PN-AAL-216

931 1055

Report of a Study by the
Berkeley Project on Education and Nutrition

Analysis of Longitudinal Data Collected in an Experimental Nutritional Supplementation Program:

An examination of the INCAP four-village study to investigate relationships between nutrition, health, family background, children's activities and school attendance and performance

Concise Title:

EDUCATION & NUTRITION: PERFORMANCE AND POLICY

Charles S. Benson, Co-Principal Investigator
Sheldon Margen, Co-Principal Investigator
Judith B. Balderston, Project Director
Maria E. Freire
Mari S. Simonen
Alan B. Wilson

April 1980

University of California
Berkeley

Supported by USAID
Contract # AID/DSPE-C-0021

Report of a Study by the
Berkeley Project on Education and Nutrition

Analysis of Longitudinal Data Collected in an Experimental Nutritional Supplementation Program:

An examination of the INCAP four-village study to investigate relationships between nutrition, health, family background, children's activities and school attendance and performance

Concise Title:

EDUCATION & NUTRITION: PERFORMANCE AND POLICY

Charles S. Benson, Co-Principal Investigator
Sheldon Margen, Co-Principal Investigator
Judith B. Balderston, Project Director
Maria E. Freire
Mari S. Simonen
Alan B. Wilson

April 1980

University of California
Berkeley

Supported by USAID
Contract # AID/DSPE-C-0021

TABLE OF CONTENTS

	Page #
INTRODUCTION	1
SCOPE OF THE STUDY	3
STRUCTURE OF THE REPORT	6
PART I. BACKGROUND	8
A. The Guatemalan Longitudinal Study: INCAP's Experiment and Findings -- by Sheldon Margen	8
B. A Brief Description of the Villages	36
PART II. ANALYTICAL RESULTS FROM THE BERKELEY STUDY	51
A. Longitudinal Analysis of Nutrition, Growth, Verbal Development, and School Achievement -- by Alan B. Wilson	54
B. School Enrollment and Achievement: Effects of Nutrition, Health, and Need for Children's Work -- by Judith B. Balderston	211
C. Education and Agriculture Efficiency -- by Maria E. Freire	251
D. Education and Fertility in Rural and Semi-Urban Guatemala -- by Mari S. Simonen	300
PART III. SYNTHESIS OF FINDINGS AND POLICY IMPLICATIONS by Judith B. Balderston	349
Appendix I	373

ACKNOWLEDGEMENTS

Our greatest debt is owed to Dr. Robert Klein and the members of his staff at the Institute of Nutrition of Central America and Panama without whose ten years of data collection this study could not have taken place. Because of their enormous efforts, we were able to carry out the work reported here. We also thank AID for their financial backing of our project and NICHD for their support of the original INCAP study. We owe special thanks to the Rockefeller Foundation for their role in support of Rand actvities which were extremely important in our work. We are very grateful to Dr. William Butz and members of the staff at Rand for their assistance in making the data collected by Rand accessible to us.

We thank the staff at AID and the members of the research review committee, Dr. Sam Wishik, chairman, and Dr. Laurence Lau and Dr. Henry Ricciutti who provided us with beneficial comments.

Several members of the INCAP research group were especially helpful to us. Besides Dr. Klein, we would like to mention Patricia Engle, Marc Irwin, John Townsend, Charles Yarbrough, William Owens, and many other members of the INCAP staff who enabled us to do so much work in a relatively short time. Carol Clark, a post-doctoral fellow of the Rockefeller Foundation, provided essential information and advice at a late stage of the research.

Several members of the Berkeley staff who do not appear as authors participated in important activities. Jane Fraser organized the data base at the outset and did preliminary analyses and writing. Vien Phan and Patsy Fosler did valuable work preparing files for analysis. Jo-Ann Work shared much of the administrative load. Joan Haller and Lynne Reilly energetically kept up with the paper flow of so many authors. We thank them all. The ultimate responsibility is with those of us whose names appear on this report.

INTRODUCTION

The work that is described in this report represents the results of a search, for policy purposes, on the relationships between nutrition and education and the out-of-school performance of young people. The study was undertaken (1) to test hypotheses regarding the relationships between education, health, and nutrition, particularly with respect to the effect of nutrition on school performance and activities of children, and (2) to analyze how choice to use nutritional supplementation, health care, contraception, and education are related to family income, parental literacy, and education. At the time the study was proposed and funded, it was suggested that it would "be useful in predicting the effects of early malnutrition on later stages in the life cycle and how early growth and development are related to success in school and in the performance of economic activities."

To carry out such an investigation requires the collection of data embracing a wide variety of areas -- physiological, nutritional, socio-economic, psychological, and a natural setting environment where the interaction among areas would take place and be observed. It was therefore extremely fortunate that the ambitious and costly longitudinal study by INCAP funded by the National Institute of Child Health and Development for data collected between 1969 and 1978, and by Rand, funded by the Rockefeller Foundation, for data collected in 1974 and 1975, had already occurred in the four villages in Eastern Guatemala. Out of these projects, a large and varied data base had been amassed, and it was therefore possible to make the connections between children's individual growth and development, their families' economic and social conditions, and the environment of the village economies and schools. It was hoped that the results obtained through the use of appropriate analytical models would be useful for the formulation of policies in international agencies, mainly by stressing the importance of inter-sectoral planning in less developed countries.

At the time of the proposal, it was recognized that, although much was already know about effects of severe malnutrition, the result of chronic malnutrition on human functioning deserved much more investigation. Nutritional science, well-developed in the laboratory, had not yet been able to translate known chemical and biological relationships into functional outcomes. The purpose of the ambitious study begun by INCAP in 1969 was to relate some of the nutritional and health inputs into growth and development outcomes. The objective of the Berkeley study was to enable educational and other planners to see how nutritional change might affect educational outcomes.

Integrating the plans made by organizations in separate fields, such as nutrition, public health, rural development, and education would potentially improve the cost-effectiveness of their programs. To do such integration requires knowledge of relationships between the separate sectors, and it was suggested that data from the INCAP and Rand studies in the four villages in Guatemala were capable of providing many important research results.

This report represents the findings of the Berkeley Project on Education and Nutrition, financed by USAID under contract #DSPE-C-0021 during the period October 1, 1978 to January 1, 1980. It is hoped that these findings will be of interest and importance to those with responsibility for making plans in a variety of sectors. The data base is so rich that we have necessarily only been able to tap a small part of it for our purposes. We expect and hope that others will follow in utilizing the information from INCAP and Rand as well as in undertaking related studies in other places. We do not expect that there will be such an ambitious collection of data made again, but that, instead, the lessons learned from the strengths and weaknesses of the INCAP and Berkeley efforts will assist those who would follow. We believe that some results of our investigation are clear and could indicate definite policy steps, while others will need to be further clarified.

The Scope of the Study

The pervasiveness of malnutrition and its relationships through both causes and effects to so many other sectors can be seen in the range of disciplines that are employed to study it. Research findings in many areas are recognized as essential to both the understanding of the problem and to the formulation and implementation of adequate policies.

This study is an example of such an interdisciplinary approach. While we began with a detailed examination of the links between early nutrition, growth, development, and school performance which was the primary concern of the study, we also moved into other related areas which must be addressed for the proper formulation of development policy.

Economists tend to see the lack of food available to the household and/or the individual in a malnourished population as a problem of insufficient production and distribution of food. Those trained in other disciplines point to diverse reasons for nutritional deficiency -- poor health, excessive and poorly spaced numbers of children per family, and poor distribution of food within the household. Educational policy makers are concerned with whether nutritional deficiency impairs school performance. Development planners attempt to measure how nutritional status affects work productivity. Nutritional planners search for adequate measures of expressing nutritional status in functional terms. Besides short and medium range policy problems, policy makers are concerned with the longer range pay-offs of nutritional policies such as how the outcomes of current nutritional policies will themselves be reflected in the future supply of food and the well being of children and families.

The current study addresses some of these problems and interrelationships. Using data from the malnourished rural population in Eastern Guatemala which was the subject of the INCAP/Rand investigation, we

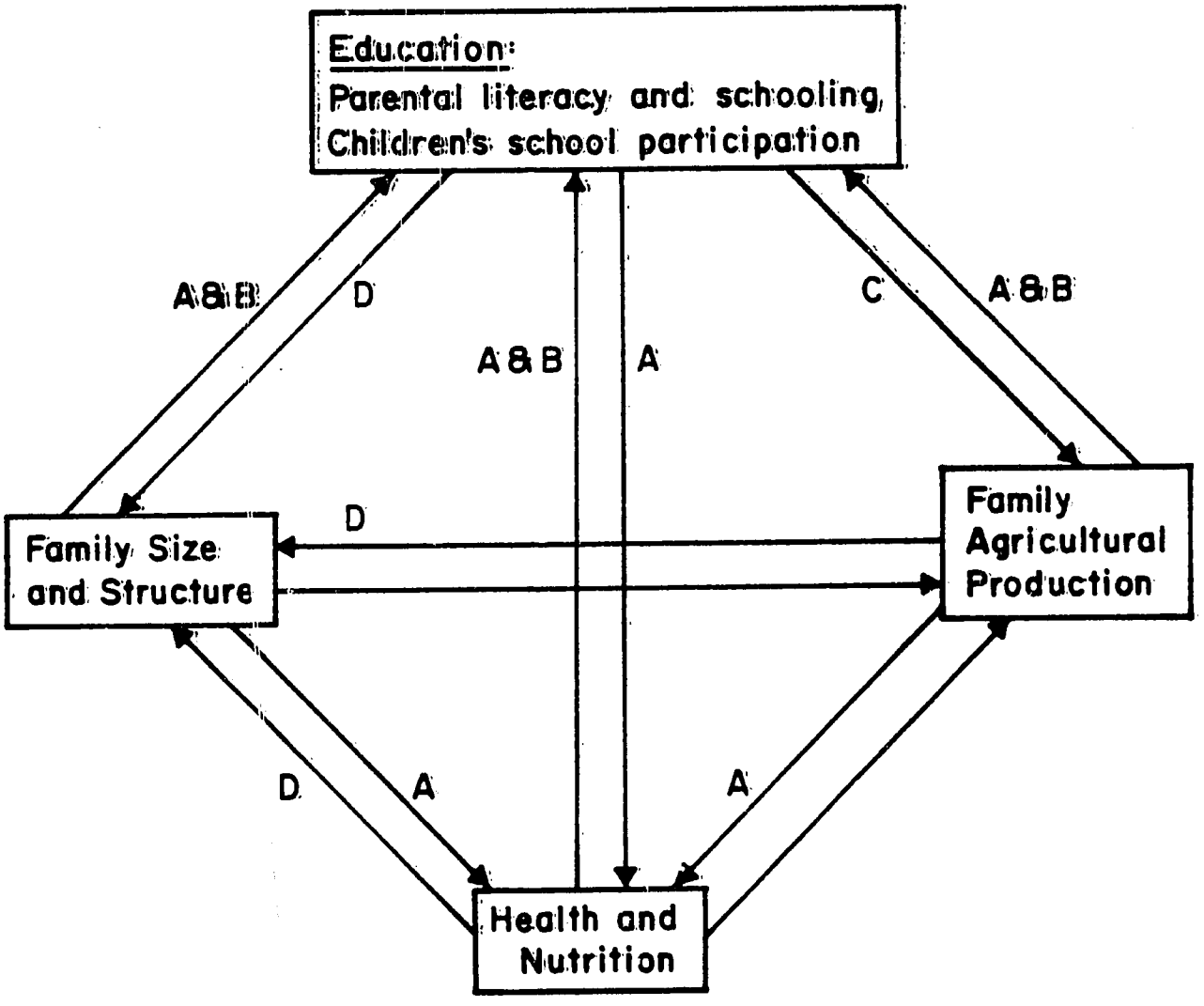
analyze the links between early childhood nutritional intake, health, growth, and school attendance and performance. Additional links between children's schooling, family work, family size, parental literacy, and agricultural efficiency are also investigated.

