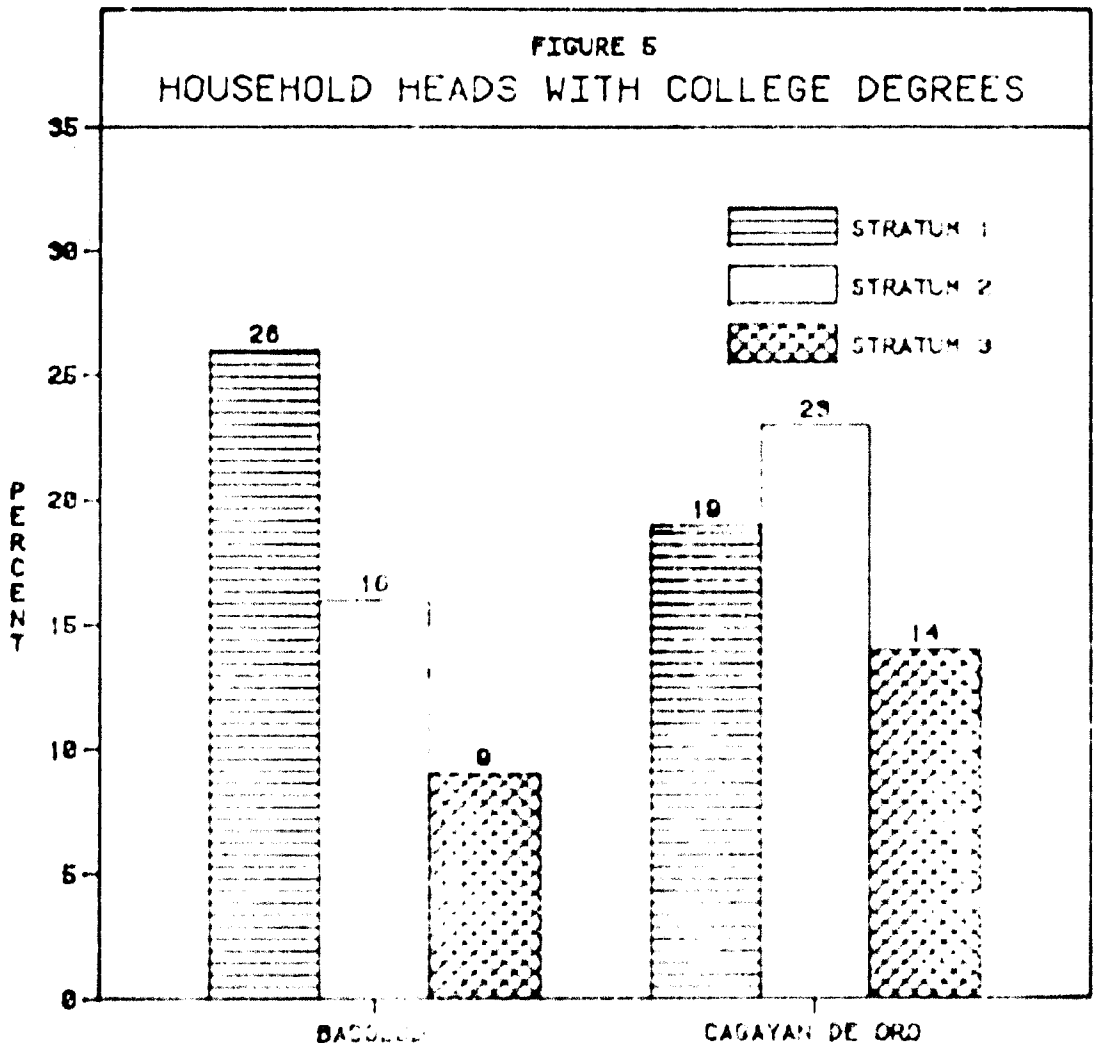
The distribution of households by income level is shown in Table 3. In Bacolod, there is a significantly larger percentage of households at the lowest level of income in strata II and III than in stratum I. In Cagayan de Oro, a larger percent of the stratum II households reported incomes at the lowest level (₱15,000) than stratum III households in both cities.

Table 3. Percent of Households: Household Income by Stratum, 1977-78

Annual Household Income (in pesos)	Cagayan de Oro Experimental Stratum			Bacolod Experimental Stratum		
	I	II	III	I	II	III
All Households (N)	100	100	100	100	100	100
1 - 4,000	16	19	25	31	40	48
4,001 - 8,000	24	30	33	32	33	35
8,001- 15,000	32	23	24	17	14	9
more than 15,000	28	28	17	20	14	8

2. Educational Level of Household Head

In Bacolod, the stratum I households had a significantly higher percentage of household heads with college degrees than the strata II and III households. In Cagayan de Oro, the percentage of household heads with college degrees was marginally greater for strata II when compared to stratum III. (See Figure 5.)

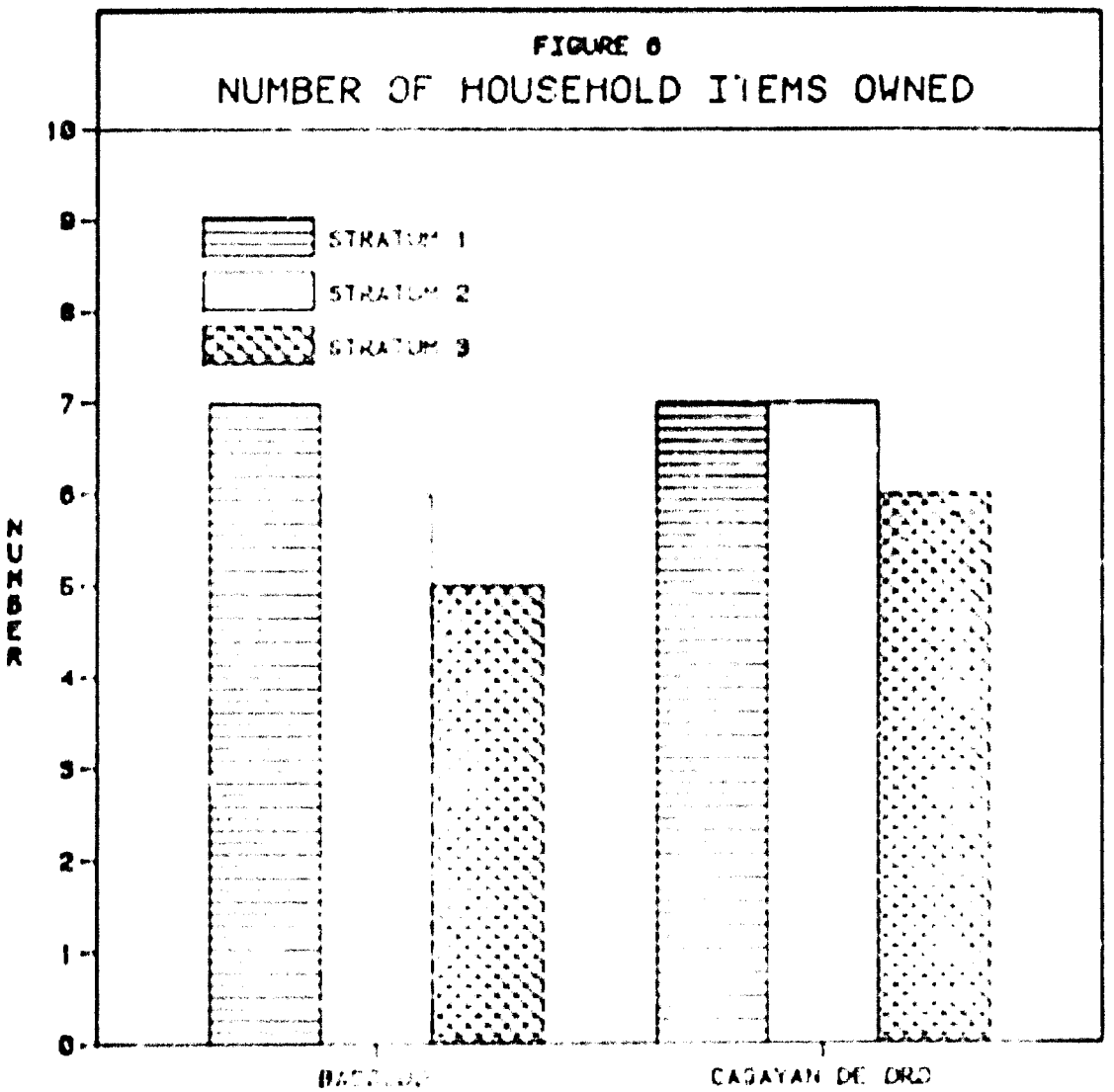


1977-78

3. Number of Household Items Owned

The households in stratum I, II, or III do not differ significantly with respect to the number of household items owned ex ept in Bacolod, where there is a significant, but small diffe ence between stratum I and II households and between stratum I and stratum III households. Household items include a sofa set (considered 1 item), dining room set (1 item), dressing table, clothes closet, radio, phonograph, stereo set, television, range, refrigerator, air conditioner, washing machine, sewing machine, electric fan, chinaware set (1 item) clock, and shelf of books (1 item). (See Figure 6.)

The results for Cagayan de Oro for this indicator might be inconclusive due to unusually high nonresponse.



1977-78

4. Classification of House Type

Only 17 percent of the Bacolod households and 34 percent of the Cagayan de Oro households live in structures described as "permanent, concrete and wood". Most of the households live in semi-permanent dwellings which were made of cheap materials and were physically run-down. In Bacolod, approximately 16 percent live in structures even less permanent.

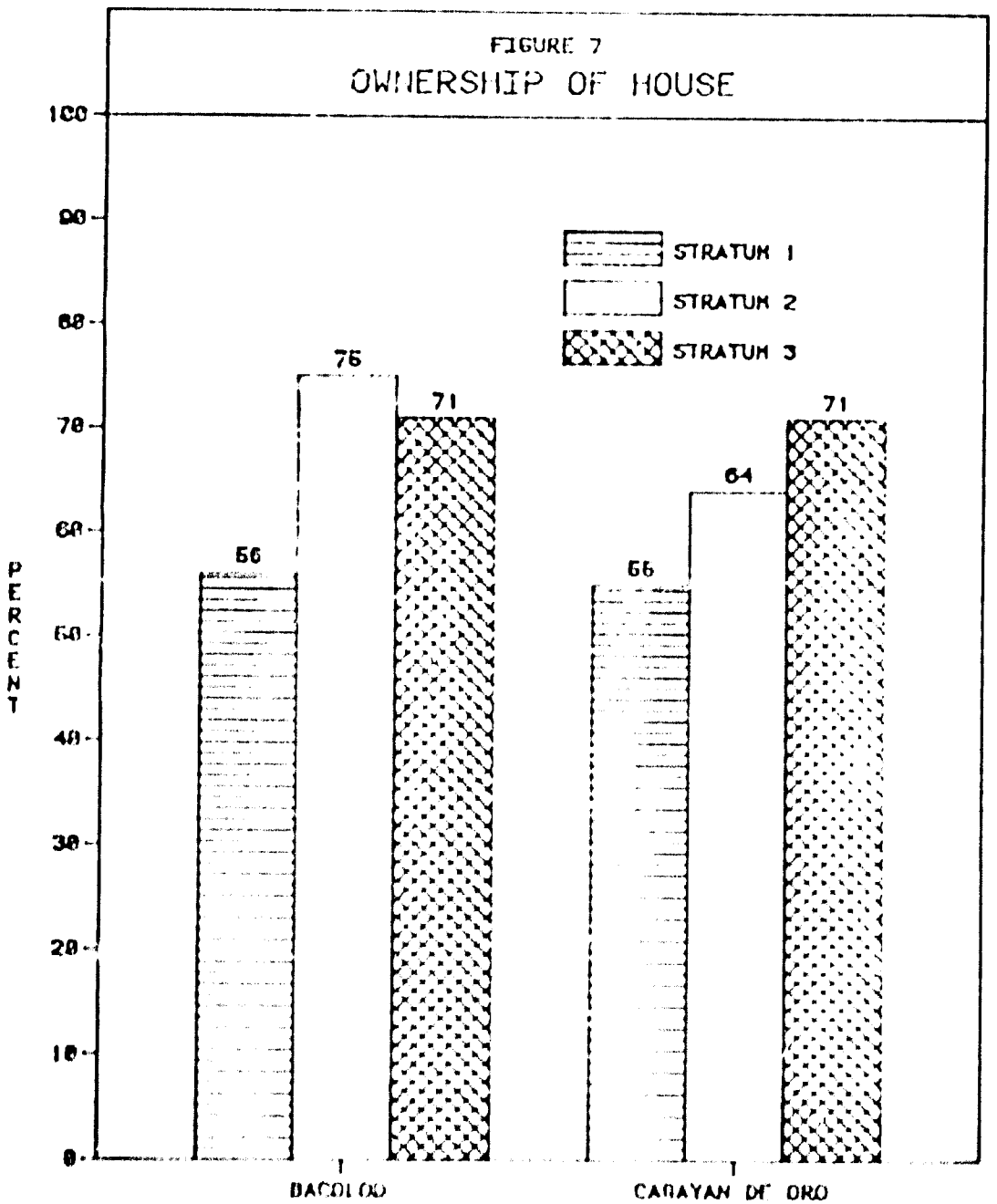
In Bacolod, more of the stratum I households live in the permanent structures than stratum II or III households. In Cagayan de Oro, a greater percentage of stratum I households than stratum III households were observed to live in the permanent structures.

Table 4. Percent of Households: Classification of House by City and Experimental Stratum, 1977-78.