Figure 1 depicts, in simplified terms, the principal sectoral links that are addressed in analyses in this report. Each arrow is labeled with a letter indicating the section in Part II in the text which relates to the specific relationship depicted.

In each of the following sections, we will present research results and, wherever possible, the policy significance of these conclusions. In some cases, research findings are strong and clear enough to lead directly to policy conclusions while, in other cases, more research is needed. At the conclusion in Part III, we shall present a more complex version of Figure 1 from the perspective of the child which reflects more properly the results of our study.

Figure 1

Nutrition, Education, Agricultural Production, and Fertility: Interactions Among Sectors



Note: A, B, C, D refer to the section in Part II of this report that investigates the linkages.

Structure of the Report

This report is organized into three major parts. Part I-A describes the experiment carried out by INCAP (Institute for Nutrition in Central America and Panama) and summarizes INCAP's findings. Part I-B describes the villages where the experiment took place.

Part II presents the methods of analysis and findings. Alan Wilson's contribution in Section II-A describes a series of longitudinal models which relate the influence of early nutrition to growth, verbal development, school attendance, and school performance. The models, using information for children studied from birth, include how nutritional intake (calories and proteins, as well as home and supplementation) affects physical growth, how nutritional intake and physical growth then enter models of verbal development, school enrollment, and achievement.

Sections II-B, C, and D describe three cross-sectional analyses carried out with data from Rand-Rockefeller surveys and some elements from the longitudinal study. Although their populations vary in composition, their features are similar enough to allow their integration in the final summary of findings. In II-B, Judith Balderston analyzes how household decisions to send children to school and/or work (including both paid work and help to the family) appear to be determined by nutrition, health, and family economic factors. Utilizing Alan Wilson's findings relating height, nutritional intake, and health, Maria Freire's economic groupings (shown in II-C), and Mari Simonen's measures of perceived child economic utility (shown in II-D), the models presented in II-B focus on the significance of family and village economic conditions upon children's school attendance. In this part, older children's work activities are also examined in relation to family and village economic factors.

Sections II-C and D address the pay-offs of nutritional policies. In II-C, Maria Freire investigates the links between education and agriculture production for 500 farmers using data collected in the Rand-Rockefeller project. The relation of formal education and literacy to farm production and profits is analyzed, as well as a search for the main factors which affect farm production and, ultimately, food availability. Findings from this study can be used (1) to speculate on the intergenerational effect of nutritional policies since, with results from II-A and B, we can see how nutrition affects education which affects food production (with a generational lag) and, consequently, nutritional status; (2) to complement II-B and analyze children's economic contribution, and (3) to detect the constraints that may impinge upon the availability of food for the household as a whole.

In II-D, Mari Simonen analyzes relationships between education, fertility, and family size. Her findings are important in several ways: (1) differential behavior is observed between rural and urban populations (using data for two additional "semi-urban" villages not included in the longitudinal study); (2) family size and fertility are explained in terms of economic and social variables which contribute to inter-sectoral planning; (3) it complements II-B and II-C by explaining fertility and family size in terms of parents' perceptions of the economic utility of children.

In Part III, the main findings of each section are summarized. Then, viewing the analytical results as interrelated from the research policy standpoints, the report concludes with a discussion of the policy implications of various interventions.

In Appendix 1, we list the variables included in the models. Appendix 2 is a detailed description of the four villages. Appendix 3 shows the links between education and measures of agricultural performance. Appendix 4 is a report on fertility and education in the four rural villages as well as two additional semi-urban villages.

PART I

I-A. The Guatemalan Longitudinal Study: INCAP's Experiment and Findings by Sheldon Margen and Maria Freire

In order to understand the purpose of the quasi-experiment carried out by INCAP in Guatemala and to recognize what, in retrospect, may have been certain design flaws, it is necessary to realize that this work was based on assumptions about malnutrition which were prevalent during the 1960's.

During the 50's and 60's, protein-energy malnutrition (PEM) was believed to be a serious problem, particularly in the developing countries. Little was published, however, about the long-range effects of PEM because the literature dealt almost exclusively with the most severe cases in which the majority of the children were hospitalized and many of them died. Obviously, this represents only a small percentage of the estimated half to two-thirds of the children in the developing world who are alleged to be suffering from chronic mild-to-moderate PEM. Furthermore, at that time, it was believed that PEM was primarily due to dietary protein deficiency, and new high protein foods were developed while ways were sought to distribute these to pregnant women and infants, the most vulnerable members of the population.

The design of the Guatemalan study was predicated on this assumption, that protein was indeed the limiting factor. However, by the 1970's, we began to realize that PEM (except in certain isolated areas, e.g., of Africa), where the diet consisted mainly of tuberous materials, was not due to primary protein deficiency but rather to low caloric intake which has the effect of utilizing protein stores by energy rather than for building tissues. In other words, the amounts of dietary protein were adequate if the calories were adequate.

In order to study chronic mild-to-moderate PEM and determine its effect on mental development (specifically cognitive performance during the pre-school years), groups of Guatemalan children and pregnant women were given a high protein supplement. A control group was to be given no supplementation. The children were followed during the pre-school years, and many measures were taken in order to determine what effects the environment and the supplement might have on a host of outcome variables. Although, as originally designed, the INCAP study did not necessarily expect to produce a clear-cut association between chronic mild-to-moderate PEM and mental development, the investigators did hope to learn whether or not this condition resulted in impaired cognitive performance during the pre-school years. They further hoped to follow these individuals through school years, adolescence, and adulthood to see whether the effect, if any, was permanent.

The investigation was three years in preparation, during which time the investigators identified two pairs of matched villages and carried out extensive anthropological and ethnographic surveys of them. The study itself began early in 1969, and the data collection continued for 8 years, until March 1979.

One pair of villages contained about 850 residents each and, the other pair, approximately 500 residents each. Based on the assumption that the operant deficiency was protein, one village in each pair received nutritional intervention consisting of a high-protein substance called atole. The other village was controlled. In order to help minimize the effect of the intervention <u>per se</u>, the control villages were provided with a non-protein, low caloric drink, fresco. All villages received free preventive and curative out-patient medical services for the duration of the study.

Data were collected on growth, morbidity, home diet, family socio-economic status, and measures of mental development. The villages were chosen because they were similar in a number of characteristics such as age distribution, size, agricultural economy, isolation from urban areas, and their Ladino (Spanish speaking, non-Indian) culture. The villages, of course, are not identical. Village 8 has a milder climate, and during the study period, electricity was introduced. Women in this village earned money by selling hand-made palm products. Village 3 has two cement factories nearby that provide employment. Village 6 has a heavily travelled main road through its center that provides more contact with the outside and also has an active Protestant group. Village 14 is the smallest (about half the size of village 6) and has the largest amount of seasonal migration to work harvesting cotton.

Supplementation centers were opened in all four villages in January 1969, and data collection started. In two of the villages (6 and 14), the supplement was atole which contained 11.5 grams of high quality protein and 163 kcal energy per 180 ml, 28% of which was derived from protein. The other two villages were provided with a low energy beverage called fresco which had about a third of the energy concentration (approximately 59 kcal/180 ml) of the atole and contained no protein. In October 1971, both vitamins, minerals, and flouride were added to the supplements in order to guard against the possibility that these may have been limiting in the average home diet of the study families. This obviously complicated the analysis because the amount of the vitamins and minerals varied depending on both the home diet and the amount of supplements ingested.

Distribution of the supplement took place seven days a week for two hours in the morning and two hours in the afternoon. Attendance at

supplementation centers was voluntary, and the supplement was available to all members of the family in the community ad libitum. Since both supplements were liquid, it was possible to accurately measure the amount taken. Daily atole and fresco intakes were measured and recorded to the nearest centiliter.

Table 1

Some Information About Supplementation

Consumption of supplement is measured in centiliters:

1 cl = 10 ml = 1/18 cup
1 cl of atole = 9.05 calories (kcal) and .638 grams of protein
1 cl of fresco = 3.27 calories and 0 grams of protein
1 cup of atole = 163 calories
1 cup of fresco = 59 calories

Yearly Average of Supplement Consumed in Cl/Week by Year of Child's Life

	1	2	3	4	5	6	7
Fresco	5.925 (9.412)	24.552 (29.257)	58.426 (59.299)	95.639 (82.085)	134.609 (94.212)	154.742 (102.880)	188.408 (102.718)
Atole	45.475 (54.770)	83.939 (78.174)	105.839 (93.361)	110.261 (101.816)	105.440 (98.569)	107.271 (109.979)	114.188 (113.773)

Key: 1st number is mean, and standard deviations are shown in ( ).

At the start of the project in January 1969, all children under seven years of age were inclued in the sample, as were all pregnant and lactating women. The village census was updated quarterly, and when a child reached age seven, s/he was dropped from the sample. Newborn and immigrating children under age seven were added to the sample. The last children added were born in February 1973, bringing the total sample to 1400. Only those children born after January 1969 were included in the Berekely analysis.

Study Variables

1. Amount of food consumed. The supplementation was measurable with great accuracy and reliability. In addition, home dietary surveys were made in order to estimate the total food intake (home diet plus supplement). In field investigations such as this, there is always a question as to whether the administration of supplement is actually supplementary or is merely a replacement. Therefore, in order for these data to be meaningful, they must be analyzed in terms of the total food consumed, and an estimate of the potential error of home dietary measurements should be made. In this instance, 24-hour recall data were collected at three-month intervals. Mothers were questioned about the previous day's intake, and interviewers recorded information on the dietary patterns of pregnant and lactating women, as well as for all children 15 months to 84 months of age. At INCAP's headquarters in Guatemala City, specific nutrient intake was computed from the food composition tables prepared for Latin America, and this was summarized in terms of calories, total protein, animal protein, calcium, iron, vitamin A, vitamin C, niacin, riboflavin, and thiamine.

2. Anthropometry. Standard anthropometric techniques were utilized. The children were examined at birth, at 15 days, and every three months thereafter up to 24 months, then every six months between 30 and 48 months and yearly up to 84 months. Data on nine variables were collected. These included the supine length, total body weight, medial calf circumference, condylar breadth of femur, triceps, subscapular and medial calf skinfolds. In addition, dental eruption data were collected, and radiographs of the left hand and wrist were made on study children at the beginning and at the times that the children were otherwise examined.

3. Biomedical data. The biomedical data collected during the study consisted of medical care records, clinic information on attendance, and

symptomotology, pregnancy, delivery, and prenatal information. For a sub-sample of the study population, the delivery information included Apgar Score, examination of the infant, determination of mother's milk production, gestational age of the child, and urinary area/creatinine ratio determinations in lactating mothers and in children.

4. Morbidity surveys. Morbidity data collection began in 1970 on children seven years old or less and in pregnant and lactating mothers. The morbidity survey was made by retrospective home interviews of mothers every two weeks. These interviews were summarized monthly for each subject.

5. Social and economic indicators. A limited amount of data on a number of family and economic social indicators were also collected. Among these were qualitative ratings of the family house, parents' clothing, and mother's report of interaction with pre-school children. These data were collected in 1968, '72, '73, and '74. In addition, a vast amount of socio-economic, family, educational, and agricultural production data were collected in 1974 in a more detailed study to be described below.

6. Mental development. A wide variety of assessment techniques were used to measure mental development, and many of these were developed specifically for this study. The measurements obviously varied with the age of the child at the time of testing.

a.) Infant assessment. A composite infant scale was developed from four widely used standard psychomotor infant tests: the Bailey Scale of Psychomotor Development, the Cattell Infant Scale, the Merrill-Palmer Scale, and the Gesell Scale. In 1972, two other infant assessments were added. One of these was the Brazelton Neonatal Scale. The second one, the Infant Cognitive Battery, was introduced to assess mental changes during the last half of the first year of life.

b.) The pre-school assessment. Within two weeks of their third, fourth, fifth, sixth, and seventh birthdays, children in the longitudinal study received the INCAP Pre-School Battery which is composed of 21 tests. Twelve of the tests were developed in 1967 and 1968 to assess a variety of cognitive processes, and the remaining 9 tests were developed between 1970 and 1971 to examine other areas that seem to be of theoretical importance. Also included were recognized I.Q. tests such as the WPPSI and the Stanford Benet.

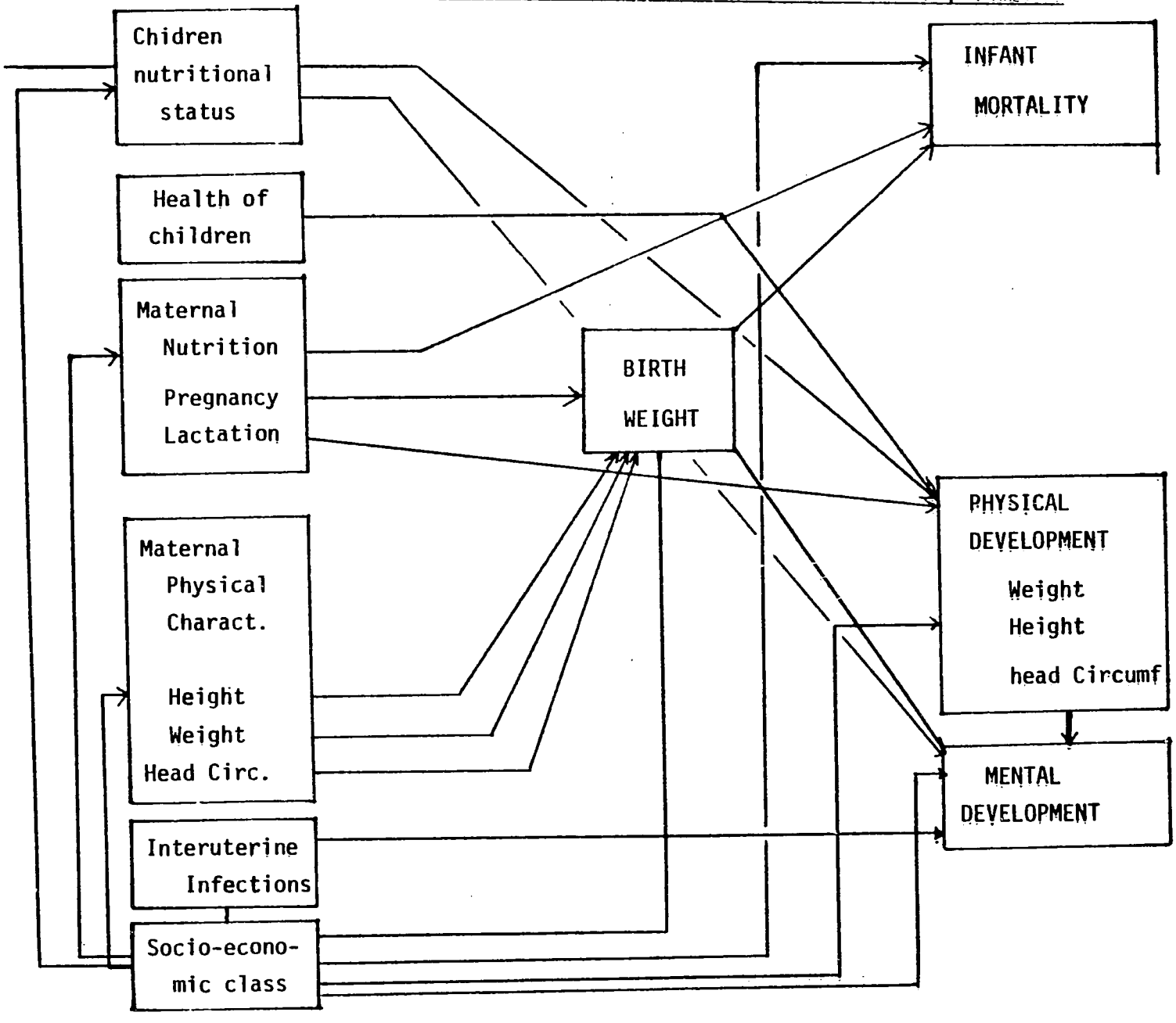
In 1974, the Rand Corporation, with funding from the Rockefeller Foundation, carried out an extensive study using questionnaires aimed at determining agricultural wealth and production, an opinion survey of the parents, and a questionnaire concerning the activities of children. These questionnaires were designed to answer questions not included in the main nutritional hypothesis and served as a rich resource to be used in the analysis of the data. Appendix I gives a list of sources of information from the Rand/Rockefeller files.

Results of the Experiment

The following is a rather brief summary of the majority of the published results of the experiments carried out in Guatemala by INCAP. Although a series of more detailed and complete summaries have been written, these are still in draft form and are not available for review. We have chosen to place the results in the general framework developed by Maria E. Freire (Figure 1). The main assumptions of relationships and causality among the variables are shown in this figure. The arrows represent the causality hypothesis being tested in the experiment. The descriptions fall under three main headings: (1) the determinants of birth weight; (2) the determinants of physical growth; and (3) the determinants of mental performance. Under each heading, we begin with a short abstract, derived

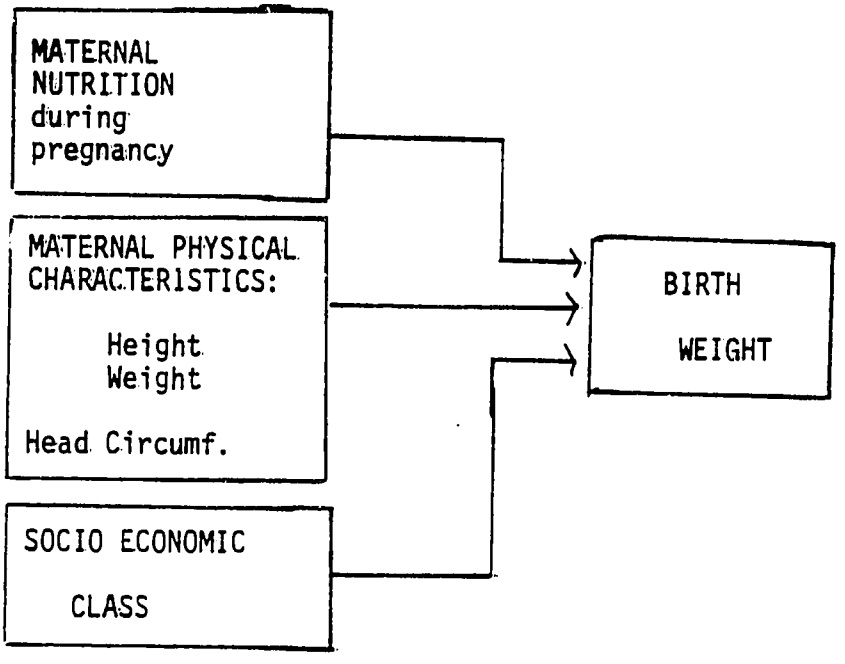
Figure 1

Network for the Causality Implicitly Assumed in the INCAP's Experiments.



from the papers published in this area, and follow with references to the papers and a short summary of the findings.

1. DETERMINANTS OF BIRTH WEIGHT



Abstract:

a. - Birth weight is mainly determined by the height and weight of the mother at the time of conception. These two factors are more important than the nutritional status of the pregnant woman.

b. - Since maternal height and weight are dependent on the nutritional status of the preschool girl (2-7 years old), low birth weight tends to continue through the next generation unless supplementation programs are given to the groups of women at more risk of having low weight babies.

c. - Supplemented nutrition should be given to shorter and thinner women since these are the women more likely to have lower birthweight children.

d. - Shorter women come from lower income classes reflecting what will be referred to in II.

e. - Differences in weight of placenta are associated with differences in calories ingested by the pregnant women. Proteins and other nutrients appear to have no effect on the placental weight nor in birth weight.

f. - When other variables are controlled, such as the ones obtained from dietary surveys, clinical and obstetrical conditions, and socioeconomic class, the effect of supplemented nutrition on birth weight is the same. An additional 10,000 calories during pregnancy is estimated to affect birth weight by 25 to 80 additional grams.

g. - Illness of mother affects birth weight. Mothers with high morbidity tend to have three times as many low birth weight children.