Classification of House	Cagayan de Oro			Bacolod		
	Experimental Stratum			Experimental Stratum		
	I	II	III	I	II	III
Total households.....	100	100	100	100	100	100
Permanent: concrete and wood..	37	37	24	23	14	13
Semi-permanent: cheap materials....	42	35	50	29	25	25
Permanent or semi-permanent in poor repair or physically run-down.....	19	26	26	39	41	41
Barong-Barong.....	3	2	-	9	19	22

5. Ownership of House and Lot

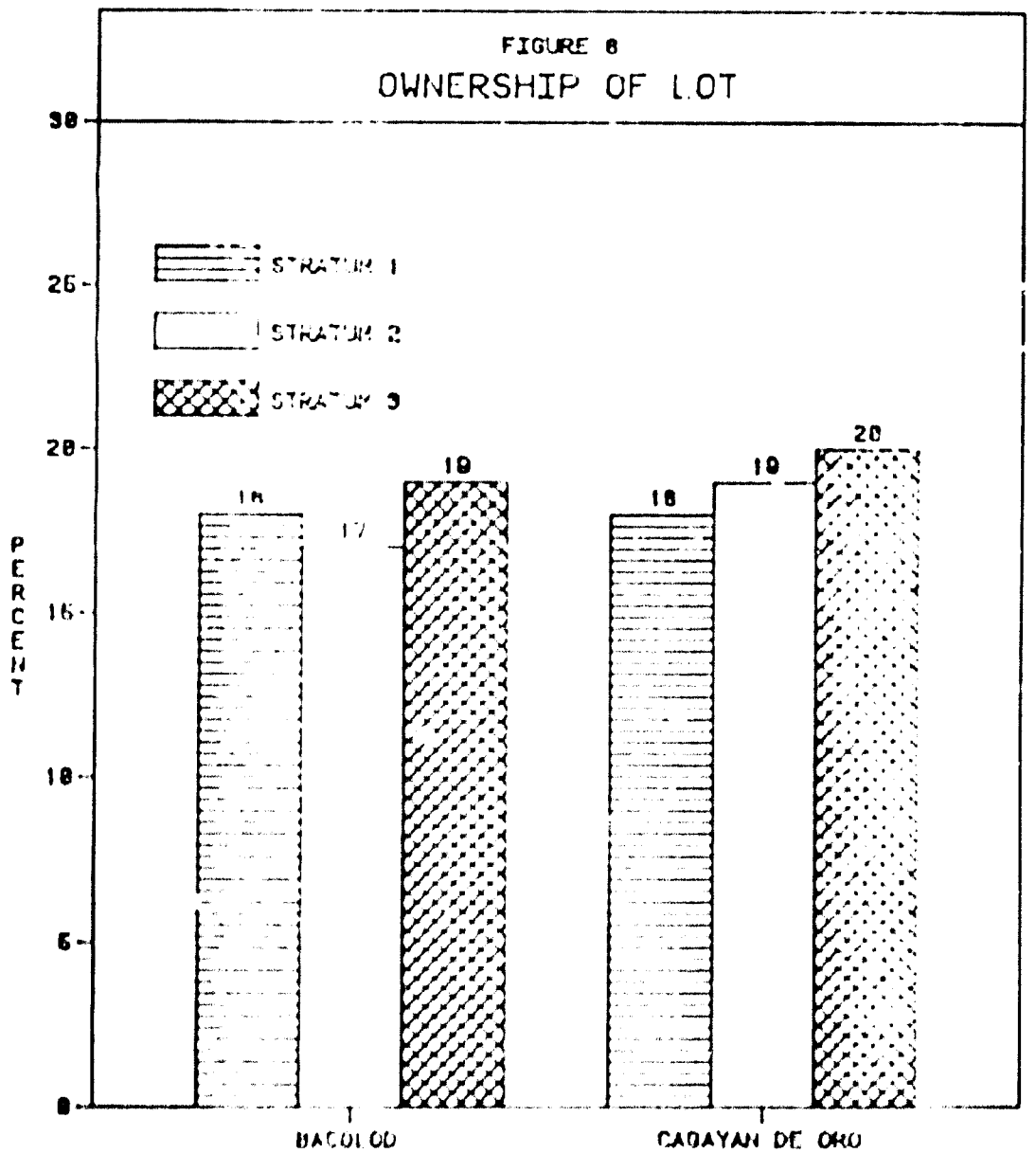
The survey results show that more of the stratum II households in Bacolod and the stratum II and III households in Cagayan de Oro own their houses than the stratum I households. (See Figure 7.)



1977-78

Surprisingly, less than 20 percent of the households in both cities own their lot. There was no significant variation between the strata with regard to lot ownership in either city. (See Figure 8.)

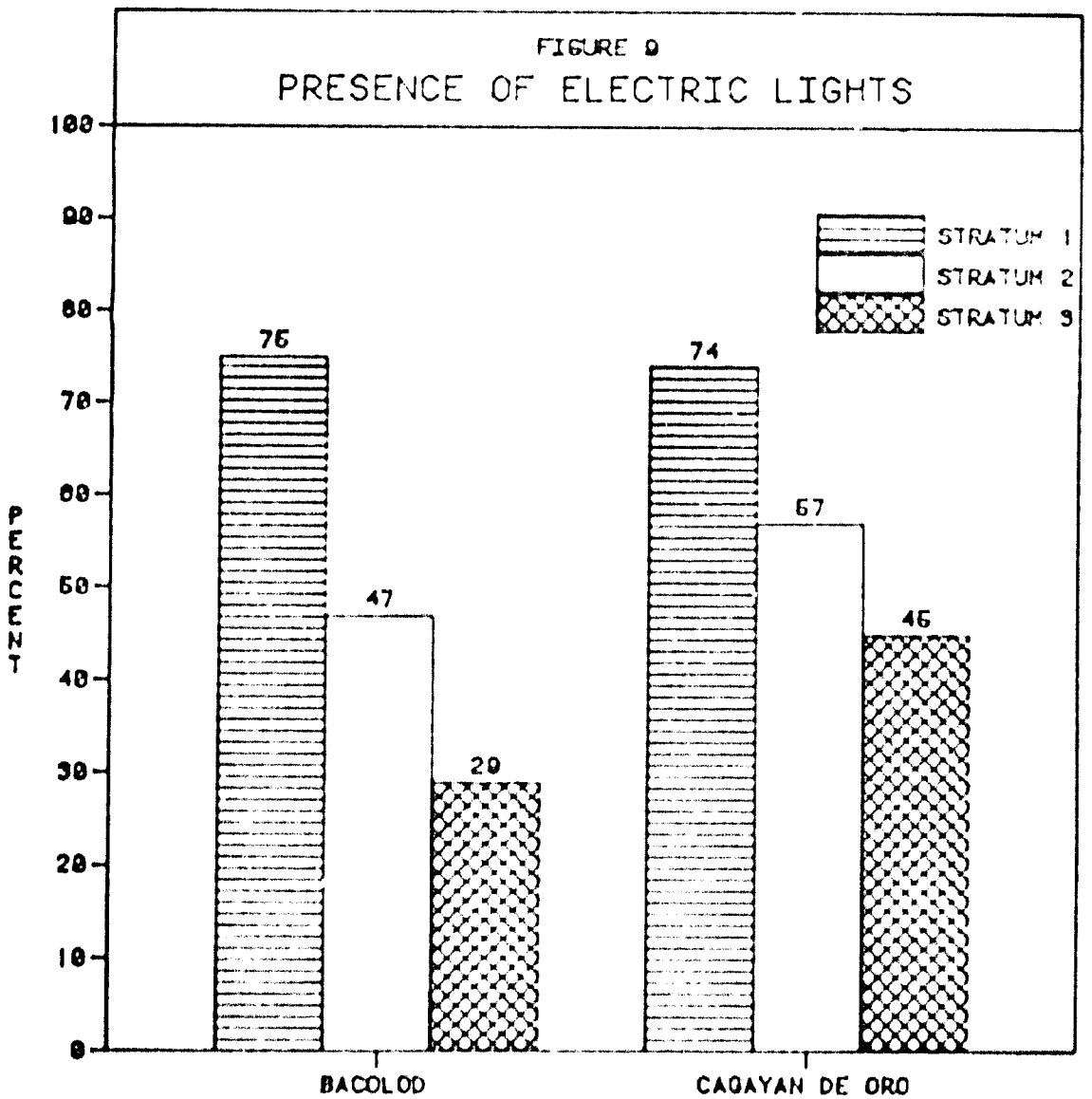
House and lot ownership does not seem to be a good socioeconomic indicator because a poor household might own their "barong barong" house while a relatively more affluent household might not own their "permanent" structure.



1977-78

6. Source of lighting

Almost all the households in both cities used either electric or gas lighting. Generally, the presence of electric lights in these areas is an indication of a higher standard of living. In both cities, 75 percent of the stratum I households used electric lighting. Accordingly, they were relatively better-off than the stratum II and III households, where in Bacolod 50 percent (stratum II) and 33 percent (stratum III) used electric lights and in Cagayan de Oro 47 percent (stratum II) and 29 percent (stratum III) used electric lights. (See Figure 9.)



1977-78

C. Water Sources in the Study Areas

More than half of the households in the study areas used piped water before the new systems were built. Seventy-one percent of the Cagayan de Oro households and 40 percent of the Bacolod households reported that they used piped water as their main water source. In most cases the piped water was from the existing city water system. Based on the 1978 records of the Cagayan de Oro and Bacolod Water Districts, there were 6,444 residential service connections in Cagayan de Oro and 3,937 in Bacolod. (See Table 5.)

Table 5. Number of Service Connections, 1977-78

Type of Connection	Cagayan de Oro		Bacolod	
	1977	1978	1977	1978
Total, all types	4,507	6,677	-	4,375

Source: Cagayan de Oro and Bacolod Water Districts Records, 1978

Only 23 percent of the households in the study areas (27 percent in Cagayan de Oro and 19 percent in Bacolod) reported that their major source was water piped into the house. The other users of piped water utilized outside sources.

Other households reported using open or covered wells, rivers, streams, rainwater, and purchased water as their main water source. Thirty-five percent of the Bacolod households and 12 percent of the Cagayan de Oro households used covered wells, 22 percent in Bacolod and 7 percent in Cagayan de Oro used open wells, and 3 percent in Bacolod and 11 percent in Cagayan de Oro used one of the other sources mentioned above.

Table 6. Percent of Households: Main Water Source By City and Experimental Stratum, 1977-78

Main Water Source	Cagayan de Oro Experimental Stratum				Bacolod Experimental Stratum			
	Total	I	II	III	Total	I	II	III
All households	100	100	100	100	100	100	100	100
Piped inside	27	32	25	10	19	42	9	10
Piped outside	44	67	16	56	21	58	3	5
Covered well[a]	12	Z	26	23	35	Z	54	40
Open well[b]	7	1	12	7	22	-	31	34
Other sources[c]	11	-	21	3	3	-	3	11

[a]Includes artesian well and electrically pumped covered wells.

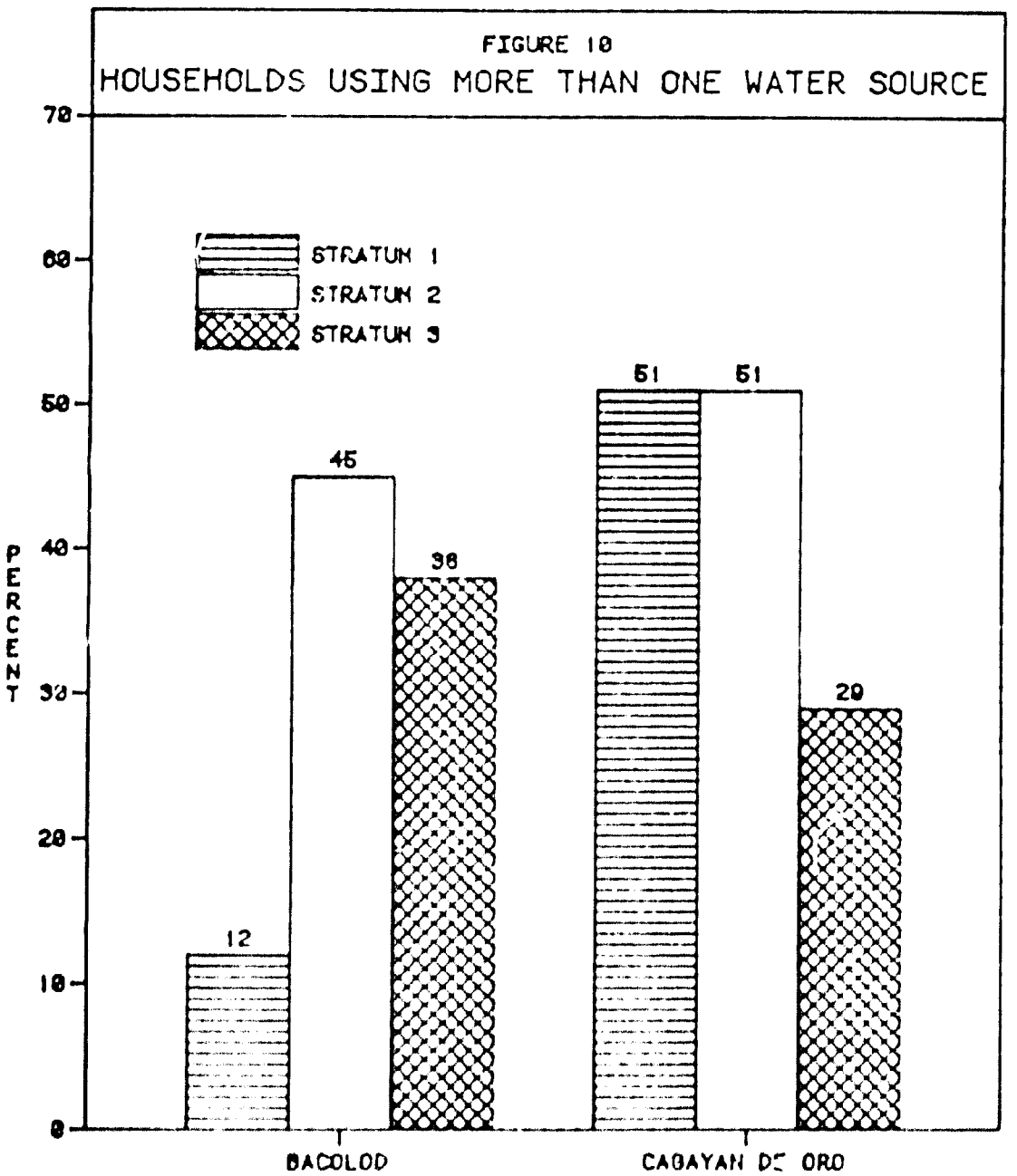
[b]Includes pumped and manually-drawn open wells.

[c]Includes rivers, streams, rainwater, bought water, and other sources not included in previous categories.

Z = less than 1/2 of 1 percent.

- = no cases

Forty-one percent of all households (51 percent in Cagayan de Oro and 33 percent in Bacolod) reported that they utilized at least one additional water source. Many more of the households in the city water area (strata I and II) used additional sources than those in noncity areas in Cagayan de Oro. In Bacolod, less of the stratum I households utilized additional water sources than either the stratum II or stratum III households. (See Figure 10.)



1977-78

D. Households' Opinions On Their Main Water Source

To provide a comparison between the City Water System users (stratum I) and nonusers (strata II and III), the households' opinions on their main water source was obtained. Below is a table presenting the percentage of households by the level of satisfaction with their water source.