The above conditions are based upon a number of papers, some of which state that low socio-economic groups tend to have placental weights 15% lower than high socio-economic groups and that the amount of supplements ingested can decrease the difference. Intake levels of proteins and other nutrients appear to have no effect on placental weight. The relationship between placental weight and birth weight was not studied (1, 2).

As the experiments proceeded, it was pointed out that, utilizing data of the four villages, for low levels of supplementation, the number of low birth weight babies was around 40%, but for high levels of supplementation, it decreased to about 25% (3).

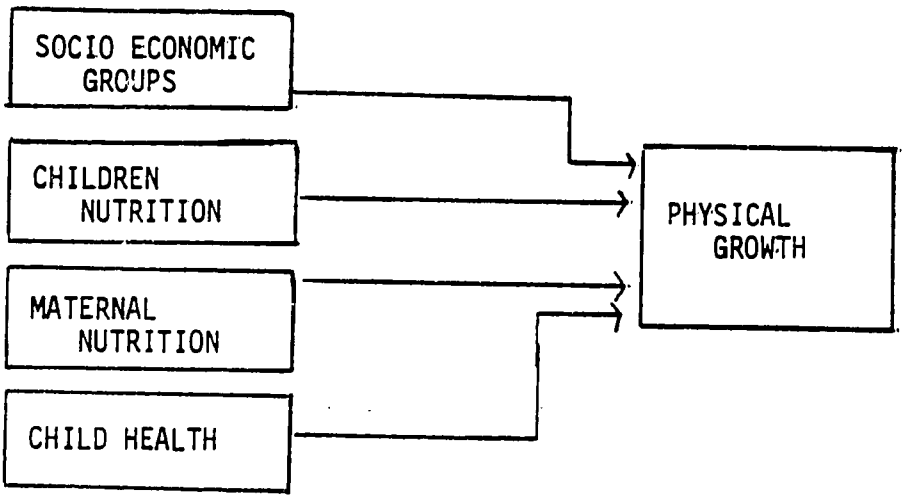
Over the four year period from 1969 to 1973 (671 births), the low supplemental group had 19% low birth weight babies, and the high supplemental group had 9% low birth weight babies. The additional birthweight per 10,000 calories was estimated at 29 grams (4, 5). In an earlier paper analyzing 418 births (6), the low supplementers were reported to have 17.9% low birth weight, and the high supplementers 10.6% low birth weight. The additional birth weight for 10,000 calories was estimated at 22 grams (6). In testing for other factors which might affect birth weight, such as the age, height, weight, and parity of the mother, none seemed to explain the relationship; only 18% of the variability in birth weight was associated with the mother's characteristics.

In another paper, the author studying 405 children again reported a change in birth weight of 29 grams per 10,000 supplemented calories (7). In a study of 1,083 children born between 1969 and 1973, it was found that the mother's illness affects the birth weight (5). The results show that the group that suffered repeated illness had 33.3% low birth weight, while the healthy group, 10.5%. It seemed that those with the high illness rate were less well nourished, therefore, the link between nutrition and birth weight seems to remain. In a paper rather early in the experiment in 1973, 852 children were studied. It was found that the mother's characteristics showed partial correlations (β) with child's birth weight as follows: mother's height with child's birth weight, $\beta = .22$; mother's weight with child's birth weight, $\beta = .20$; and birth weight with supplementation, $\beta = .17$. The mother's height and weight were highly correlated at .50. The taller the mother, the less supplement she seems to require to have a normal baby. For mothers of smaller size with a higher risk of low birth weight newborns, the supplement seems to have a strong effect. It seems that there is a 20% reduction in low birth weight babies for mothers who receive 15,000 calories of supplementation during pregnancy to about 5% low birth weights for mothers who consume 20,000 kcal (8). According to their method of analyses, the effect is due to calories and not protein since atole and fresco supplementation seemed to be equal in effect. It is further stated that maternal height determines the birth weight and that maternal height may be determined by malnutrition of the mother during her own early childhood at 2-7 years of age. It is suggested that birth weight would be improved by supplementing pre-school girls, causing them to be taller and heavier mothers (9). The effect of different supplementation is estimated as follows: 10,000 calories supplementation will lead to an increase of

1.2 grams in birth weight if administered to pre-school girls, 19 grams if administered to pre-pregnant women, and 28 to 80 grams if administered to pregnant women.

In one of the earliest papers in this series, the timing of the supplement to the mother was examined; it appeared that the influence of the supplement seemed to be slightly more important in the first trimester (10).

2. DETERMINANTS OF PHYSICAL GROWTH



Abstract:

a. - The main objective of studying the determinants of physical growth is to find an association between nutritional level of a child with his physical growth. Due to the difficulty in measuring nutritional level of the children, physical growth is taken as a good measure for nutritional status. INCAP's research to this problem followed two main lines: defining good indices for physical growth and finding the determinants of physical growth.

b. - The findings on the determinants of physical growth yield the following conclusions:

1. The main indices of physical growth being used by INCAP are height, weight, and head circumference, although other measures are being studied.

2. Length and weight seem to be highly correlated with each other, and the path of growth follows a linear relationship.

3. Growth is determined mainly by environmental variables, nutritional status, and health conditions. Ethnic factors do not seem to have explanatory power for the differences in attained size among different societies.

4. Diarrheal diseases accounts for 10% of physical retardation for children at age 7.

5. Nutrition has a strong effect on attained size at 7 years of age. Low supplemented children are three times more likely to be physically retarded when seven years old.

The data upon which these conclusions are based are as follows: in about 9,119 Guatemalan children from 0-7 years of age studied between '69 and '72, the weight and height were associated by the following formula, weight = 0.42 x height (in meters) -8.85 (r=0.979). These results showed a pattern of growth similar to a Denver group, and no differences appeared to exist between sexes. Well-nourished children seemed to present the same characteristics as those seen in the Denver study (11). In a paper published in 1974 (12), it was pointed out that extreme malnutrition affects growth, but the question was asked what about mild-to-moderate malnutrition. In Guatemala, it was found that children suffering clinical malnutrition showed a weight deficiency of about 2,250 grams, moderate malnutrition 1,250 grams, and mild malnutrition about 1,000 grams. A well-nourished child shows no decrement. The effects of the nutritional deprivation require about 6 months post partum before they are readily apparent. Weight seems to be more responsive than height. Head circumstance also seems to be a good indicator for children to about the age of two.

In the study of about 581 children followed from birth to 36 months (5) where malnutrition and growth retardation were studied (where growth retardation was defined as weight, height, and head circumference below the tenth percentile of the population), results comparing a low and high supplemented group show that the low supplemented groups seem to have three times as much physical retardation as the high supplemented groups. Low supplemented groups consumed, on the average, less than 5000 kcal per quarter, and the high supplemented group more than 10,000 kcal per quarter. The supplementation program is believed to reduce the growth differences

between Guatemala and the United States by 50%. This finding had been reported previously in an article which stated the physical growth retardation is apparently due about 50% to disease and about 50% to nutrition factors (13).

In a series of studies of growth patterns of 0-7 year olds measured at intervals reported previously for anthropometry, it was found that the correlation between height and calorie supplement was equal to 0.15, and the correlation between weight and calories is equal to .11. At three months, the low supplemented groups were reported to have a 0.9 centimeter deficiency in height and a .21 kilogram deficiency in weight; at six months, a 0.6 centimeter deficiency in height and .17 kilogram deficiency in weight (14). In 1972 (14), reporting on 1,119 children, the effects of atole (the protein supplement) were compared with fresco (the calorie supplement); it was found that the children who received atole grew faster and their weight also increased more between nine months and seven years. No effect was found from 0-6 months reflecting the breast feeding period (14). However, a paper in 1973 (15) refutes the former conclusion on the effects of protein. No effects of protein were found in the group of children from 1-3 years old. This was a small sample of about 50 children. Only the urea/creatine index seem to show some differences.

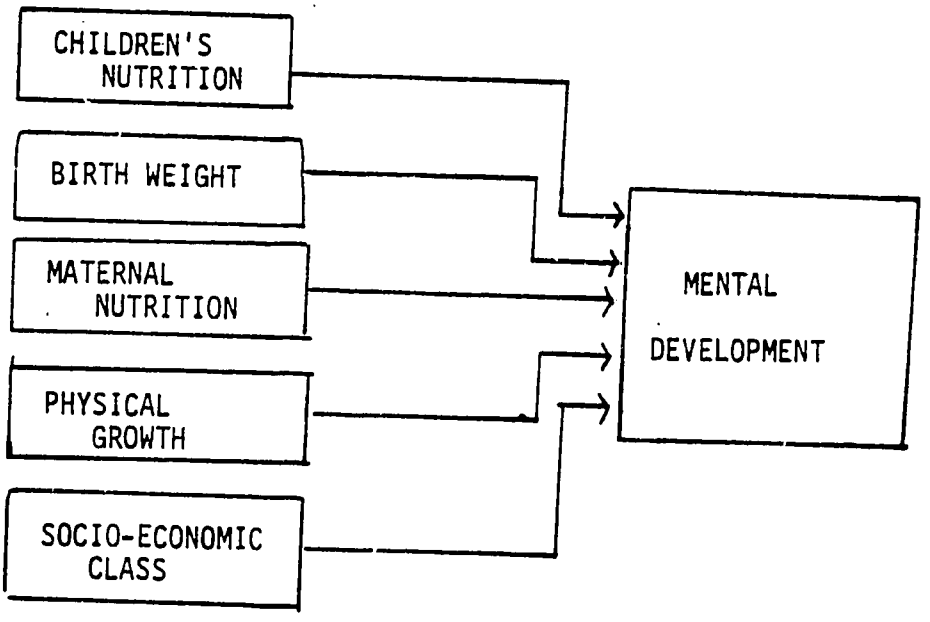
In the studies of the relationship between morbidity and growth, it was pointed out (16, 17) that diarrhea is responsible for at least 10% of the growth retardation at seven years of age. This was done in a study which followed children aged 15 days to seven years for two years and shows a height and weight differential of 3.5 cm. and 1.5 kg. at 7 years of age between high and low disease groups. Analyzing the type of illness (fever, respiratory, diahrrea), regression analysis of age, sex, diet

supplement, type of illness, growth, shows that diarrhea was the most important of the illnesses affecting the growth in both height and weight. This is typical of findings in the developing countries (18).

These findings were reported in further detail in a more detailed paper (19). In a methodological paper, the authors questioned the accuracy of under-reporting of morbidity data, utilizing their forthnightly recall method. They concluded that, depending upon the severity and the time lapse, under-reporting tends to increase. Therefore, even in the case of a two-week recall, a substantial amount of under-reporting is expected with minor illnesses (20).

As a final conclusion, the authors state that physical growth is determined only by environmental factors, mainly diseases and nutritional status. They believe that studies from Japan, India, the United States, and Guatemala confirm this ideal. Obviously, they are speaking of population groups and assuming a "common" genetic pool (21).

3. DETERMINANTS OF MENTAL DEVELOPMENT



Abstract:

a. - For low supplemented children, 11% of the cases show mental retardation; for high supplemented children, 2% show retardation.

b. - Multiple regression analysis, taking mental and motor performance as dependent variables and social class, nutritional status (in most instances measured by physical growth indices) and infant bearing practices as independent variables, show that:

1. For children from 0 to 15 months, mental performance is associated with low birth weight. Other variables such as mother's head circumference, mothers third trimester weight, caloric supplementation during pregnancy are also important.

2. For children from 0 to 24 months, growth measures appear correlated with mental tests scores at 24 months and 6 months; child rearing practices and social class are also correlated with mental performance.

3. For children of 3-7 years old, physical growth, measured by head circumference and height, seems more correlated with special kinds of tests. The association holds differently for girls and boys. For example, height is more correlated with vocabulary test scores, mainly in those villages where the supplements are rich in protein. Height is a good predictor for memory test scores for girls and head circumference for boys. Socio-economic factors seem to affect response time tests for boys and memory for designs in girls. When previous test scores are included in the multiple regression as an independent variable, it holds most of the explanatory power for the variation in mental test scores.

In an early study of children from 0-15 months (19), it was noted that the low supplemented children had an 11 percent incidence of "mental retardation," whereas the high supplemented group had only two percent as measured by the indices used at that age.

In two studies (23, 24), a sample of 64 children aged 0-24 months was analyzed. The object was to study the relationship between social class, measured by the type of house, nutritional status, developmental test performance, and infant caretaker interaction. The results show that the socio-economic characteristics, the infant caretaker interaction, and nutritional status were all strongly related. Family social status interacts with nutritional status at 16 months but not before. The family

social status interaction correlates negatively with playing alone (with an r=.43). Since child-rearing practices and family income affects mental performance and nutritional status, the interaction becomes complex because nutritional status is also affected by mental performance. In studies of older children, a correlation between a) nutritional status at 5-7 years (measured by height and head circumference), b) nutrient intake measured by number of days of attendance at supplement center, c) family socio-economic status (measured by house-type), and d) psychological performance (battery of tests) was determined. The results of the study showed that seven years mental development was strongly associated with attained age and social class. Multiple regression analysis of these variables, separate for boys and girls, show that neither attained physical size nor estimates of protein intake (from supplement plus home diet) were strongly correlated with differences in psychological tests. These tests were performed on a smaller group of children and seem to show some contradiction with the prior reported findings.

In a prior quoted paper, the relation between maternal nutrition and motor and mental performance was measured (5). These data were collected for children between the ages of 0-36 months followed between 1969 and 1973. Here, it was found that from the ages of 6-36 months, the amount of supplement was correlated with mental scores and/or motor performance at various ages; the most important finding was that the lower the social class of the mother, the higher the association between the supplement and psychological performance.

In another analysis in 1974 of the same children as above (13), for 352 children, it was found that the birth weight correlates with mental performance at 6 months. Furthermore, physical growth and mental performance at the ages from 3-7 shows results varying among tests. Taking the vocabulary

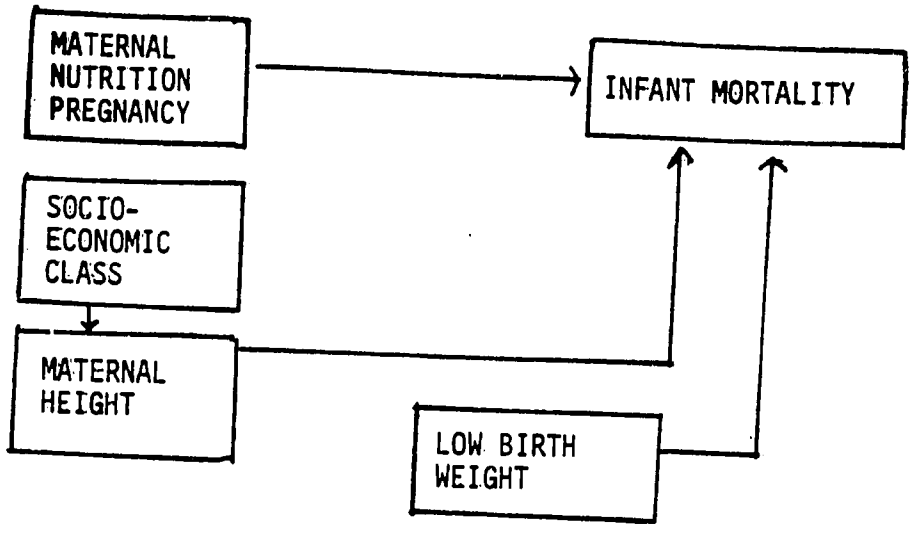
test scores as a measure of mental performance, the supplement is correlated for boys at 3 and 4 years of age in the villages with protein rich supplements and is almost negligible with villages where the supplement is only calories. In fact, the indices of physical growth, such as head circumference and height, seem more correlated with vocabulary than nutrients.

In a very early paper of the sample reported above (25), again relating physical growth as a proxy for nutritional status to mental performance measured by language, memory, and perception tests, it was found that the proportion of the variance accounted for by the physical growth measure was highest in the perception tests. In one of the earliest papers studying seven year olds (26), it was shown that head circumference and vocabulary scores are the most highly correlated with socio-cultural variables.

Again in another paper (27) in 1974 on 405 children between the ages of 3 and 7, it was shown that physical growth measured by height and head circumference is most highly associated with vocabulary tests. In another paper (28) based on the same sample, birth weight and mental development showed strong correlations, even after controlling for more than fifty variables.

Attempts to separate the effects of socio-economic, cultural, and physical growth variables as factors in the development of height and mental development seemed rather difficult (29, 30). The studies show that the cultural status interacts with social class, and both interact with nutrition and mental development. In another paper (31) on birth weight and mental performance at 6 months, the relationship between birth weight and psycho-motor performances at 6 months was shown to be related to the low birth weight rather than to other variables of maternal malnutrition. Birth weight was directly associated with scores on all tests. Other variables such as mother's head circumference, the mother's third trimester weight, and calorie-supplementation during pregnancy were also important.

4. DETERMINANTS OF INFANT MORTALITY



Abstract:

a. - Differences in infant mortality are due 68% to birth weight deficiency.

b. - Low birth weight babies are three times more likely to die than babies of adequate birth weight.

c. - With respect to supplemented nutrition for pregnant women, results show that:

The Amounts of Supplement in Kcals	Infant Mortality
5,000	56/1,000
5,000 - 9,000	54/1,000
20,000	24/1,000

d. - For the whole population, the following data show the benefits of the INCAP program:

	64-69	70-71
Neo-natal deaths	48.7	16.0
Infant deaths	104.0	48.5

e. - It is believed that the decrease in infant mortality, which decreased from 200/1,000 to 80/1,000 could be reduced to 55/1,000 if nutrition policies are implemented.

The outline on the previous page is a general schematization of the factors leading to increased infant mortality reported from the Guatemalan study.

In addition to this outline, there are several papers discussing the problems of measurement and methodological considerations (32-38). These papers discuss INCAP's attempt to overcome many of the difficulties which face field quasi-experiments such as: a) improvements in measuring physical growth, b) the development and interpretation of mental test scores and intelligence, c) the problems of inter-country and cultural comparisons, and d) the statistical problems of data analysis.

a.) The improvements in physical growth measurements. Very often, because of the difficulty of obtaining accurate information on nutritional intake, physical growth measurements are used as substitutes for that sort of information. However, it is often pointed out that chronological age may not equal physical age. Because of this, several studies on physical growth and age have been undertaken. Furthermore, in a child who is stunted and small for his age, the question is whether the size for age (physical age) or chronological age should be used in the analysis. Obviously, this question remains to be answered. However, a number of studies were carried out by the INCAP group to relate physical growth and age. These included the measuring of chronological age by bone age, the measuring of head circumference and chest cirumference and their importance, the use of the upper arm anthropometric measurements as indicators of nutritional status, the development of height and weight standards for pre-school age considering international and national measures, and the time of the deciduous tooth eruption and nutrition level (39, 40, 41, 42).

b.) Mental tests and intelligence. The obvious enigma on how to measure mental development remains. For very early ages, sensory motor tests are used. Although these are meaningful for the study of the adaptative behavior, they are probably not useful for testing intelligence. Measurement of I.Q. appears to be of little use in looking at nutrition and intelligence; the I.Q. test tells little about those aspects of mental development which are of interest in studying the impact of nutrition on behavior.

c.) School attendance and achievement. As the subjects of the supplementation study begin to reach school age, perhaps the most important questions emerge: have these experimental interventions had an effect on learning ability or on other attributes of social competency? This is the question to which our analysis is directed. One paper (43) examines questions related to this, on the relationships between pre-school test performance, family characteristics, and school performance for children born prior to the supplementation experiment. The conclusions of this paper are that school attendance depends on the child's development and the family's economic and educational background.

Summary

The review of the published material from INCAP is rather difficult to summarize. Since much of the data was analyzed before all the studies were complete and since different samples were the basis for different papers, contradictions do appear. It is difficult at present to reconcile these differences since the final papers, based on the entire experiment, have not yet been published by INCAP. In addition, there are two important points which we wish to make. First, most of the INCAP analyses were based solely upon the amount of supplement consumed and ignored home diet.