Table 7. Percent of Households: Level of Satisfaction with Service of Their Present Water Source By City and Experimental Stratum, 1977-78

Level of Satisfaction	Cagayan de Oro Experimental Stratum			Bacolod Experimental Stratum		
	I	II	III	I	II	III
Very Satisfied	2	3	4	5	3	6
Satisfied	45	27	38	67	70	68
Dissatisfied	48	62	57	27	27	25
Very Dissatisfied	5	7	1	1	z	1

As mentioned earlier, the survey estimated that 50 percent of the households in Cagayan de Oro and 32 percent in Bacolod City utilized a City Water System source (not necessarily their own). Of these, very few in both cities reported that they were very satisfied or very dissatisfied with their water source.

In Bacolod City, the majority of the households in all strata were satisfied with their water source. The most frequent stated reason for their satisfaction was that water was always available. Most of the households who were dissatisfied reported either the water flow was inadequate or the source was too far from their residence.

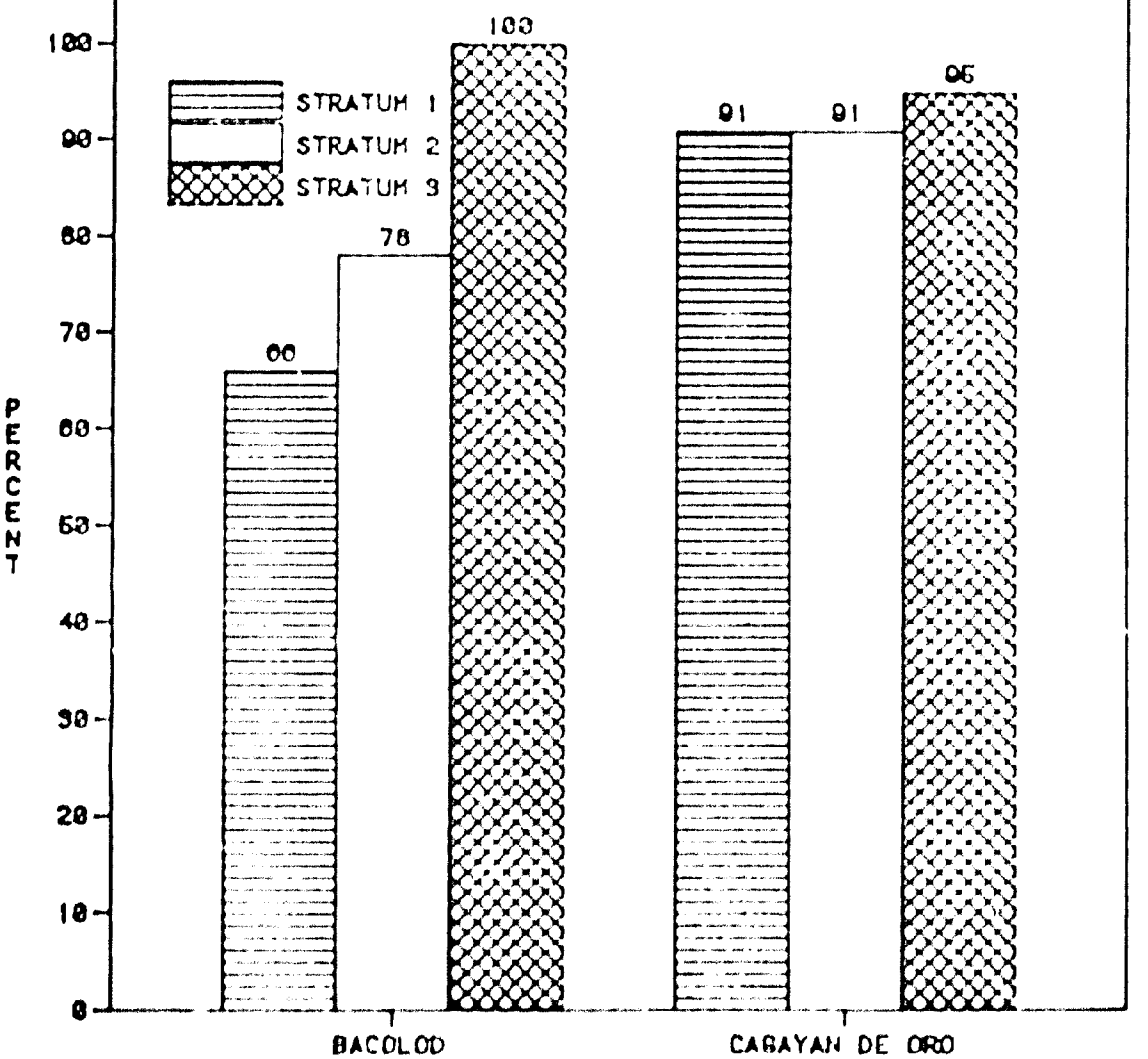
In Cagayan de Oro, more households in all strata reported that they were dissatisfied than satisfied with their water source. The reasons for dissatisfaction among the city water users were related to the inadequacy or unreliability of the water flow, while nonusers reported that water was too expensive or unreliable. The satisfied households reported that water was always available for their needs.

E. Water Quality

The results of the water quality tests conducted for the baseline study showed that most (82 percent) of the samples tested in both cities were below the Philippine national standards for safe drinking water. In Bacolod, the proportion of households with unsafe water was highest in stratum III (100 percent), followed by stratum II (78 percent), and stratum I (66 percent). There were no significant differences in Cagayan de Oro. (See Figure 11.)

FIGURE 11

HOUSEHOLDS WITH UNSAFE WATER



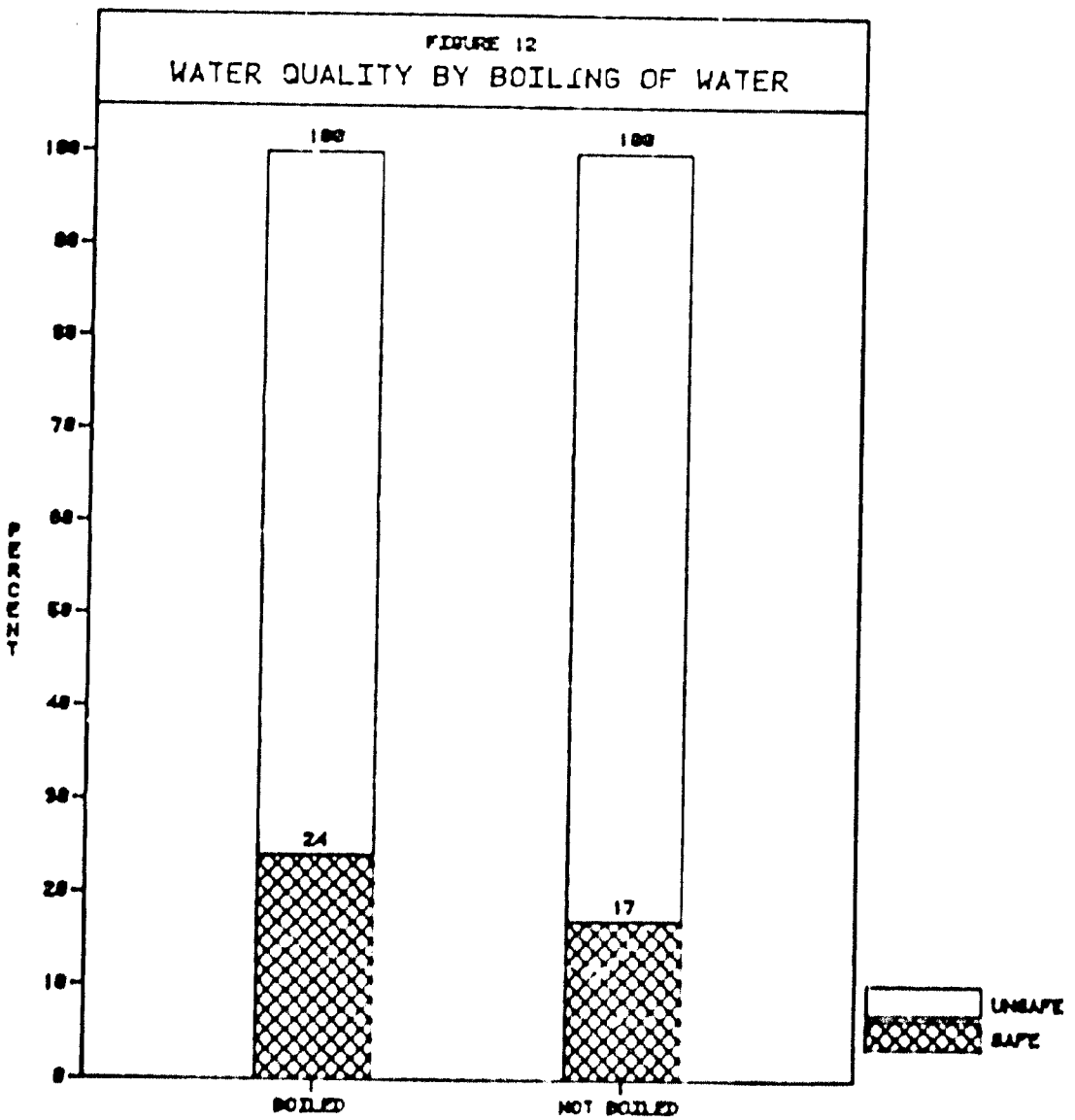
1977-78

More of the households whose main water source was piped into the house had safe water. (See Table 8.)

Table 8. Percent of Households By Water Quality Results and By Main Water Source, 1977-78

Water Quality	All Sources	Piped Inside	Piped Outside	Covered Well	Open Well	Other Source
Both Cities	100	100	100	100	100	100
Safe	18	28	16	15	14	9
Unsafe	82	72	84	85	86	91
Cagayan de Oro	100	100	100	100	100	100
Safe	9	15	7	6	3	5
Unsafe	91	85	93	94	97	95
Bacolod	100	100	100	100	100	100
Safe	23	39	27	17	16	17
Unsafe	77	61	73	83	84	83

The survey results showed a marginally significant relationship between the boiling of household drinking water and the quality (safety) of the sampled drinking water. Households reporting that they usually boiled their water were more likely to have safe water than those who did not. (See Figure 12.)



1977-78

No conclusive statement could be drawn from the data on household treatment of water with chemicals because too few households (1.4 percent) reported treating their drinking water with chemicals.

F. Water Availability and Consumption

The indicators that were obtained to determine the relative availability of water were:

percent of households storing water
percent of households fetching water
average distance from house to water source

The quantity of water consumed by the households was not directly measured in the baseline survey. Instead, some partial measurements were used, including number of gallons of water fetched per household member per week and number of gallons of water bought per household member per week.

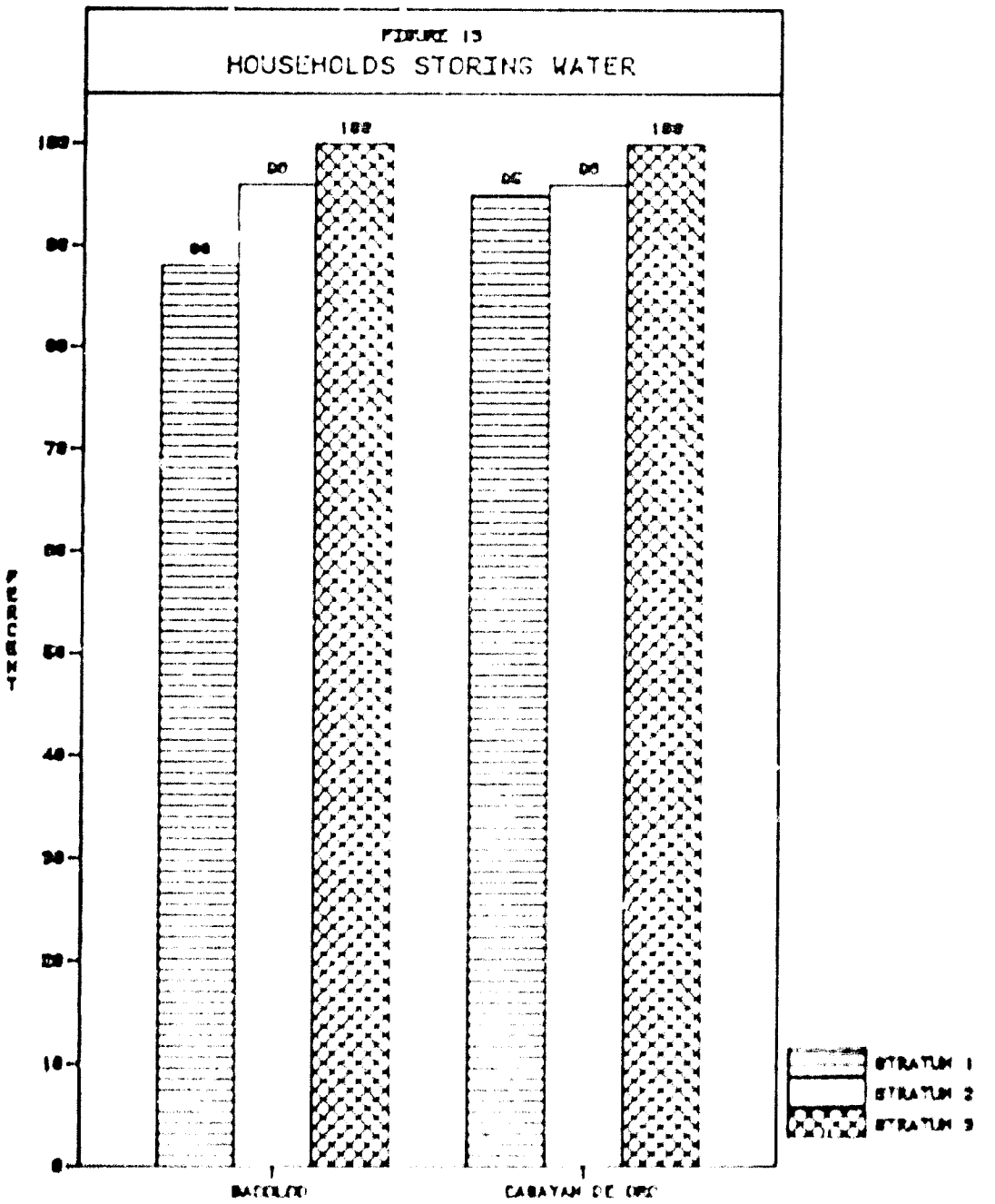
A summary of these indicators is presented in the following table.

Table 9. Water Availability and Consumption By City and Experimental Stratum, 1977-78

Indicators	Cagayan de Oro			Bacolod		
	Experimental Stratum			Experimental Stratum		
	I	II	III	I	II	III
Percent of households storing water	95	96	100	88	96	100
Percent of households fetching water	63	71	87	57	83	86
Average distance (meters) from house to main water source ($\bar{X}$)	84	96	96	21	36	51
Number of gallons fetched/household member/week ($\bar{X}$)	8	8	8	11	13	11
Number of gallons bought/household member/week ($\bar{X}$)	6	12	2	5	5	4

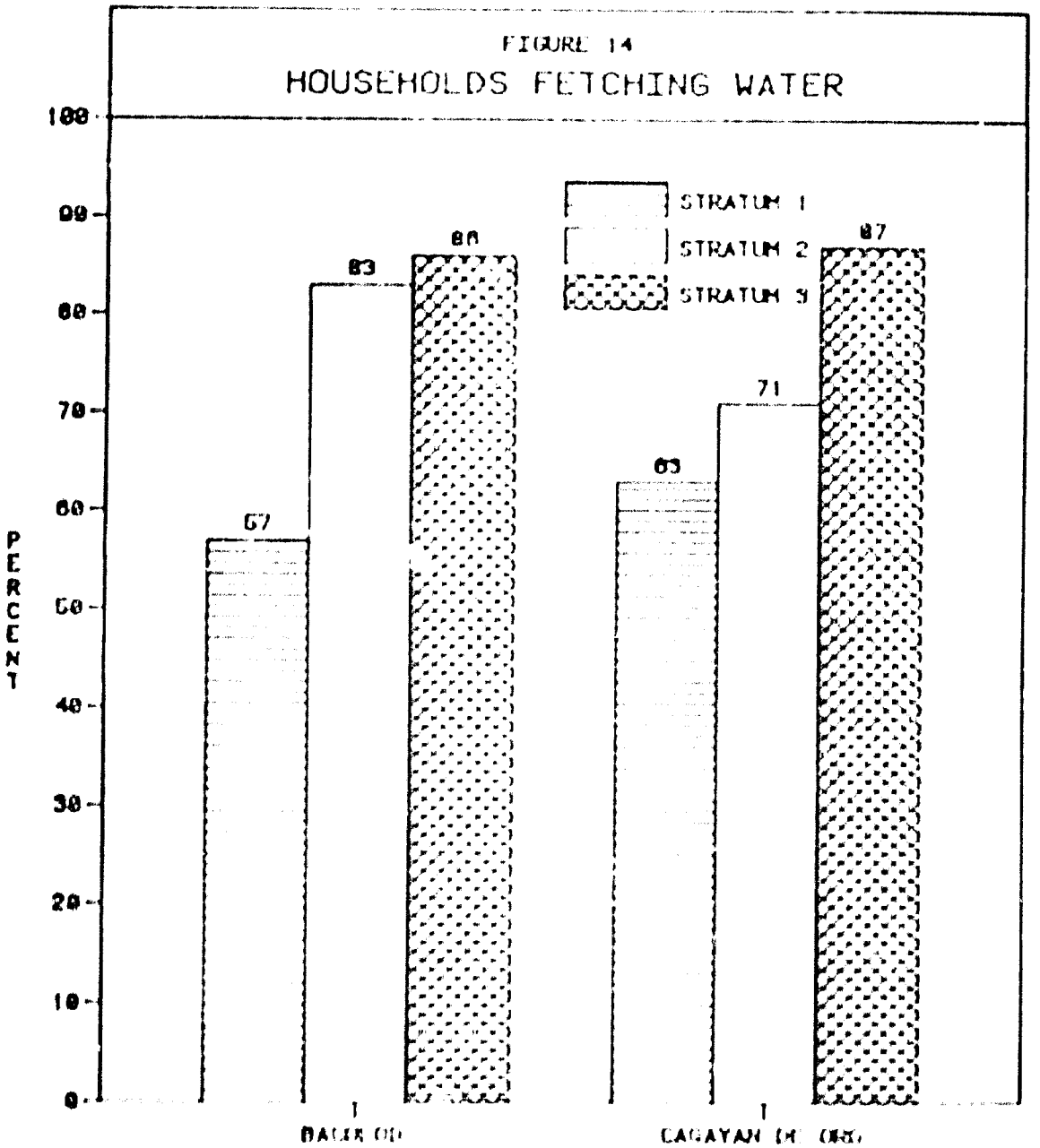
1. Storing of Water

Nearly all households in both cities, (95 percent) stored their water for household uses, indicating that either the water supply might not have been sufficient 24 hours a day, or that the household preferred to have stored water available. (See Figure 13.)



2. Water Fetching

Seventy-three percent of the households in both cities fetched water. This included 61 percent of the stratum I households, 79 percent of the stratum II households, and 86 percent of the stratum III households. In each city, more than 85 percent of the stratum III households fetched water. (See Figure 14.)

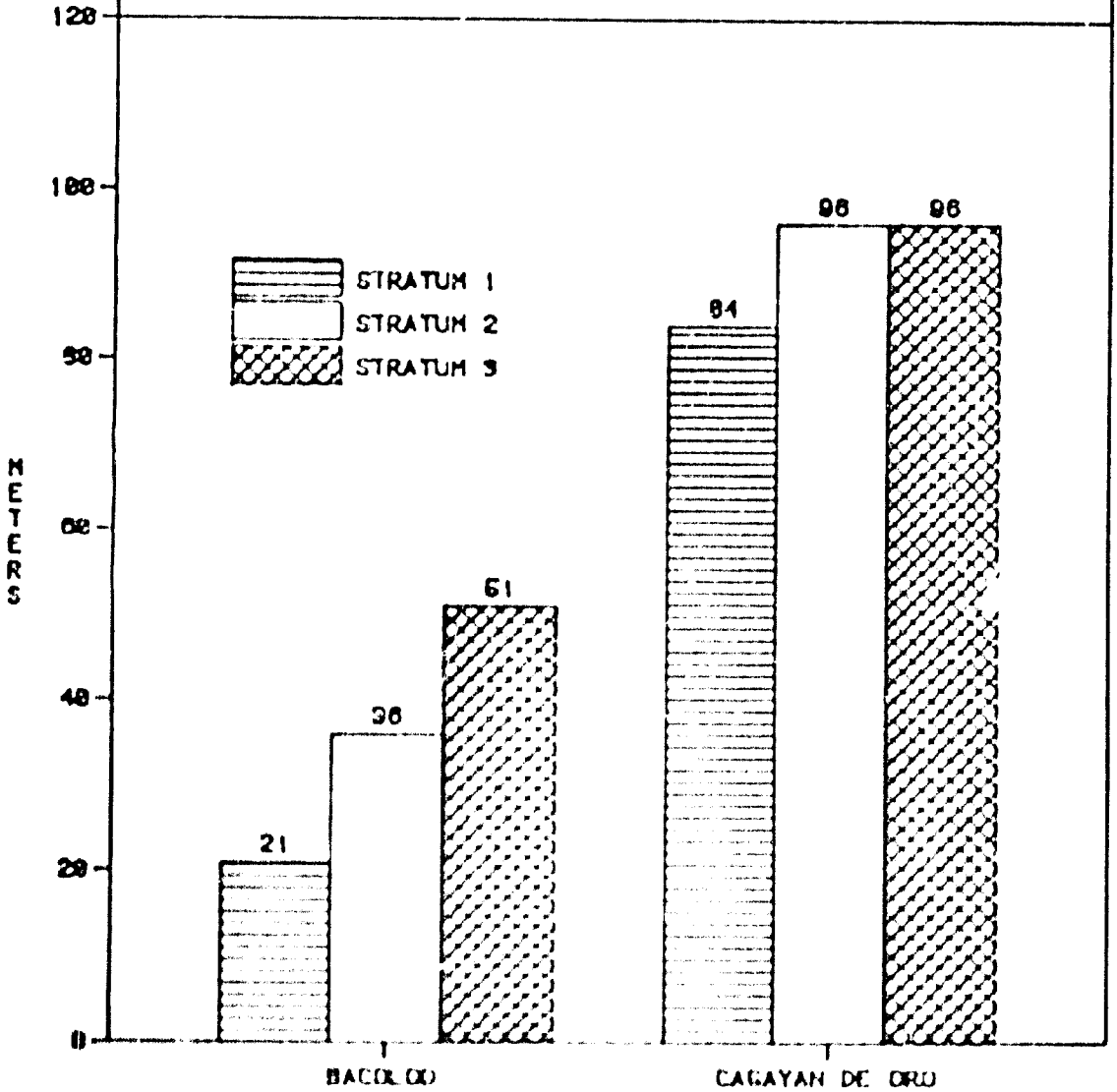


1977-78

3. Distance from House to Main Water Source

As previously mentioned, 27 percent of the Cagayan de Oro households and 19 percent of the Bacolod households used water sources that were piped into the house. The other households had to walk some distance to obtain water from the source. In Cagayan de Oro, the average distance from the house t the source for these households did noit vary between the strata. In Bacolod, the stratum I households were closer to the water source, on the average, than the stratum II or III households. (See Figure 15.)

FIGURE 16
AVERAGE DISTANCE FROM HOUSE TO MAIN WATER SOURCE



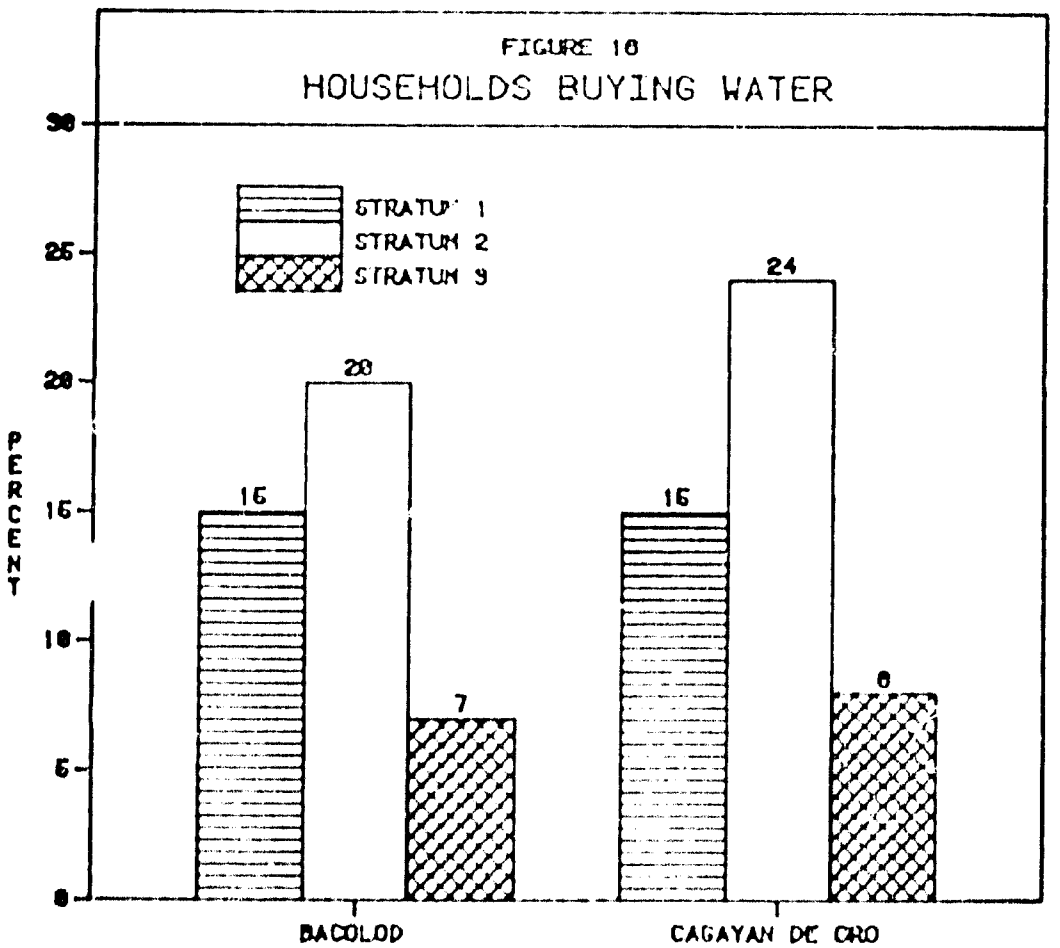
1977-78

4. Average Gallons of Water Fetched

On the average, a household fetched about 73 gallons of water per week -- 11 gallons per household member. In Cagayan de Oro, households in each of the three strata fetched about 8 gallons per household member. In Bacolod, stratum I and II households fetched an average of 11 gallons per household member while stratum III households fetched 13 gallons per household member.

5. Bought Water

As shown in Figure 16, most of the households in both cities reported that they did not purchase water. Nevertheless, a significantly lower percentage of stratum III households bought water than stratum II households. The percentage of stratum III households who bought water was also lower than the percentage of stratum I households (this relationship was only marginally significant in Cagayan de Oro).



In Cagayan de Oro, the average number of gallons of water that a household bought in a one-week period varied substantially between strata with stratum II having the highest estimate (12 gallons). In Bacolod, the average ranged from 4 to 5 gallons in the three strata with little variability. Table 10 shows the survey results in terms of gallons of water bought per household member.

Table 10. Average Number of Gallons of Water Bought Per Household Member, 1977-78.

	Cagayan de Oro Experimental Stratum			Bacolod Experimental Stratum		
	I	II	III	I	II	III
Average Gallons Per Household Member	0.9	1.7	0.3	0.7	0.7	0.5

G. Cost of Water in the Study Areas

The survey estimated that among the 59,335 households in the study areas, 12,596 or 21 percent received water bills. Of these 12,596 households, almost all (96 percent) were in stratum I and 62 percent had water piped into their houses. For households with inside connections, the average water bill for the month preceding the interviews was ₱25.3 ($3.33) -- ₱25.4 in Cagayan de Oro and ₱25.30 in Bacolod. For households who used at least one source that was part of the city water system, but not piped into their houses, the average bills were ₱16 ($2.13) overall -- ₱12 in Cagayan de Oro and ₱24 in Bacolod. However, almost half of the households in both cities received a bill of less than ₱1 ($0.13).

As previously mentioned, there were also 10,580 households in the study areas that purchased water from vendors or neighbors. This included 15 percent of the stratum I households, 21 percent of the stratum II households and 7 percent of the stratum III households. These households paid an average of ₱0.08 ($0.01) per gallon of water -- ₱0.13 in Cagayan de Oro and ₱0.04 in Bacolod. (See Table 11 for the average centavos paid per gallon.)