To do this requires the assumption that the supplement acted as a true incremental increase in the dietary intake of the individual and that, therefore, the amount of supplement consumed can serve as a proxy measure of total intake. However, close analyses of the data shows that there is an unknown degree of error in the home diet, and that the degree, direction, and variability of the bias in reporting home diet is unknown but important for interpretation of the experiment. It certainly appears that at least a portion of the supplement consumed is actually replacement of home diet; therefore, the amount of supplement cannot serve as a measure of the increase in diet. Second, in view of the attempts to analyze the data, primarily in terms of the supplement intake, the problem of differentiating between protein and caloric effects has been a very difficult one using INCAP's method of analysis. This is a problem which the Berkeley group recognized from the onset and that must be kept in mind when examining the published work to date.

The experiments which were carried out by INCAP and described in their published papers which we have summarized suggest the following interpretations:

1.) Birth weight is determined mainly by the physical characteristics of the mother at the time of conception, particularly height and weight. Since these two characteristics are strongly dependent on social class and the early nutritional status of the mother (when she was between the ages of one and seven), the problem of low birth weight is likely to continue for generations unless specific measures are considered. In particular, interventions that are directed specifically to shorter and thinner mothers may correct this condition. INCAP's experiments suggest that, for pregnant women, supplementation of 20,000 additional calories during pregnancy increases appreciably the weight of the newborn baby. The effect of an

additional caloric supplement of 10,000 calories increases the newborn's weight another 25-80 grams. There appear to be sex differences; the effect of increased calories during pregnancy is greater for female newborns than for males. The Guatemalan group believes that their results show that growth effects can be explained entirely on calories and that protein played no role.

2.) Birth weight appears to influence mental development during the first fifteen months of life. Mental performance also seems to be associated with certain physical characteristics of the mother, such as head circumference.

3.) Physical growth is often used by INCAP as a proxy of the child's nutritional status. Not only height, weight, and head circumference have been studied, but other anthropometric measurements such as bone growth and ossification, tooth eruption, and various types of arm measurements are used by INCAP to reflect the nutritional status of the child. Important conclusions can be derived from the INCAP data which show that physical growth is mainly dependent on environmental conditions, such as nutrition and health conditions, and independent of genetic factors. For developing countries, the most important health factor explaining delayed physical growth appears to be diarrheal disease, supposedly responsible for 10-15% of the growth retardation in younger children. Nutrition (including nutrition of pregnant women) seems to be the other important factor explaining the physical growth retardation. Results of the Guatemalan experiment show that low supplemented children are three to seven times more likely to be physically retarded than children who are receiving high supplements. It is estimated that adequate supplemental feeding programs would reduce the differences of growth between Guatemalan and U.S. children by 50%.

4.) The causes of poor or lower mental performances have not been uniquely determined. Certain measures of physical growth seem to have some predictive power for specific mental tests, but results vary by age and sex and reflect the powerful influence of socio-economic factors such as infant rearing practices, economic status, etc. Analysis of mental development between the ages of five and seven years "shows that prior test results are the best predictors of seven year old's performance." One might infer that the mental characteristics of the child, therefore, are settled at five years and no subsequent changes can result from an alteration in environmental conditions. Whether this is a correct interpretation of the results of the Guatemalan experiment or whether this conclusion will hold up in the final analysis is obviously open to serious question.

5.) Final Conclusions. Many studies report the influence of birth weight on mental performance during the life of the human being. These results have been accepted at this time virtually as "universal truth." If this is true, then, the main policy implications of the Guatemalan data would suggest that one of the principal methods for avoiding mental performance impairment would be to begin intervention primarily at the pre-school level, concentrating on girls to avoid their physical retardation, lower weight new-borns, and high-risk mothers. Finally, one might assume from the INCAP experiments that nutritional factors are strongly associated with physical growth but that their influence on mental tests of performance is not yet well determined over the long-run. For policy formulation, the short-run effects measured by psychological tests will need to be translated into measures of school performance and social competency. This is the question that will be considered in analyses that were examined by the Berkeley group and that are described in the present report.

BIBLIOGRAPHY

(1) Lechtig, A.; Behar, M.; et al., "Influence of Maternal Nutrition on the Characteristics of the Placenta and Fetal Growth." Paper presented at the symposium, "Effects of Inadequate Nutrition on Human Development," from the Continental Meeting on Science and Man, held June 20 to July 4, 1973, Mexico, D.F.

(2) Lechtig, A.; Yarbrough, C.; et al., "Effect of Maternal Malnutrition on the Placenta." American Journal of Obstrectics and Gynecology 123 (1975): 191-201.

(3) Lechtig, A.; Delgado, H.; Lasky, R.E.; et al., "Maternal Nutrition and Fetal Growth in Developing Societies. Socioeconomic factors." American J. Dis. Child. 129 (1975): 434-437.

(4) Lechtig, A.; Habicht, J. P.; et al., "Effect of Food Supplementation During Pregnancy on Birthweight." Pediatrics 56 (No. 4) (October 1975).

(5) Klein, R. E., et al., "Effect of Maternal Nutrition on Fetal Growth and Infant Development." Paper presented at the Fourteenth Meeting of the PAHO Advisory Committee on Medical Research, Washington, D.C., July 7-11, 1975.

(6) Habicht, J. P.; Lechtig, A.; et al., "Maternal Nutrition, Birth Weight and Infant Mortality." Size at Birth, Ciba Foundation Symposium 27. Amsterdam: ASP, 1974.

(7) Lechtig, A.; Habicht, J. P.; et al., "Maternal Nutrition and Birth Weight." Fourteenth International Congress of Pediatrics, Buenos Aires, October 1974.

(8) Habicht, J. P.; Lechtig, A., et al., "The Effect of Malnutrition During Pregnancy on Survival of the New-Born." Hearings on the Select Committee on Nutrition and Human Needs, U.S. Senate, Washington, D.C., June 1973.

(9) Lechtig, A.; Yarbrough, C.; et al., "Influence of Maternal Nutrition on Birth Weight." The American Journal of Clinical Nutrition 28 (November 1975): 1223-1233.

(10) Habicht, J. P.; Lechtig, A.; et al., "The Effect on Birth Weight of Timing of Supplementation During Pregnancy." IX International Congress of Nutrition, Abstract of Short Communications, September 1972, Arroyo, Barta, et al. (eds.), p. 149.

(11) Yarbrough, C.; Habicht, J. P.; et al.; "Length and Weight in Rural Guatemalan Ladino Children: Birth to Seven Years of Age." American Journal of Physical Anthropology 42 (1975): 439.

(12) Yarbrough, C.; Habicht, J. P.; et al., "Anthropometry as an Index of Nutritional Status." Nutrition and Malnutrition. New York: Plenum Press, 1974. Volume 29 of Advances in Experimental Medicine and Biology.

(13) Division de Desarollo Humano -- Nutrition, Crescimiento y Desarollo. Boletin de La Oficina Sanitaria Pana Americana, January 1975.

(14) Habicht, J. P.; Lechtig, A.; Yarbrough, C., and Klein, R.E.: "The Timing of Supplementary Feeding on the Growth of Rural Preschool Children." Proc. 9th International Congress of Nutrition, Mexico City, 1972.

(15) Habicht, J. P.; Schwedes, Arroyave, G; and Klein, R. E. "Biochemical Indices of Nutrition Reflecting Ingestion of a High Protein Supplement in Rural Guatemalan Children." American Journal of Clinical Nutrition 26 (1973): 1046-1052.

(16) Martorell, R.; Yarbrough, C.; et al., "Diarrheal Diseases and Growth Retardation in Preschool Griata Malan Children." American Journal of Physical Anthropology 43: 341-346.

(17) Martorell, R.; Yarbrough, C.; et al., "Diarrea y Retardo en Crecimiento Fisico, INCAP National Document, DDD ≠ 90.

(18) Lasky, R. E.; Lechtig, A.; Delgodo, H.; Klein, R. E.; Engle, P.; Yarbrough, C.; and Martorell, R. "Birth Weight and Psychomotor Performance in Rural Guatemala." Amer. J. Dis. Child. 129 (1975): 566-570.

(19) Martorell, R.; Habicht, J. P.; Yarbrough, C.; Lechtig, A.; Klein, R. E.; and Western, K.A. "Acute Morbidity and Physical Growth in Rural Guatemalan Children." Amer. J. Dis. Child. 129 (1975): 290.

(20) Martorell, R., Habicht, J. P.; Yarbrough, C.; Lechtig, A.; and Klein, R. E. "Under Reporting in Fortnightly Recall Morbidity Surveys." Env. Child Health 129 (June 1976).

(21) Martorell, R.; Lechtig, A.; et al. "Normas Antropometricas de Crescimiento Fisico Para Paises en Desarollo: Nacionales o Internacionales?" Publicación INCAP E-833.

(22) Lechtig, A.; Delgado, H.; Lasky, R.; et al. "Effect of Improved Nutrition During Pregnancy and Lactation on Developmental Retardation and Infant Mortality." Paper presented at the Western Hemisphere Nutrition Congress, Symposium on Perinatal Nutrition, Bal Harbour, Florida, Miami, August 19-22, 1974.

(23) Klein, R.; Yarbrough, C.; et al., "Correlations of Mild to Moderate Protein · Caloric Malnutrition Among Rural Guatemalan Infants and Preschool Children." Symposium of the Swedish Nutrition Foundation XII, 1974.

(24) Klein, R.; Lasky, R.; et al., "Relationship of Infant/Caretaker Interaction, Social Class and Nutritional Status to Developmental Test Performance Among Guatemalan Infants." INCAP unpublished document, June 1973.

(25) Klein, R.; Lester, B. M.; et al., "On Malnutrition and Mental Development. Some Preliminary Findings." Proc. 9th International Congress of Nutrition, Mexico, 1972, Vol. 2, (Kargen-Basel 1975), pp. 315-321.

(26) Klein, R., Lester, B., et al., "Cross-cultural Education of Human Intelligence." Lipids Malnutrition and the Developing Brain, a Ciba Foundation Symposium, October 1971. Amsterdam: ASP.

(27) Engle, P.; Klein, R. E.; et al., "Effecto de la des Nutricion Sobre el Desarollo Mental." Serminario on Organization of Mental Retardation Services, Scientifique Publication No. 293, Cartagena, Colombia, 1973. Washington, D.C.: OPS, 1974. pp. 73-80.

(28) Delgado, H.; Lechtig, A.; et al., "Maternal Nutrition: Its Effects on Infant Growth and Development and Birthspacing." Harold C. Mack Symposium, "Nutrition and Human Reproduction: Biochemical and Clinical Aspects," Detroit, Michigan, October 1975.

(29) Nerlove, S.; Roberts, J.; et al., "Natural Indicators of Cognitive Development: An Observational Study of Rural Guatemalan Children." School of Urban Public Affairs, Reprint No. 124, 1974.