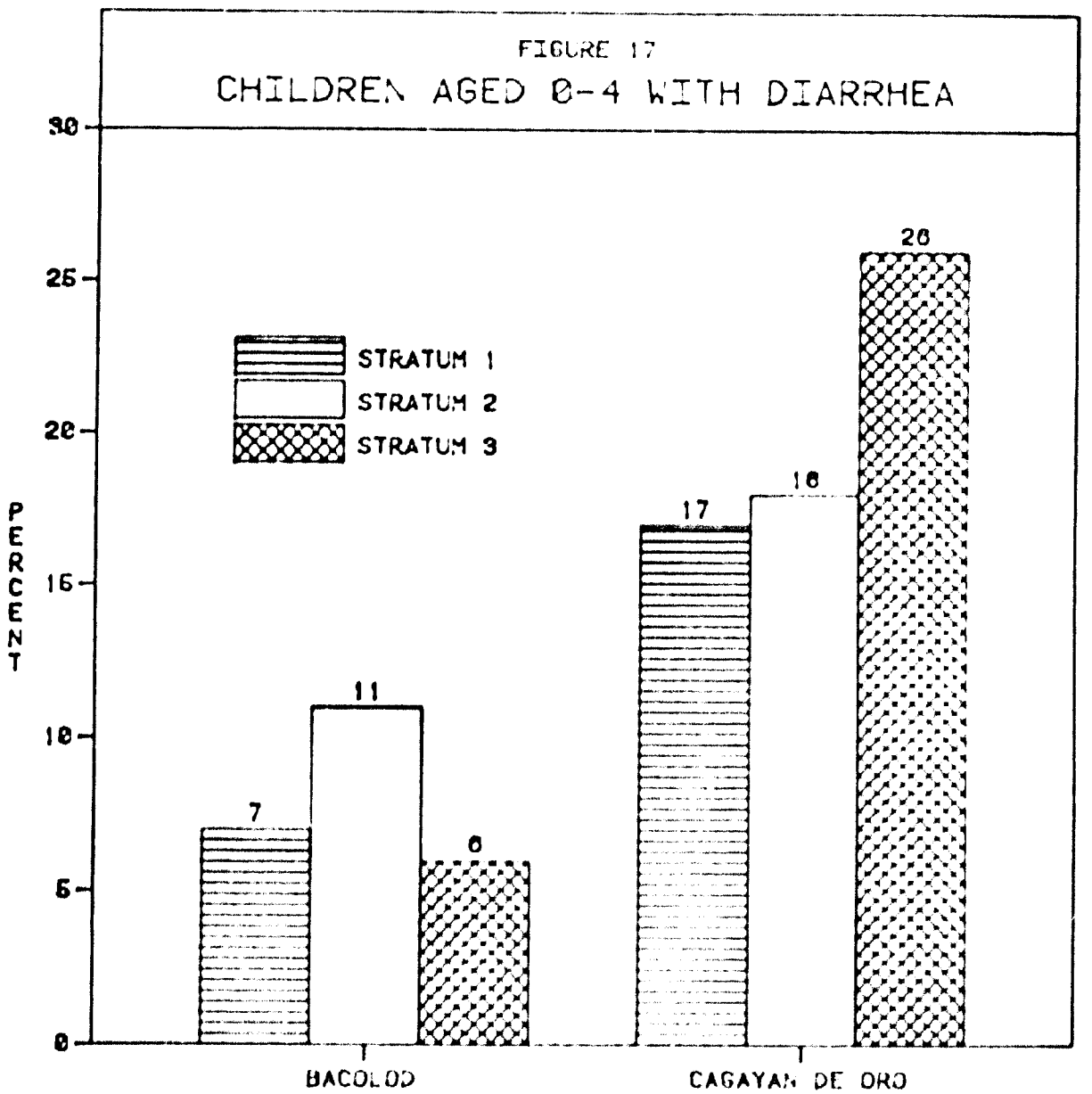
Table 11. Average Centavos* Paid Per Gallon of Water By City and Experimental Stratum, 1977-78

	Cagayan de Oro Experimental Stratum				Bacolod Experimental Stratum			
	Total	I	II	III	Total	I	II	III
Average Centavos Paid Per Gallon	13	18	9	11	4	3	4	1

* A centavo is 1/100 of a peso or about 1/8 of a U.S. penny.

H. Health and Nutrition

For this study, the main indicators of water-related impacts are the nutritional status of children and incidence of diarrhea among children. The baseline survey obtained measures of height, weight, and incidence of diarrhea during the day before interview for children aged 0-4 years. The data shows that there were no significant differences between groups of children in the three experimental strata with respect to these variables. The differences between cities are significant, however. The baseline measures are shown in the following figure and table.



1977-78

Table 12. Percent of Children 0-4: Nutritional Level[13] By City and Experimental Stratum, 1977-78

Nutritional Status	Cagayan de Oro Experimental Stratum			Bacolod Experimental Stratum		
	I	II	III	I	II	III
Normal	40	44	41	37	37	22
1st degree malnourished	41	41	43	34	36	29
2nd degree malnourished	15	12	11	26	24	37
3rd degree malnourished	3	3	4	7	6	12

An objective of the baseline survey was to observe associations between these main indicators and other variables collected. The following analysis describes relationships between incidence of diarrhea, nutritional status, and other variables for children aged 0-4 in the baseline survey. The analysis presented here may be divided into two parts. The first part, a "bivariate analysis", examines associations between pairs of variables, for example, the association between diarrhea and age. The second part, a "multivariate analysis", expresses more complex relationships that may arise among three or more variables.

The statistical methodology employed was log-linear analysis. This technique allows researchers to express both simple and complex relationships present in cross-classified data with a flexibility comparable to that provided linear regression and related techniques in the case of continuous variables. Numerous books and technical articles have appeared on the subject.

Although most of the statistical literature on log-linear analysis pertains to the case of simple random sampling, recent developments have extended the results to complex samples. The statistical tests reported here were computed by the program CPLX (Contingency Table Analysis for Complex Sample Designs), developed at the U.S. Bureau of the Census. The tests take into consideration both the weights applied to the sample cases and the manner in which the sample was selected.

---

[13]The standards for nutritional level (normal, first degree, second degree, and third degree malnourished) were determined by the National Nutrition Council (NNC) in the Philippines. Weight Chart for Preschoolers (0-59 months) tables produced by the NNC were used in this analysis.

In the bivariate part of this analysis, the question examined is whether two variables are associated in a statistically significant manner; that is, whether the observed differences in the distribution of one variable at the different levels of the other variable are beyond the range that could be attributed to sampling variability alone. The familiar Pearson chi-square test is often used for this purpose in the case of simple random samples.

The CPLX program used in this analysis produces a test statistic, a jackknifed chi-square test, that performs an equivalent function. The multivariate analysis extends the bivariate analysis by considering the joint effect of variables on each other. This general notion may be illustrated by the following example. Suppose that diarrhea was shown in the bivariate analysis to be related to both age and nutritional status separately. Multivariate analysis of the three variables simultaneously could produce any of the following outcomes.[14]

1. Diarrhea is associated with age, but nutrition has no additional effect when age is considered.
2. Diarrhea is associated with nutrition, but age has no separate effect when nutrition is considered.
3. Both age and nutrition have effects on diarrhea. The separate effects of age and nutrition combine in an additive manner; i.e., the effect of age does not depend on the level of nutrition, and vice versa.
4. Both age and nutrition have effects on diarrhea, but the effect of one variable depends on the level of the other.

In the first two outcomes, one of the two variables "accounts for" or "explains" the effect of the other variable on diarrhea observed in the bivariate analysis. In the last two outcomes, both variables affect the incidence of diarrhea, with neither fully explaining the effect of the other. In the third outcome, the effect of each variable on diarrhea does not depend upon the level of the other variable, whereas in the last outcome, both variables may be said to interact in their effect. The distinction between 3 and 4 rests upon the logistic scale implied by the log-linear analysis, i.e., in 3 the effects of the two variables are additive on a logistic scale.

In the case of simple random sampling, chi-square tests are again used to distinguish between these outcomes. CPLX again provides "jackknifed" versions of these tests for the case of complex samples.

It should be recognized that the bivariate analysis does not identify true causes of diarrhea. The multivariate analysis begins to approach a causal model to the extent that some variables may be identified as explaining the apparent effect of others. However, even with the multivariate analysis, cause, in a strict sense, is not firmly established.

---

[14]This example utilizes incidence of diarrhea as the dependent variable. In other situations, nutritional status could be the dependent variable.

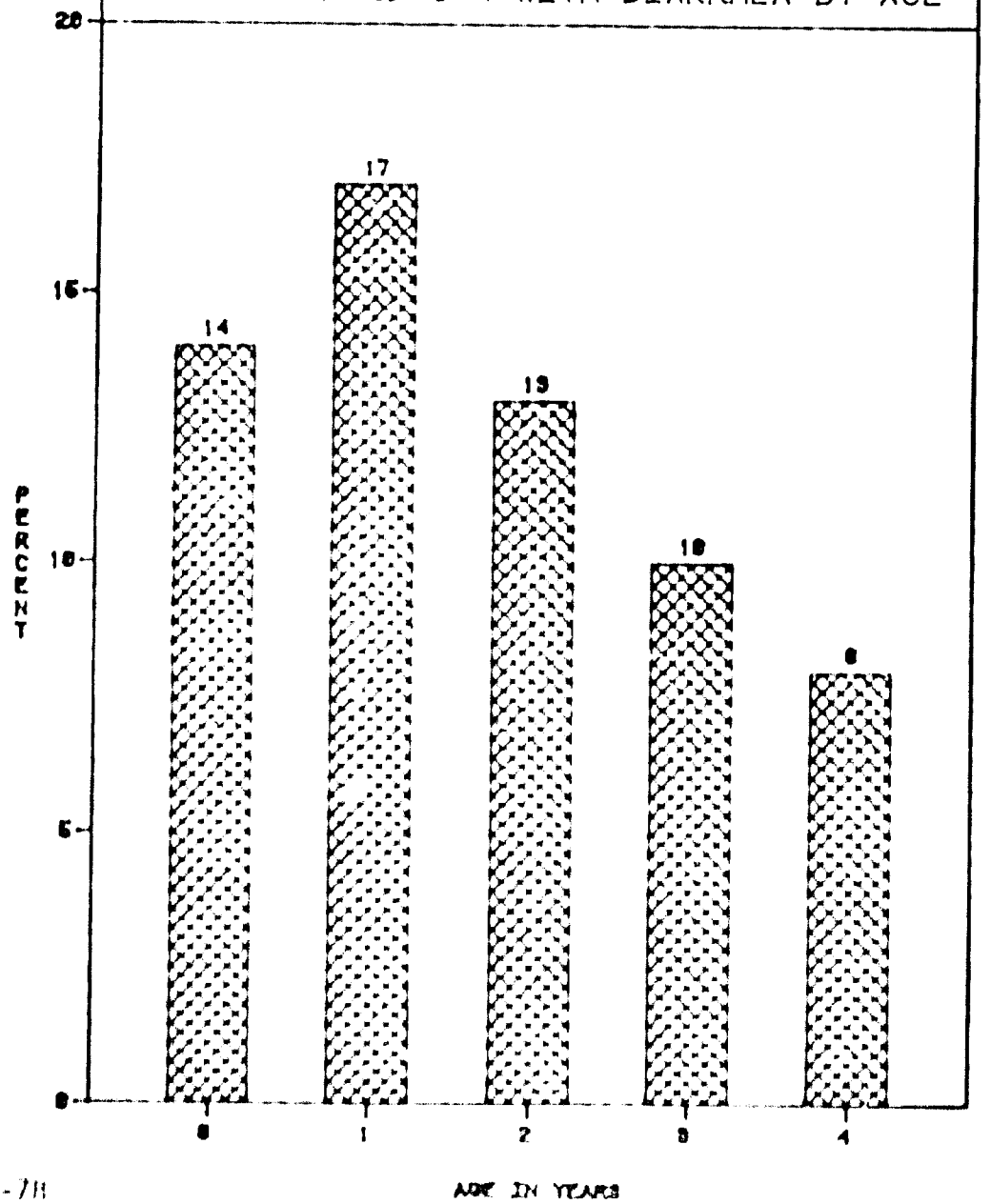
1. Bivariate Analysis - Diarrhea

The variables that were found to be associated with incidence of diarrhea among children were the child's age, the child's nutritional status, the household's method of excreta disposal, the household's type of bathroom, the distance between the house and the toilet, the frequency of handwashing after using the toilet and after children were playing outside, and the method of garbage disposal. These are elaborated upon below.

There is a high association ($p \leq .01$) between the age of a child and the incidence of diarrhea. Children between the ages of 0 and 2 had a much higher diarrhea rate.

FIGURE 18

CHILDREN AGED 0-4 WITH DIARRHEA BY AGE



1977-78

The method of excreta disposal is associated with incidence of diarrhea ($p<.01$). In households having automatic flush toilets, fewer children had diarrhea than in households with other systems.

Table 13. All Children: Diarrhea in the Past 24 Hours By Method of Excreta Disposal, 1977-78

Method of Excreta Disposal	DIARRHEA			
	Yes	Percent	No	Percent
Automatic flush	280	5	5,031	95
Manual flush	2,724	14	16,509	86
Stays in pit	1,713	13	11,430	87
Wrap and throw	[illegible]2	17	535	83
Other	760	11	6,304	89

The nutritional level of the child is associated with incidence of diarrhea ($p<.01$). The rate is higher for the most extreme form of malnutrition, but there is little difference among the other categories.

Table 14. All Children 0-4: Diarrhea in the Past 24 Hours By Nutritional Status, 1977-78

Nutritional Status	DIARRHEA			
	Yes	Percent	No	Percent
Normal	1,728	11	13,845	89
First degree malnourished	2,147	13	14,496	87
Second degree malnourished	1,134	11	8,795	89
Third degree malnourished	638	24	2,953	76

The diarrhea rate among children living in houses with fully enclosed bathrooms was marginally lower than those living in houses where the bathroom was partially or completely exposed to the elements.

Table 15. All Children 0-4: Diarrhea in the Past 24 Hours By Whether Bathroom Is Fully Enclosed Or Not 1977-78

Type of Bathroom	DIARRHEA			
	Yes	Percent	No	Percent
Total	5,663	12	40,143	88
Enclosed	1,685	10	14,765	90
Not enclosed	3,978	14	25,378	86

Generally, as the distance from the house to the toilet increased, the incidence of diarrhea among the children was higher ($p<.01$).

Table 16. All Children 0-4: Diarrhea in the Past 24 Hours By Distance Between House and Toilet, 1977-78

Distance (meters)	DIARRHEA			
	Yes	Percent	No	Percent
Inside house	1,428	9	13,929	91
1-4	159	6	2,416	94
5-9	680	12	5,221	88
10-14	876	17	4,333	83
15-24	649	12	4,853	88
25-49	1,116	19	4,697	81
50-99	309	14	1,827	86
100 or more	320	17	1,526	83
Total	5,538	12	38,803	88

Children whose hands were washed more frequently after they played outside had a lower diarrhea rate as a group. This association was only marginally significant.

Table 17. All Children 0-4: Diarrhea in the Past 24 Hours By Frequency of Washing Childrens' Hands After Playing Outside, 1977-78

Frequency of Handwashing	DIARRHEA			
	Yes	Percent	No	Percent
Total	5,755	13	39,547	87
Always or Sometimes	5,231	12	37,633	88
Rarely or Never	524	21	1,914	79

The frequency of handwashing after using the toilet is associated with incidence of diarrhea among children ($p<.05$). Fewer children had diarrhea in households reporting that they washed "always" or "sometimes" than in households that washed "rarely" or "never".

Table 18. All Children 0-4: Diarrhea in the Past 24 Hours By Frequency of Washing Hands After Using the Toilet, 1977-78

Frequency of Handwashing	DIARRHEA			
	Yes	Percent	No	Percent
Always or sometimes	5,099	12	37,396	88
Rarely or never	742	20	3,031	80

A lower percentage of children living in households where garbage was collected by a service was observed to have diarrhea than in households where garbage was collected by some other method ($p<.05$).

Table 19. All Children 0-4: Diarrhea in the Past 24 Hours By Type of Garbage Disposal, 1977-78

Type of Garbage Disposal	DIARRHEA			
	Yes	Percent	No	Percent
Total	5,841	13	40,413	87
Collected	948	9	10,117	91
Other	4,893	14	30,295	86

A number of other variables were tested for a significant relationship with diarrhea, but were found not to be significant. They include: whether the sample of water from the household was classified as pure, whether the household had piped water as the principal source of water, boiling of water, income, whether the child is currently breastfed, frequency of baths, type and distance of bath, whether the bath is shared, disposal of bath water, child's washing of hands before eating, handwashing by food preparer, cleaning of table before eating, and washing of fruits and vegetables.

2. Bivariate Analysis - Nutrition

The nutrition variable reported on here is based on comparison of weight to age against standards developed for the Philippines. Alternative definitions of nutrition in the literature have been based on comparison of weight to height or height to age. The restriction of the analysis here to the weight to age comparison represents a compromise in the interest of timeliness. The question of which standard is most appropriate will merit further study in the final report.

The analysis here will examine a limited number of variables: city, age, income, whether the child is currently breastfed, water purity and source, and method of excreta disposal. The inclusion of the last variable is motivated by the significant role that it played in the analysis for diarrhea.

The strong effect of city on the levels of the nutrition variable was shown earlier. The overall nutritional level was considerably higher in Cagayan de Oro than Bacolod according to the survey data.

There is a high association ($p<.0001$) between the age of a child and the incidence of diarrhea. Children under age 1 had the highest rate for normal level of nutrition, whereas the highest rates for second and third degree malnutrition occur for age 1.

Table 20. Percent of All Children 0-4: Nutritional Distribution By Age of Child, 1977-78

Age (Years)	NUTRITION			
	Normal	First Degree	Second Degree	Third Degree
0	56	30	9	4
1	29	31	33	7
2	25	43	26	6
3	30	42	23	6
4	38	45	12	5

There is a high association ($p<.0001$) between household income and nutritional status. Families with high income are clearly more able to nourish their children.

Table 21. Percent of All Children 0-4: Nutritional Distribution By Income, 1977-78

Income	NUTRITION			
	Normal	First Degree	Second Degree	Third Degree
0-4,000	30	36	27	7
4,001-8,000	32	37	26	6
8,001-15,000	41	41	15	4
15,001 or more	51	37	9	3

There is a high association ($p<.0001$) between current breastfeeding and nutritional status. Currently breastfed children appear more frequently in the normal category.

Table 22. Percent of All Children 0-4: Nutritional Distribution By Current Breastfeeding Status, 1977-78

Currently Breastfed	NUTRITION			
	Normal	First Degree	Second Degree	Third Degree
Yes	47	28	19	7
No	33	40	22	5

There is a clear association ($p<.0001$) between nutrition and the method of excreta disposal. Families with automatic flush toilets have children with considerably higher levels of nutrition. Those with manual flush toilets are next best, followed by families disposing of waste in the pit under their toilet facility and those using other options.

Table 23. Percent of All Children 0-4: Nutrition By Method of Human Waste Disposal, 1977-78

Method of Disposal	NUTRITION			
	Normal	First Degree	Second Degree	Third Degree
Automatic flush	54	32	10	4
Manual flush	39	41	16	5
Stays in pit	32	38	25	5
Other/wrap & throw	25	30	36	9

There is a clear association ($P<.0001$) between nutrition and the principle source of water: 50 percent of children in households with piped water fall in the normal category, while only 33 percent of the remaining children are classified as normal. There was no significant relationship between nutrition and the purity of the water sample.

3. Multivariate Analysis - Diarrhea

a. Breastfeeding

As a first illustration of multivariate analysis of factors associated with diarrhea, consider the implications of the following two facts:

i. Current breastfeeding is strongly related to a child's age; most breastfed children are under 2 years old.

ii. Children under 2 have high incidence of diarrhea, especially 1 year olds.

These two facts, working together, could make breastfeeding appear to have no or a negative effect on the health variable, whereas actually breastfeeding, taking into account age, could have a beneficial effect.

The preceding hypothesis may be examined by considering the complete cross-classification of diarrheal status by age by current breastfeeding status. CPLX may be asked to consider two models: one in which the rate of diarrhea is associated with age but not current breastfeeding status, and a second which allows for separate effects of both variables on diarrhea. If the second model had fitted the data significantly better than the first, then this would be evidence of a significant relation between the health variable and breastfeeding, once age is controlled. Furthermore, CPLX could estimate the parameters of the second model, which would show the direction of association between diarrheal status and breastfeeding, controlling for age.

In point of fact, the second model did not fit significantly better than the first; thus, the multivariate analysis came to the same negative finding as the bivariate analysis. Breastfeeding did not have an effect on the incidence of diarrhea.

b. Income

The possible effect of income was examined in a similar manner. Just as for the variable "currently breastfed", income had not shown a statistically significant relationship with diarrhea status in the bivariate analysis. This negative result is contrary to the typically strong effects that income has on other

variables in many other studies, and is also surprising given the observed relationships between income and variables in this study that do show statistically significant associations with diarrhea, such as nutrition and method of waste disposal. Again, multivariate analysis was employed to assess whether this apparent lack of relationship on the bivariate level was a consequence of more complex relationships with other variables.

The following observations appeared to offer an explanation for the bivariate outcome:

(1) Average household income was measured to be higher in Cagayan de Oro than Bacolod.

(2) Diarrhea rates were higher in Cagayan de Oro than Bacolod.

These two factors could have worked to suppress an underlying positive relationship between income and health status. When this hypothesis was examined by means of a three-way table (diarrhea status by income by city), however, income failed to show a statistically significant association with diarrheal status when city was controlled.

For the sake of completeness, additional models incorporating factors of (human) waste disposal and nutrition were examined. Again, income failed to make an independent contribution to the model. Consequently, just as for breastfeeding, the multivariate analysis did not substantially alter the conclusion of the bivariate analysis. The household's income level did not affect the incidence of diarrhea.

c. City

As noted earlier, the observed rates of diarrhea differed substantially between the two cities in this study. One interpretation would be to ascribe the difference simply to seasonal differences due to the interviewing periods for the two cities, but the difference also could reflect socioeconomic and other ecological differences between the two.

Besides the overall comparison of the two cities, it is of interest to assess the extent to which the impact of other variables on the rate of diarrhea is affected by city. To measure this association or "interaction", the three-way cross-classification of diarrhea by the explanatory variable by city must be considered. Five variables related to diarrhea in the bivariate analysis were evaluated in this manner: age, method of human waste disposal, distance to the toilet, handwashing after using the toilet, and nutrition. (Other variables mentioned earlier, type of bathroom and garbage disposal, were dropped from further analysis on account of their marginal levels of significance.) Of the five variables, three showed

no interaction with city in their effect on diarrhea: distance to toilet, handwashing after the toilet, and nutrition. The lack of significant interaction on these three variables is a finding of substantive interest. For example, the lack of interaction between nutrition and city in their effect on diarrhea is evidence that very low nutrition (in terms of weight per age) is associated with higher rates of diarrhea for each of the two cities to approximately the same extent.

The remaining two variables studied, age and method of human waste disposal, showed statistically significant interactions with city. In both cities, the highest rate of diarrhea were for 1 year olds, but for ages 3 and 4, an anomoly appeared where the diarrheal rates were higher for age 3 than age 4 in Cagayan de Oro (thus, continuing the general decline in rates after age one), but the rates were curiously lower for age 3 than age 4 in Bacolod. This reversal was the principle source of the statistically significant interaction between the effects of age and city.

Method of waste disposal also interacted with city in the effect on diarrhea rates. In both cities, rates were lower for households with automatic flush toilets than for households with manual flush toilets. In Bacolod, however, the differential is significantly more pronounced than in Cagayan de Oro. As with the basic differences between cities mentioned earlier, this interaction may be due to seasonal differences. (Quite possibly, during a period of generally low diarrheal incidence, the association of diarrhea with the sanitation variable may be quite strong, but during a period of generally high incidence, the association may be more blurred.) Again, however, the interaction may be a consequence of ecological differences between the cities.

d. Distance to Toilet, Washing of Hands After Using Toilet, Method of Human Waste Disposal

Three variables associated with the type and use of the toilet are all significantly related to diarrhea in the bivariate analysis. The multivariate question here addresses whether each of these related variables has a direct effect on diarrheal rates distinct from the contribution of the other two.

This question was approached by first examining the relationship of distance to the toilet and method of human waste disposal. These two variables are highly related in the survey results, in particular, virtually all reported automatic flush toilets were within the respondents' houses. To examine whether each of these two variables had a separate effect on the diarrheal rate, the cross-classification of diarrheal status by distance to toilet by method of human waste disposal was considered. Because the effect of human waste disposal was found to interact with city, the cross-classification also included city as the fourth variable.

A number of different models for this table were evaluated using CPLX. When no relationship between diarrhea and method of disposal was incorporated in the model, distance to toilet was found to have a significant effect; with the variables reversed, method of disposal was found to be significant when the effect of distance to the toilet was ignored. Of course, these outcomes simply replicate the earlier results of the bivariate analysis, since the effect of only one variable is considered in the models.

When the main effect of method of disposal and its interaction with city are included in the model, distance to the toilet no longer contributes significantly to the model. The opposite is not true, however, when either the main effect of distance to the toilet, or the main effect and interaction with city are included, method of disposal and its interaction with city continue to contribute significantly o the model. Thus, in the first case, the method of disposal appears to "explain" tl bivariate effect of distance to toilet, while in the second case distance to toilet fails to explain fully the contribution of the method of disposal.

These findings serve to identify method of disposal as more directly related to diarrheal rates than distance to the toilet. Although it is far too much of a simplification to say that method of disposal is the cause of diarrhea, the inference from the preceding models is nonetheless causal in nature: that distance to toilet appears to have no statistically significant effect on diarrhea that is not explained by the method of waste disposal and city.

A similar analysis was conducted to assess the relative contributions of handwashing after use of toilet and method of human waste disposal. The outcome of the modeling parallels the preceding analysis: method of disposal, and its interaction with city, but not handwashing, or its interaction with city, continue to contribute significantly to the model when the other variable was considered. Hence, the three significant bivariate associations appear to be all a consequence of a single underlying relationship between diarrhea and the method of human waste disposal.

4. Multivariate Analysis - Nutrition

a. Interactions With City

As with the analysis of diarrhea, the effect of other variables on nutrition was examined for variation by city. Remarkably, two variables with extremely significant effects on nutrition did not show any interaction effect with city: income, and method of excreta disposal. Thus, within sampling error, high income and automatic flush toilet, appear to be positively associated with nutritional level in the same manner in Cagayan de Oro and Bacolod.

Age does interact with city ($p<.01$) in the effect on diarrhea. The somewhat different patterns in the variation of nutrition by age is shown in the following table.

Table 24. Percent of All Children 0-4: Nutritional Distribution By Age of Child and City, 1977-78

Age (Years)	NUTRITION			
	Normal	First Degree	Second Degree	Third Degree
	Cagayan de Oro			
0	63	27	6	4
1	32	42	22	4
2	31	48	19	2
3	40	46	11	3
4	40	46	12	2
	Bacolod			
0	52	32	12	4
1	27	22	43	9
2	12	39	30	9
3	23	39	32	7
4	36	44	13	8

b. Breastfeeding

The bivariate analysis indicated an extremely strong and beneficial relationship between breastfeeding and nutrition. The bivariate tables for age, however, show that nutrition is highest, according to the weight standards used in this study, for children under age 1, the same children who are most likely to be breastfed. A question immediately arises, therefore, as to whether the apparent association between nutrition and breastfeeding derives simply from their common dependence on age.

Parallel analyses for two different age groupings were performed in order to test this hypothesis. One grouping was confined to the first year of age, split into six 2-month intervals. The second was based on three half-year intervals starting at 6 months

and ending at 2 years. In each case, for the three-way classification of nutrition by age by breastfeeding, the last variable failed to show a statistically significant relationship with nutrition, once the relationship between nutrition and age was considered.

The lack of statistically significant relationship between nutrition and breastfeeding, once age is taken into account, may be considered counter to other literature on this subject. In order to test the validity of this finding, an even more complex model was examined, which could have offered an explanation for the observed results.

Particularly after 4 or 6 months, the choice of whether to breastfeed the infant depends quite strongly on the household income of the mother; high-income mothers are more likely not to breastfeed. The previous bivariate tables also showed that high income was generally associated with high levels of nutrition. Thus, the three-way table relating nutrition to age and breastfeeding may not correctly represent the hidden effect of income.

This second, more complex, hypothesis was analyzed for two separate age groupings: the first half-year of age, split into three 2-month intervals; and the next year and one-half, split into three half-year intervals. Both age and income were allowed to affect nutrition in the model, but the separate contribution of breastfeeding to nutrition was examined. Again, however, breastfeeding showed no statistically significant relationship with nutrition, once the other variables were taken into account.

The reader should refrain from superficial generalizations of these findings beyond the context of this project. Nonetheless, the results indicate that any beneficial effects of breastfeeding do not express themselves in terms of the child's weight to a degree sufficient to yield a statistically significant relationship from the survey results. The negative outcome holds equally when only age is considered or when effects of both age and household income are taken into account. Should the follow-up survey replicate this finding, then the variable breastfeeding can be expected to play little or no role in the final analysis.

c. Method of Excreta Disposal

The bivariate analysis indicated an extremely strong association between nutrition and method of excreta disposal. Since the choice of method is highly related to income and income has a strong effect on nutrition, an appropriate question is whether the observed relationship between nutrition and method of disposal is due to particular properties of the latter variable as an expression for sanitary practices or is the result of their common association with income. In other words, does method of

disposal continue to have a significant effect on nutrition, once income is considered?