(30) Klein, R.; Freeman, H.; et al., "Is Big Smart? The Relation of Growth to Cognition. Journal of Health and Social Behavior 13 (September 1972).

(31) Lasky, R. E.; Lechtig, A.; Delgodo, H.; Klein, R. E.; Engle, P.; Yarbrough, C.; and Martorell, R. "Birth Weight and Psychomotor Performance in Rural Guatemala." Amer. J. Dis. Child. 129 (1975): 566-570.

(32) Kagan, J., Klein, R., et al. "Memory and Meaning in Two Cultures." Child Development 44 (1973): 221-223.

(33) Klein, R.; Freeman, H. F.; and Millet, R. "Psychological Test Performance and Indigenous Conceptions of Intelligence." The Jouranl of Psychology 84 (1973): 219-222.

(34) Kagan, J., and Klein, R. "Cross-cultural Perspectives on Early Development." American Psychologist, November 1973, pp. 947-961.

(35) Lester, B.; Kotelchuck, M.; et al. "Separation Protest in Guatemalan Infants: Cross-cultural and Cognitive Findings." Developmental Psychology 10 (No. 1) (1974): 79-85.

(36) Kagan, J.; Klein, R.; and Lester, B. "Recall and Clustering Among Culturally Isolated Children." Unpublished working paper of INCAP.

(37) Lester, B., and Klein, R., "The Effect of Stimulus Familiarity on the Conservation Performance of Rural Guatemalan Children." The Journal of Social Psychology 99 (1973): 197-205.

(38) Romney, A. K.; Kieffer, M.; and Klein, R. "A Normalization Procedure for Correcting Biased Response Data." Social Science Research 2 (1973): 307-320.

(39) Martorell, R.; Yarbrough, C.; et al., "Upper Arm Anthropometric Indicators of Nutritional Status." The American Journal of Clinical Nutrition 29 (January 1976): 46-53.

(40) Delgodo, H.; Habicht, J. P.; et al., "Nutritional Status and the Timing of Deciduous Tooth Eruption." The American Journal of Clinical Nutrition 28 (March 1975): 216-224.

(41) Martorell, R.; Yarbrough, C.; et al., "The Head Circumference/Chest Circumference Ratio in Mild-to-Moderate Protein-Calorie Malnutrition." Environmental Child Health (August 1975).

(42) Lechtig, A., Martorell, R., Yarbrough, C.; Delgodo, H.; and Klein, R. E. "The Urea/Creatinine Ratio: Is It Useful for Field Studies?" Environmental Child Health 121 (June 1976).

(43) Irwin, M.; Engle, P. L.; Yarbrough, C.; Klein, R. E.; and Townsend, J. "The Relationship of Prior Ability and Family Characteristics to School Attendance and School Achievement in Rural Guatemala." Child Development 49:415 (1978).

I-B. A Brief Description of Villages

This description is based primarily on the INCAP/Rand family census of early 1975, the various Rand questionnaires, the description by Victor Mejia Pivaral[2], descriptions of the villages prepared for Rand in 1975 (in Spanish), and our own work, based on analysis of data by village.

Both the four rural and the two semi-urban villages are included in this description. We shall begin with the four rural villages and subsequently comment on the two semi-urban villages as to their distinct features from the rural ones.

The Four Rural Villages

All four villages are located in the Department of El Progresso at distances of 36 to 102 kilometers from Guatemala City. There is no public transportation directly to the villages; roads are in relatively poor condition and, thus, traffic is light, and travel between each of the villages and the outside world is done with some difficulty. For this reason, the villages were considered as good choices for the longitudinal experiment.

The rainy season for all villages is during the months of May to September. Climatic conditions are similar for villages #3, #6, and #14, but village #8 lies at a lower elevation and receives less rain and has warmer temperatures. Village #8 is also distinct from the other villages in that it is considerably further from Guatemala City than the others. Village #8 is also closer to a larger town than the others, close enough, in fact, for its children to attend school in that town.

---

(1) Coded as Villages #31 and #32.

(2) Guatemala Indigena, Vol. 7, #3, July-September 1972, in Spanish.

The culture of all the villages is Ladino, meaning that the people speak Spanish and have adopted the practices of the modern rather than traditional Indian culture. Most people in the villages are nominally Catholic, although there are also active Protestant groups.

Literacy of the villagers is less than 50% of those age seven or older. The highest literacy is in village #8 (42%) and the lowest in village #14 (27%). Men are more likely to be literate than women. People working in non-agricultural activities are more likely to be literate than agricultural workers. Literacy is most important in getting work outside the villages, for example, for young women who want to work as domestics in Guatemala City.

More than half of the villagers over age seven have passed no grades in school. The sex differences in grades passed are small, except in village #8 where many more men have schooling than women. Each village has a primary school; some teachers commute daily from Guatemala City (especially in village #3), while others live near the vilages. Children may start school at age seven, but many do not start until eight, nine, or ten. Repetition of grades is common. Of those aged ten, about 70% are students, but percentages differ markedly by village. (See Table II-1 for village differences and Section II-B for a fuller discussion of school participation.)

Boys' involvement in school is, in general, higher than girls', the largest difference between boys and girls being in village #8. Surprisingly, in village #6, girls enroll much more than bodys do, and the number of years enrolled is also higher (see Table 1-1).

Children begin to help their parents when they are quite young. Girls help their mothers preparing tortillas, caring for younger siblings, and carrying water, and boys help their fathers in the family plots. (Section

II-B describes in greater detail the involvement in various activities in the home and on the family farm plots.)

Contraception is rarely used among the villagers, although a national radio campaign has made them more aware of its availability. Women have children at close intervals and breast feed them for long periods. (Section II-D describes at greater length family planning knowledge and practice in the villages.)

Houses are constructed of local materials. Floors are usually dirt, roofs are tile (since the earthquake, many are now metal), and walls are reed or adobe. (Village #8 has a milder climate so that houses there need not be as well-built as in the other villages.) Most of the houses which are generally owned by the occupying family consist of one or two rooms. Only village #8 has electric power, and fewer than half of the families in village #8 have electricity (at the time of the survey) in their own homes. In villages #3 and #14, most families carry their water from public wells in the village. In #6 and #8, approximately 15% of the families have their own well, but one-third must carry their water from a river nearby. Over half of the families have radios, and about 10% have sewing machines.

The economy is mainly agricultural, with corn and beans grown as staple crops, tomatoes and chilis as cash crops, and maicillo (sorghum) primarily as animal feed. (Section II-C is devoted to a fuller description of agricultural production.) Almost 90% of the male heads of families have agricultural occupations, including some who are day laborers for other farmers. Some families leave the villages in the dry season to earn wages as harvesters. Village #8 has the lowest rate of migration, and village #6 the highest. About 80% of the female heads of families do not work for money, except in village #8 where about three-fourths

TABLE I-1

Comparison of Educational Indices by Sex and Village for Children Born Between 1962 and 1972

Villages	Vil. 3		Vil. 6		Vil. 8		Vil. 14		All	
	Cases	%	Cases	%	Cases	%	Cases	%	Cases	%
Enrolled before or at age 7	46	23.3	26	12.6	42	28.6	27	16.4	141	19.7
Ever attended school	83	41.9	56	27.2	69	46.9	45	27.3	253	35.3
Participation Rates(1)	242	58	190	42	169	60	146	49	747	52
Boys	129	61	84	35	89	66	75	52	377	52
Girls	113	55	106	50	80	55	71	46	370	51
Average number years enrolled										
Boys		2.85		2.26		3.17		3.08		2.84
Girls		3.26		2.54		3.04		2.69		2.89
Rate of Promotion(2)										
Boys		.84		.84		.76		.73		.80
Girls		.82		.83		.79		.76		.80

(1) Number of children completing at least one year of school and proportion of children of school age enrolled in school for at least one year.

(2) Number of years passed in school/years spent at school.

TABLE I-2

LITERACY AND SCHOOLING* OF MALE HEAD OF HOUSEHOLD, RURAL VILLAGES.

Percentage and Number of Male Heads
of Household That Can Read and Write by Villages

Village	3		6		8		14		All	
	N	%	N	%	N	%	N	%	N	%
0	77	44.8	100	51.0	42	32.3	63	55.8	282	46.2
1	27	15.7	25	12.8	20	15.4	19	16.8	91	14.9
2	68	39.5	71	36.2	68	52.3	31	27.4	238	39.0
All	172	100.0	196	100.0	130	100.0	113	100.0	611	100.0

0 = doesn't read or write
1 = a little
2 = yes

5 cases with missing data are omitted.

Number and Percentage of Male Heads
of Household by Grades and Villages

Grades	VILLAGES									
	3		6		8		14		All	
	N	%	N	%	N	%	N	%	N	%
0	97	62.6	108	60.3	51	41.5	74	69.2	330	58.5
1	4	2.6	6	3.4	2	1.6	3	2.8	15	2.7
2	16	10.3	32	17.8	22	17.9	12	11.2	82	14.5
3	17	11.0	19	10.6	26	21.1	13	12.1	75	13.3
4 to 6	21	13.5	13	7.3	22	17.9	5	4.7	62	11.0
8	--	---	1	.6	--	---	--	---	--	---
All	155	100.0	179	100.0	123	100.0	107	100.0	564	100.0

52 cases with missing data are omitted.

---

*Literacy coded as 0 = does not read or write at all
1 = reads and writes a little
2 = reads and writes well

Schooling = Number of grades of schooling completed.

TABLE I-3

LITERACY AND SCHOOLING* OF FEMALE HEAD OF HOUSEHOLD, RURAL VILLAGES.

Percentage and Number of Female Heads
of Household that Read and Write by Villages

Village	3		6		8		14		All	
	N	%	N	%	N	%	N	%	N	%
0	101	49.0	128	62.7	108	65.1	77	61.1	414	59.0
1	46	22.3	33	16.2	23	13.9	29	23.0	131	18.7
2	59	28.6	43	21.1	35	21.1	20	15.9	157	22.4
All	206	100.0	204	100.0	166	100.0	126	100.0	702	100.0

0 = doesn't read or write
1 = a little
2 = yes

6 cases with missing data are omitted.