A preliminary answer to this question comes from consideration of the cross tabulation of nutrition by income disposal by age by city. A "main effects" model fits the table quite well, i.e., income, method of excreta disposal, and city each had separate effects on nutrition that combined additively. The contributions of income and method of disposal were each strongly significant ($p<.0001$). The interaction parameters of the model were tested and found not significant, e.g., the level of income made no statistically significant difference in the effect of method of excreta disposal on nutrition (and vice versa).

On occasion, the question of statistical significance can get confused with substantive importance, in particular, there is a danger in overinterpreting the impact of a statistically significant variable that has limited actual effect on the dependent variable. The "beta" parameters of the logistic model provide one such measure of the substantive impact of the independent variables. The parameters of the "main effects" model are given in Table 25.

Table 25. Beta Parameters of the Logistic Regression of Nutrition on City, Method of Excreta Disposal, and Income (Standard Errors in Parentheses)

Variable	LEVELS OF NUTRITION			
	Normal	1st Degree	2nd Degree	3rd Degree
City				
Cagayan de Oro	.35 (.12)	.36 (.09)	-.24 (11)	-.47 (.17)
Bacolod	-.35 (.12)	-.36 (.09)	.24 (11)	.47 (.17)
Method of Excreta Disposal				
Auto flush	.58 (.28)	-.24 (.12)	-.33 (.38)	-.01 (.40)
Manual flush	.18 (.17)	.33 (.14)	-.41 (.26)	-.10 (.25)
Stays in pit	-.07 (.17)	.23 (.17)	.17 (.16)	-.34 (.31)
Other	-.68 (.16)	-.33 (.19)	.56 (.27)	.45 (.31)
Income				
1 - 4,000	-.41 (.16)	-.24 (.14)	.35 (.22)	.30 (.32)
4,001 - 8,000	-.40 (.16)	-.24 (.13)	.51 (.19)	.14 (.25)
8,001 -15,000	.14 (.19)	.06 (.21)	-.11 (.25)	-.10 (.38)
15,001+	.67 (.20)	.41 (.25)	-.74 (.30)	-.34 (.45)

As a rough indication of the interpretation of these parameters, a coefficient of 0.6 converts a proportion at 0.50 to 0.65, while -0.6 converts the same 0.50 to 0.35. Similarly, 0.4 changes 0.50 to 0.60 (-0.4 gives 0.40), and 0.2 changes 0.50 to 0.55 (-0.2 gives 0.45). At levels other than 0.50, the effect is not symmetric: a 0.6 converts a proportion of 0.10 to 0.17, while -0.6 converts 0.10 to 0.06. Similarly, 0.4 gives 0.14 (-0.4 gives 0.07) and 0.2 gives 0.12 (-0.2 gives 0.08). Thus, a parameter of 0.2 indicates an interesting and important effect, while 0.6 may be said to denote an extremely strong difference. Thus, the parameters presented in the table denote clearly substantial effects for city, method of excreta disposal, and income. In particular, even when income is controlled, method of excreta disposal continues to have a statistically significant and substantively important association with nutrition.

Income by itself does not "explain" the effect of method of excreta disposal on nutrition. Quite possible, however, the method of disposal represents additional information on wealth and other aspects of socioeconomic status of the household beyond what information is implied by income. A hypothesis may be considered in which the association between nutrition and method of disposal is interpreted as a reflection more on socioeconomic status not represented by income than on sanitation per se.

As a test of this hypothesis, other variables that could also serve as measures of socioeconomic status were studied. Four variables were examined: type of lighting (electric vs. other), education of the head, classifiction of the house, and the number of household items. All four variables showed highly significant relationships to nutrition.

As a next step, each of the four socioeconomic variables were examined to determine if they still contributed additional explanation in a model of nutritonal status that also included the effect of income. Only one, type of lighting, did so; the remaining three could not add significantly to the explanation given by income.

The five-way cross classification of nutrition by income by city by type of lighting by method of excreta disposal represents the joint effect of all variables. When income and type of lighting are considered, method of disposal continues to make a statistically significant contribution ($p<.02$). Furthermore, the parameters of the model are not substantially altered by the inclusion of the variable, type of lighting, as can be seen in Table 25.

The baseline data thus indicate that method of excreta disposal may be of key importance in the analysis of the project result.

Table 26. Beta Parameters of the Logistic Regression of Nutrition on City, Method of Excreta Disposal, Income, and Type of Lighting (Standard Errors in Parentheses)

Variable	LEVELS OF NUTRITION			
	Normal	1st Degree	2nd Degree	3rd Degree
City				
Cagayan de Oro	.36 (.12)	.38 (.09)	-.25 (.12)	-.49 (.17)
Bacolod	-.36 (.12)	-.38 (.09)	.25 (.12)	.49 (.17)
Method of Excreta Disposal				
Auto flush	.51 (.25)	-.37 (.24)	-.24 (.40)	.10 (.43)
Manual flush	.14 (.17)	.26 (.14)	-.36 (.28)	-.04 (.26)
Stays in pit	-.04 (.16)	.30 (.16)	.13 (.16)	-.40 (.29)
Other	-.62 (.16)	-.19 (.20)	.48 (.32)	.33 (.39)
Income				
1 - 4,000	-.36 (.17)	-.14 (.15)	.29 (.21)	.22 (.35)
4,001 - 8,000	-.38 (.17)	-.19 (.13)	.48 (.20)	.10 (.27)
8,001 -15,000	.13 (.20)	.02 (.22)	-.08 (.24)	-.07 (.39)
15,000+	.62 (.21)	.31 (.26)	-.68 (.32)	-.25 (.48)
Type of Lighting				
Electric	.12 (.10)	.24 (.08)	-.15 (.17)	-.21 (.20)
Other	-.12 (.10)	-.24 (.08)	.15 (.17)	.21 (.20)

I. Sanitary Practices

As mentioned in the previous section, the variables that were found to be significantly associated with diarrhea in children 0-4 years old were:

- frequency of handwashing after playing
- frequency of handwashing after toilet use
- method of excreta disposal
- method of garbage disposal
- type of bathroom
- distance from house to toilet

This section will investigate whether differences exist among households between strata with respect to the above-mentioned indicators.

1. Handwashing After Playing Outside

In Cagayan de Oro, 91 percent of stratum I households reported that the household members washed hands after playing outside either always, most of the time, or some of the time, while stratum II reported 89 percent and stratum III, 87 percent.

In Bacolod, households members in about 1 percent of stratum I households, 3 percent of stratum II households, and 1 percent of stratum III households rarely or never washed hands after playing outside.

(See Table 27. for percentage distribution of households on frequency of washing hands after playing outside.)

Table 27. Percent of Households: Frequency of Washing Hands After Playing Outside By City and Experimental Stratum, 1977-78

Frequency of Washing Hands After Playing Outside	Cagayan de Oro Experimental Stratum			Bacolod Experimental Stratum		
	I	II	III	I	II	III
All households	100	100	100	100	100	100
Always, most of the time, some of the time	91	89	87	99	97	99
Rarely, Never	9	11	13	1	3	1

2. Handwashing After Toilet Use

Eighty-six percent of the Cagayan de Oro households and 73 percent of the Bacolod households reported that household members washed their hands "always" or "most of the time". In Cagayan de Oro, the proportion of households in stratum I reporting in the "always" category was significantly higher than in stratum II but was similar to stratum III. In Bacolod, the proportion of households reporting in the "always" category in stratum I was marginally higher than in both stratum II and stratum III. (See Table 28.)

Table 28. Percent of Households: Frequency of Handwashing After Toilet Use, 1977-78

Frequency of Washing Hands After Toilet Use	Cagayan de Oro Experimental Stratum I	II	III	Bacolod Experimental Stratum I	II	III
All households	100	100	100	100	100	100
Always	80	75	71	56	45	43
Most of the time	9	10	12	24	24	21
Some of the time	6	10	10	19	22	27
Rarely	2	3	5	2	5	3
Never	1	3	1	-	2	5
Don't know	0.5	-	1	-	1	1

3. Method of Excreta Disposal

Seventy-three percent of the Cagayan de Oro households and 55 percent of the Bacolod households used either an automatic or manual flush toilet. In Cagayan de Oro, more households in stratum II used the automatic flush type than in stratum I. However, a significantly lower percentage of stratum III households used this method than in stratum I. In the same city significantly more of the stratum I households used manual flush toilets than stratum II households. However, no significant differences could be observed between strata in Cagayan de Oro in terms of "stays in pit" and "wrap and throw" methods of excreta disposal.

In Bacolod, more households in stratum I used the manual flush toilet than in either strata II or III, but no strata variation could be established with respect to manual flush, "stays in pit", or "wrap and throw" excreta disposals. (See Table 29 for percentage distribution by method of excreta disposal.)

Table 29. Percent of Households: Method of Excreta Disposal By City and Experimental Stratum, 1977-78

Method of Excreta Disposal	Cagayan de Oro Experimental Stratum			Bacolod Experimental Stratum		
	I	II	III	I	II	III
All households	100	100	100	100	100	100
Automatic flush	17	24	9	14	14	11
Manual flush	64	42	66	60	35	24
Stays in pit	17	23	24	21	29	34
Wrap and throw	0.5	z	-	1	2	1
Other	1	11	1	4	20	30

4. Method of Garbage Disposal

Households in the two cities disposed of their garbage in a variety of ways. Sixty-three percent of the households either burned the garbage or had it collected by a garbage service. It was evident that more households in stratum I had their garbage collected by a service when compared to either strata II or III. (See Table 30.)

Table 30. Percent of Households: Method of Garbage Disposal By City and Experimental Stratum, 1977-78

Method of Garbage Disposal	Cagayan de Oro Experimental Stratum I	II	III	Bacolod Experimental Stratum I	II	III
All households	100	100	100	100	100	100
Collected by service	20	10	3	67	27	3
Burning	41	49	57	16	29	60
Compost	7	7	10	9	15	24
Dump outside yard	17	20	19	2	3	1
Dump pit	3	2	2	3	11	7
Dump inside yard	2	3	1	-	2	-
Other	9	9	8	5	14	5

5. Type of Bathroom

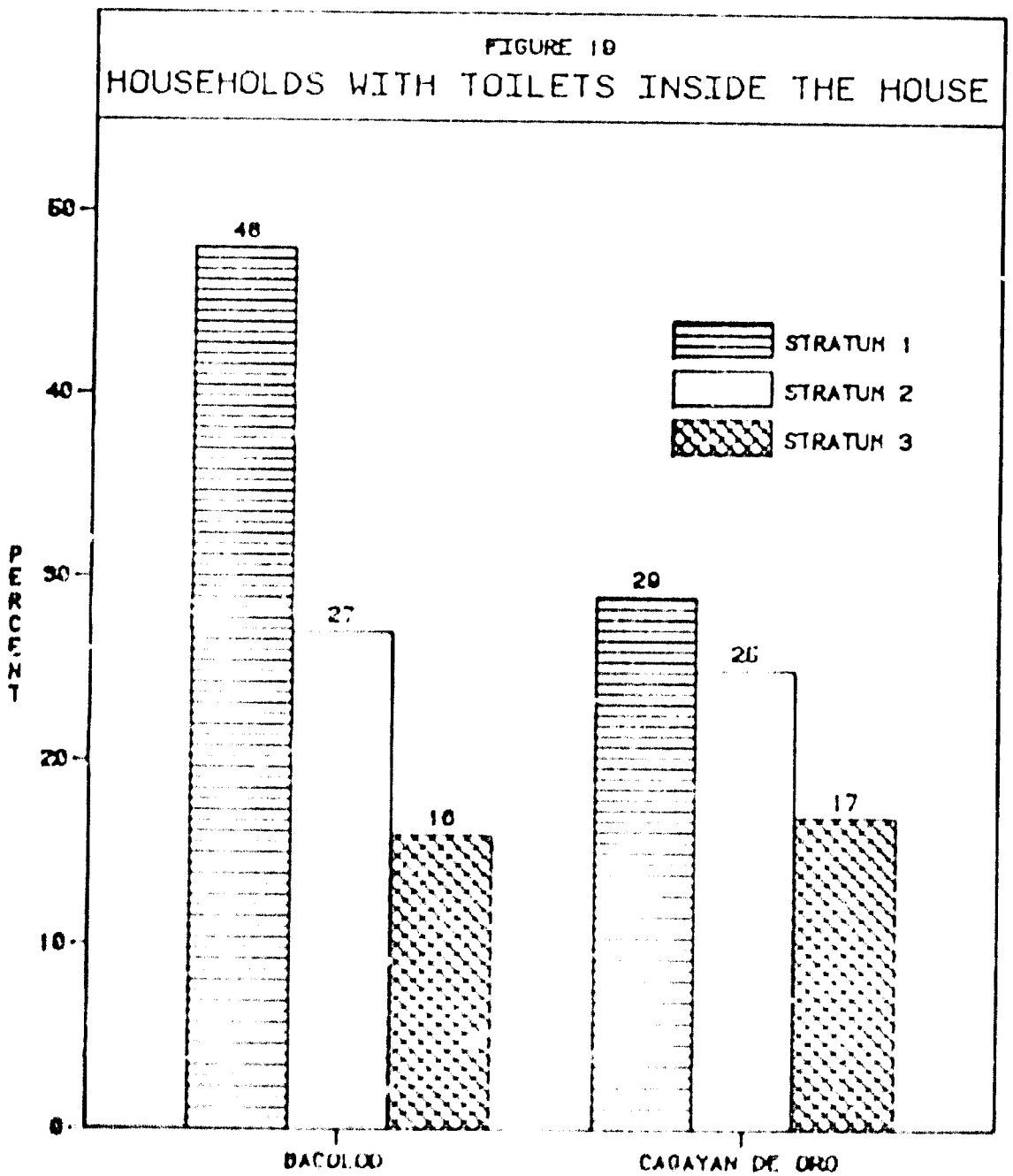
Bathrooms in both cities were either enclosed (4 walls), partially sheltered, or open. In Cagayan de Oro, more stratum I households had enclosed bathrooms than stratum III, whereas in Bacolod, the proportion of households in stratum I having enclosed bathrooms was significantly higher than either stratum II or III. (See Table 31.)

Table 31. Percent of Households: Type of Bathroom By City and Experimental Stratum, 1977-78

Type of Bathroom	Cagayan de Oro Experimental Stratum I	II	III	Bacolod Experimental Stratum I	II	III
All households	100	100	100	100	100	100
Enclosed (4 walls)	48	43	33	60	37	25
Partially sheltered	11	9	7	17	20	10
Open	41	49	60	23	43	65

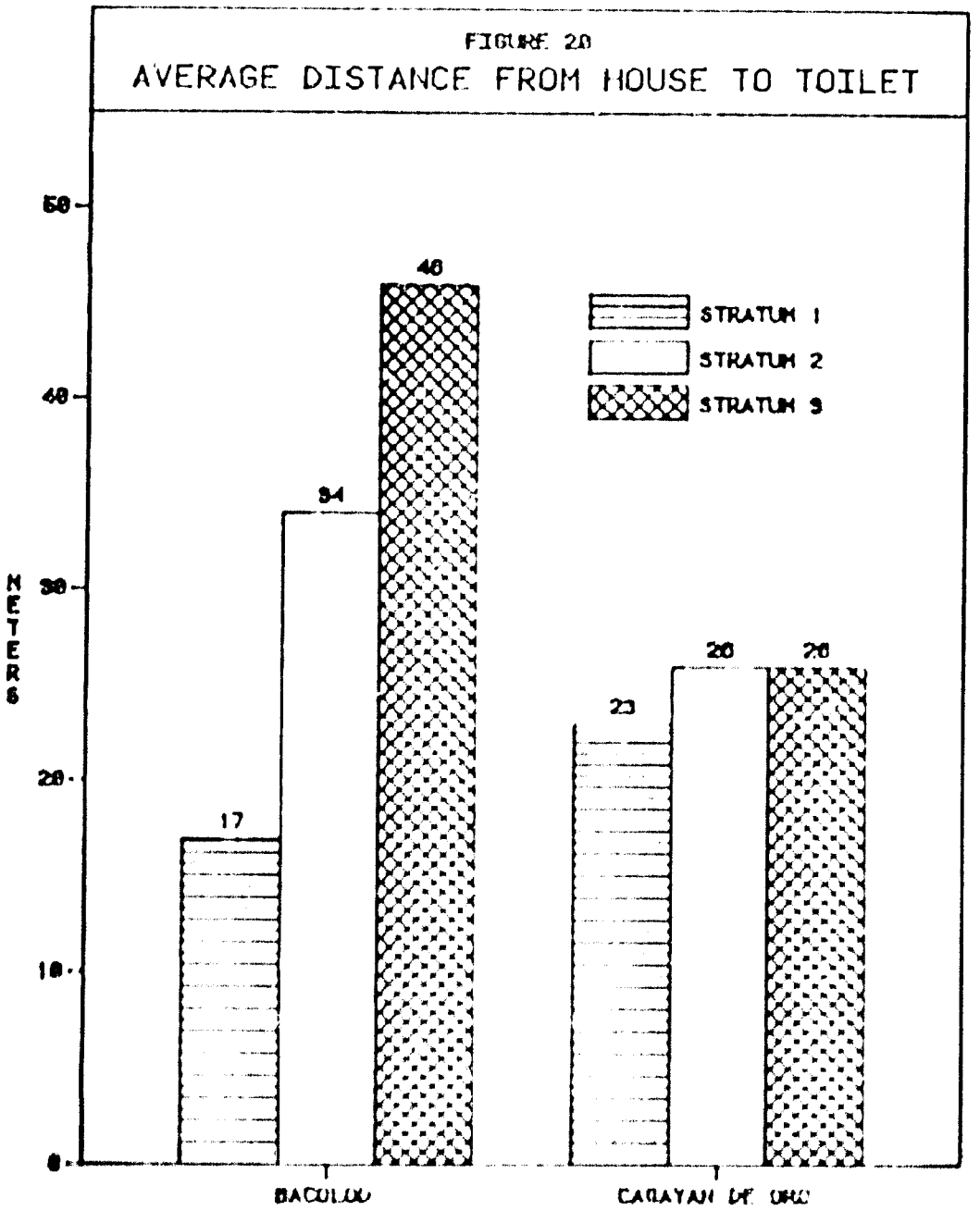
6. Distance From House to Toilet

A significantly higher percentage of stratum I households in Bacolod had toilets inside their houses than in either stratum II or III. In Cagayan de Oro the percentage in stratum I was significantly higher than stratum III only. (See Figure 19.)



1977-78

The average distance from the house to the toilet for both cities was 29 meters. In Cagayan de Oro, statum I households had an average distance of 23 meters, while the distances for stratum III households was 26 meters. The data for Bacolod were 17 meters for statum I and 46 meters for stratum III. (See Figure 20.)

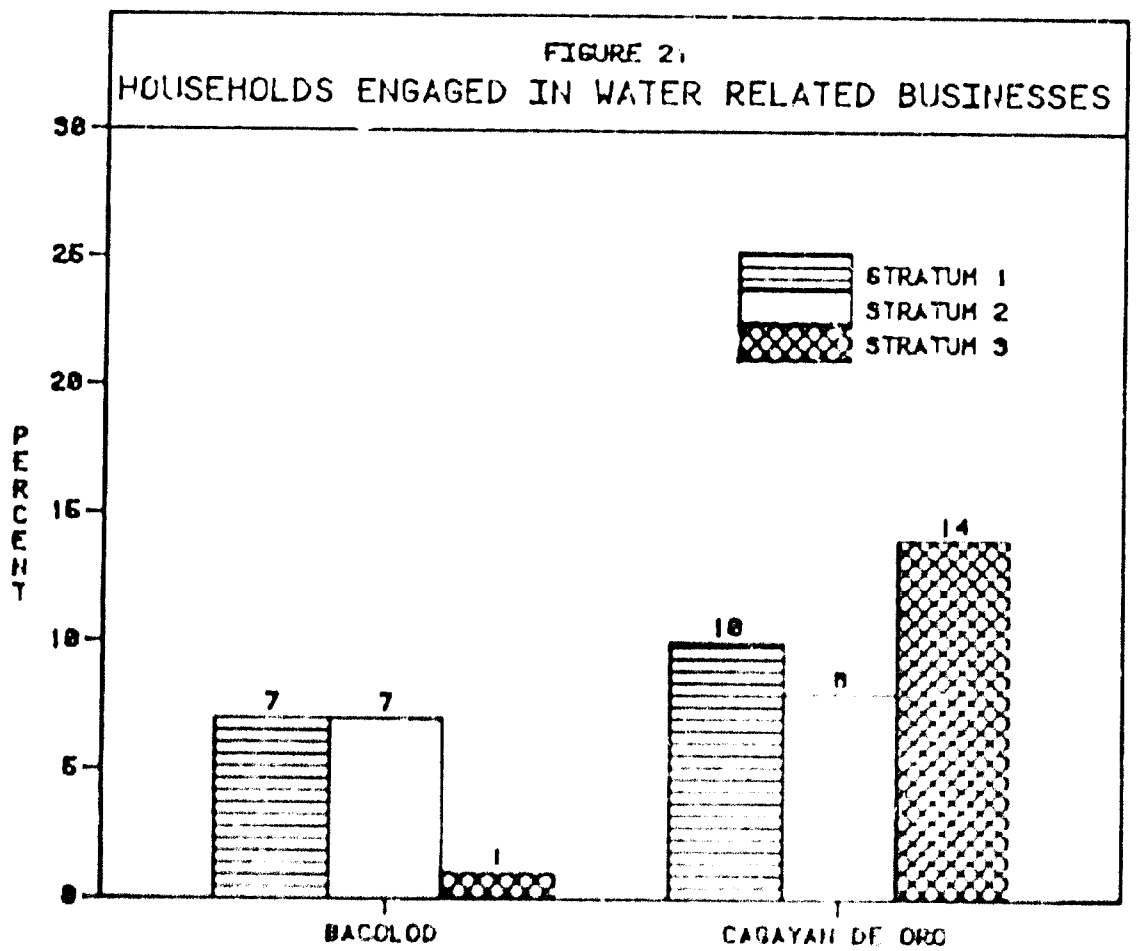


1977-78

J. Economic Uses of Water

1. Water-related Business Activities

The baseline surveys estimated that only 4,640 households (8 percent of the households) were engaged in water-related business activities in the two cities. This included 14 percent of the stratum III households in Cagayan de Oro, and only 1 percent of the stratum III households in Bacolod. (See Figure 21.)



1977-78

The most frequently reported kinds of water-related businesses were "selling water" and "pig raising". Nine hundred sixty-seven households in Cagayan de Oro (21 percent of all households with water-related businesses) sold water as their water-related business. In Bacolod, 395 households or 9 percent had pig-raising businesses.

On the average, households engaged in water-related businesses earned P137 ($18.15) per week from the business. (See Table 37.)

Table 32. Number of Households Engaged in and Average Income Earned in Previous Week (In Pesos) for Each Water-related Business: Type of Business

Type of Business	Cagayan de Oro	Bacolod
All businesses		
Number of households	2,379	2,263
Average income	153	119
Food processing		
Number of households	458	219
Average income	136	40
Water selling		
Number of households	967	62
Average income	95	4
Laundering		
Number of households	143	275
Average income	131	11
Vegetable gardening		
Number of households	94	233
Average income	58	52
Poultry raising		
Number of households	122	130
Average income	388	Not available
Pig raising		
Number of households	395	619
Average income	97	205
Restaurant		
Number of households	197	419
Average income	272	185
Other businesses		
Number of households	249	299
Average income	104	58

Although most of the households that engaged in water-related businesses did so to supplement their incomes, a few actually obtained 75 to 100 percent of their household income from the business. (See Table 33.)

Table 33. Percent of Households: Percent of Total Income Earned From Water-related Business By Experimental Stratum, 1977-78

Percent of Total Income Earned From Water-related Businesses	Total	Experimental Stratum		
		I	II	III
0	93	92	94	98
1-24	3	4	3	1
25-49	1	1	1	-
50-74	1	1	1	-
75-100	2	2	2	1

2. Perceived Adequacy of Water Supply for Business Activities

Fifty-three percent of the Cagayan de Oro households and 81 percent of the Bacolod households felt that their water supply was "very adequate" or "adequate" for business activities. The remaining households reported that the supply was not adequate. However, only 15 percent of these households said that they desired to engage in a water-related business even if the supply would become adequate.

VI. RELIABILITY OF THE DATA

The estimates developed from the sample are apt to differ somewhat from the results of a survey covering all households in the sampled cities but otherwise conducted under essentially the same conditions as the actual survey. The sampling errors - the differences between the estimates obtained and the results theoretically obtainable from a comparable complete coverage survey are unknown. Guides to the potential size of the sampling errors, however, are provided by the standard errors of the estimates.

The particular sample selected for the Philippines Provincial Water Survey is one of a large number of similar probability samples, that, by chance, might have been selected under the same specifications. Each of the possible samples would yield somewhat different sets of results and the standard errors are measures of the variation of all the possible sample estimates around the theoretical, comparable complete-coverage values.

Estimates of the standard errors have been computed from the sample data for selected statistics in this report. They were calculated in the form of relative standard errors (i.e., coefficient of variation), the standard errors divided by the estimated values to which they refer.

In conjunction with its associated estimate, the relative standard error may be used to define confidence intervals, ranges that would include the comparable complete-coverage value for specified percentages of all possible samples. The complete-coverage value would be included in the range:

i. from one standard error below to one standard error above the derived estimate for about two-thirds of all possible samples.

ii. from two standard errors below to two standard errors above the derived estimate for about 19 out of 20 of all possible samples.

iii. from three standard errors below to three standard errors above the derived estimate for nearly all samples.

An inference that the complete-coverage result would be within the indicated ranges would be correct in approximately the relative frequencies shown. Those proportions, therefore, may be interpreted as defining approximate probabilities that the estimates shown would differ from complete coverage results by as much as one, two, or three standard errors, respectively.

For example, if an estimated total is shown as 50,000 with an associated relative standard error of 2 percent, then the standard error is 1,000 (i.e., 2 percent of 50,000). The interval from 49,000 to 51,000 represents approximately a two-thirds confidence interval, i.e., in about two-thirds of the samples, the interval formed by adding and subtracting 1,000 from the sample estimate would contain the complete coverage total. Similarly, adding and subtracting 2,000 from the sample estimate (in this case 48,000 to 52,000) forms intervals that will cover the complete coverage total in about 95 percent of the possible samples. Adding and subtracting 3,000 forms on intervals that will cover the complete coverage total for almost all samples.

In addition to the sampling errors, the estimates are subject to various response and operational errors: of collection, response, coding, transcription, nonresponse, etc. These errors of response and operations would also occur if a complete canvass were to be conducted under the same conditions as the survey. Explicit measures of their effects generally are not available.

It is believed that many of the important response and operational errors were detected and corrected in the course of the review of the data for reasonableness and consistency. However, there is some indication that interviewer bias may have been a problem for some survey items, particularly in Cagayan de Oro. This became evident during the correction phase of the edit for consistency. Many of the same kinds of errors that were discovered were found to have occurred on questionnaires completed by the same interviewer(s). Although most were corrected, this type of pattern usually implies that other errors occurred due to these interviewers' sloppiness or insufficient training. An example of an error that might [illegible] is an entry of 2,000 pesos for "income from fishing, farming, the past month" on a series of questionnaires completed by the same interviewer. Since it is extremely unlikely that several different households would report the same figure for this component of income, this "error" could be due to the interviewer leading the respondent, making an entry by his/her own determination, or using his/her own series of questions rather than those on the questionnaire. This error would not be detected since "2,000" is a "legal" and reasonable figure.

As in all surveys, nonresponse was present. For this particular survey the nonresponse was item nonresponse, rather than total nonresponse since there were no complete nonresponses. For many characteristics the nonresponse rate was less than 10 percent. However, for certain characteristics the nonresponse rate was much higher. (See Appendix.) For this survey, the estimate of a characteristic was made only from those sampled households which responded. Thus, estimates may be biased. The size of the bias will depend on how alike or different the respondent and nonrespondents are. The more alike they are for a given characteristic the smaller the bias will be.

Appendix Characteristics With Nonresponse Rates Above 10 Percent

*1. Number of Household Items	Cagayan de Oro	52.6
	Stratum I	50.0
	Stratum II	52.1
	Stratum III	49.5
2. Diarrhea Rate	Cagayan de Oro	26
	Stratum I	29
	Stratum II	23
	Stratum III	20
3. Nutrition by Number Breastfeeding	Both Cities,	
Age 0-1	Across All	19
Age 0-4	Strata	20

*If the nonresponse rate is calculated allowing each household to fail to respond to one household item (there were 16 listed on the questionnaire) then the rate is below 10 percent.