Number and Percentage of Female Heads
of Household by Grades and Villages

GRADES	VILLAGES									
	3		6		8		14		All	
	N	%	N	%	N	%	N	%	N	%
0	113	55.4	131	65.2	109	65.7	81	64.3	434	62.3
1	12	5.9	4	2.0	4	2.4	5	4.0	25	3.6
2	44	21.6	39	19.4	26	15.7	24	19.0	133	19.1
3	18	8.8	15	7.5	17	10.2	14	11.1	64	9.2
4 to 6	17	8.3	12	6.0	9	5.4	2	1.6	40	5.7
15	--	---	--	---	1	.6	--	---	1	.1
All	204	100.0	201	100.0	166	100.0	126	100.0	697	100.0

*Literacy coded as 0 = does not read or write
1 = reads and writes a little
2 = reads and writes well

Schooling = number of grades of schooling completed.

TABLE I-4

PERCENT WHO READ AND WRITE AND PERCENT WITH NO GRADES OF ALL AGE SEVEN AND ABOVE, BY SEX AND VILLAGE. RURAL VILLAGES.

Percent Who Read and Write

Sex	Village				
	3	6	8	14	All
Male	38.7	30.2	57.1	31.5	39.4
Female	35.9	23.6	28.7	22.0	27.5
All	37.3	27.1	42.4	26.9	33.5

Percent With No Grades

Sex	Village				
	3	6	8	14	All
Male	56.7	62.5	33.9	57.8	53.8
Female	49.3	62.9	56.2	58.9	56.6
All	53.0	62.7	45.6	58.4	55.2

TABLE I-5

SOME ECONOMIC INDICATORS FOR AGRICULTURE PRODUCTION. RURAL VILLAGES.

Village	Vil. 3	Vil. 6	Vil. 8	Vil. 14	TOTAL
Number of cases	229	238	184	136	787
No. cases not missing	207	225	178	124	734
% cases using land	82%	83%	65%	88%	79%
Average extension of land per family (cuerdas)	97.23	83.17	36.8	70.3	68.6
Average size of parcel (cuerdas)	28.70	37.7	15.3	23.5	27.5
Average number of plots	2.90	2.1	2.3	2.8	2.5
Average price of land (quetzales/cuerdas)	1.13	2.7	3.7	.9	2.0
Average productivity of land . . .(quetz./cuerdas)	3.0	4.81	8.85	3.16	4.79
% of corn and beans in total agriculture output	84%	50%	33%	74%	55%
% of corn and beans sold (out of respective output)	14.4%	20.4%	7.7%	14.2%	16%
% agriculture output sold	23.7%	53.6%	54.9%	27.2%	45.2%
% agriculture output not sold	76.3%	46.4%	45.0%	62.8%	54.8%
Labor-Land ratio (men/cuerdas)	3.08	3.24	2.5	2.65	2.9
Fertilizer-Land ratio	.16	.39	.55	.19	.34
Hired Labor-Total Labor (%)	11.97%	19.9%	47.9%	10.1%	21.1%
Productivity of Labor (Gross) (cuerdas)	.95	1.14	1.64	.99	1.15
Productivity of Labor (Net) (cuerdas)	.90	1.02	1.42	.92	1.04
% Families Using Hired Labor	34.1%	43.0%	47.8%	28.4%	38.6%
Average Wage (cuerda/day)	.75	.85	.87	.84	.84

TABLE I-6

CHARACTERIZATION OF FOUR RURAL GUATEMALAN VILLAGES
PRODUCTION AND SALES -- FIVE MAJOR CROPS

Village	Total				Village 3				Village 6				Village 8				Village 14			
	N. Families		Av. Value	Weight	N. Families		Av. Value	Weight	N. Families		Av Value	Weight	N. Families		Av. Value	Weight	N. Families		Av. Value	Weight
Crops Grown:			quetz.				quetz.				quetz			%	quetz.	%		%	quetz.	%
Corn	568	95.0	133.4	42.2	169	89.4	143.2	70.8	183	97.3	146.0	33.9	109	84.0	173.0	32.4	107	97.3	92.5	48.6
Beans	305	51.0	74.8	12.7	97	51.3	46.9	13.1	138	73.4	91.4	16.3	16	12.3	14.9	.5	54	49.1	95.2	25.3
Tomato	91	15.2	277.7	14.1	4	2.1	156.7	1.8	83	44.1	298.8	30.6	---	----	----	----	4	3.6	12.3	2.5
Chili	42	7.0	855.6	20.0	--	-----	-----	----	24	12.8	546.7	16.7	17	13.1	1318.1	48.5	1	.9	357.8	1.8
Malcillo	114	19.1	49.1	3.1	58	30.7	51.4	8.7	4	2.1	59.0		1	.8	17.3	0.0	51	46.4	46.3	11.6
Others	130	21.7	109.3	7.9	25	13.2	74.2	5.4	27	14.4	64.0	2.2	54	41.5	158.3	18.5	24	21.8	86.5	10.1
Total. . .	598	100.0	300.2	100.0	189	100.0	201.1	100.0	188	100.0	419.5	100.0	130	100.0	355.2	100.0	110	100.0	104.2	100.0
Crops sold:[a]		% [b]	quetz.	c)		% [b]	quetz.	c)		% [b]	quetz.	c)		% [b]	quetz.	c)		% [b]	quetz.	c)
Corn	181	32.0	50.9	12.2	63	37.0	56.2	14.6	78	42.6	49.4	14.4	23	32.1	52.8	8.1	17	7.0	36.15	6.2
Beans	78	25.5	82.1	28.5	9	9.5	58.2	11.5	48	34.8	88.9	33.3	2	12.5	7.6	6.3	19	35.1	84.1	31.1
Tomato	90	98.0	274.1	97.6	3	75.0	206.5	98.9	83	100.0	283.9	97.6	--	----	----	---	4	100.0	120.4	88.7
Chili	42	100.0	641.9	75.0	--	---	----	----	24	100.0	548.4	99.9	17	100.0	790.8	60.0	1	100.0	357.8	100.0
Malcillo	47	41.2	39.2	32.9	33	57.0	39.2	61.8	1	.7	115.0	6.7	--	-----	-----	----	11	27.0	345.0	18.9
Others	130	100.0	109.3	100.0	25	100.0	74.2	100.0	27	100.0	64.0	100.0	54	100.0	158.3	100.0	24	100.0	86.5	100.0
Total. . .	363	71.4	110.7	46.4	96	56.8	81.4	20.6	136	71.0	343.3	29.4	79	62.0	293.9	50.3	52	100.0	107.2	27.4

[a] Sales evaluated at average prices.

[b] $\frac{\text{(Number of farmers who sell crops i)}}{\text{Number of farmers who grow crop i}} \times 100$

[c] $\frac{\text{Total sales crop i}}{\text{Total harverst crop i}} \times 100$

TABLE I-7

INCOME SOURCES (in quetzales)

Village	Total quetz.	%	Vill. 3 quetz.	%	Vill. 6 quetz.	%	Vill. 8 quetz.	%	Vill. 14 quetz.	%
Agric. Income:	250.8	45.5	170.9	32.2	355.2	53.6	263.8	53.9	170.8	36.7
Consumed	134.4	24.4	130.8	24.8	144.9	21.9	131.1	26.8	124.0	26.6
Sold	116.4	21.1	39.1	7.4	190.3	31.7	132.7	27.1	46.8	10.1
Non-Agric. Income:	300.0	54.5	360.5	67.8	307.5	46.4	225.4	46.1	294.6	63.3
Salaries	219.3	39.8	295.7	55.6	229.0	34.6	152.6	31.2	171.0	36.7
Other Activities	80.7	14.7	64.8	12.2	78.5	11.8	72.8	14.9	123.6	26.6
Total Income[a]	550.8	100.0	531.4	100.0	662.7	100.0	489.2	100.0	465.4	100.0
Income per Capita	131.9		136.7		167.5		111.4		88.7	

[a]Average per family.

TABLE I-8

NUMBER AND STRUCTURE OF FAMILIES BY SOURCES OF INCOME

Villages	Total	%	Vill. 3	%	Vill. 6	%	Vill. 8	%	Vill. 14	%
1. Agriculture only	103	14.4	25	12.5	52	23.4	7	4.0	19	16.0
2. Salary only	55	7.6	18	9.0	23	10.4	10	5.7	4	3.4
3. Activities only	35	4.9	9	4.5	2	.1	20	11.4	4	3.4
4. Agric. and Salary	287	40.1	111	55.5	106	47.5	17	9.7	53	44.5
5. Agric. and Act.	75	10.5	16	8.0	10	4.5	34	19.4	15	12.6
6. Act. and Salary	28	3.9	3	.2	9	4.1	15	8.6	1	.8
7. All	131	18.3	17	8.5	20	9.0	71	40.6	23	19.3
8. Only sales others	2	.3	1	.05	--	---	1	.6	--	---
9. Total	716	100.0	200	100.0	222	100.0	175	100.0	119	100.0
10. Agriculture [1+4+5+7+8]	598	83.6	170	85.0	188	84.7	130	74.3	110	92.4
11. Salary Work [2+4+6+7]	370	61.9	149	74.5	158	71.2	113	64.6	81	68.1
12. Activities [3+5+6+7]	269	45.0	28	14.0	41	18.5	160	91.4	20	16.8

TABLE I-9

FORMS OF LAND OWNERSHIP

VILLAGE	Total		Village 3		Village 6		Village 8		Village 14	
		%		%		%		%		%
Total families with land. . .	667	100.0	196	100.0	204	100.0	151	100.0	116	100.0
Nuclear	539	80.1	187	95.4	165	80.9	92	60.9	95	81.9
Extended	123	18.4	31	15.8	41	20.1	15	9.9	36	31.0
Rented	318	47.7	114	58.2	86	42.2	59	39.1	59	50.9
Municipal	63	9.4	0	---	1	.5	61	40.4	1	.8
Total families cultivating. . .	582	87.3	170	87%	187	91.7	115	76.2	110	94.8

TABLE I-10

STRUCTURE OF TOTAL LAND BY TYPES OF OWNERSHIP

VILLAGE	Total	Vil. 3	Vil. 6	Vil. 8	Vil. 14
Nuclear land	58.6%	74.5%	74.1%	25.2%	75.7%
Extended	7.0%	2.7%	11.6%	2.9%	9.7%
Rented	20.2%	22.8%	14.3%	37.6%	14.2%
Municipal	4.2%	---	0.0%	34.3%	.4%
Total	100.0%	100.0%	100.0%	100.0%	100.0%

of the mothers weave palm products that they sell to merchants who order them from the city (see Tables I-5 to I-10).

The Two Semi-Urban Villages

The additional data on the two semi-urban villages provide a useful contrast to the four rural villages, and perhaps, a projection to a more "modern" society through which the rural villages may pass.

The two semi-urban villages are only about 20 minutes driving time from Guatemala City, and many inhabitants commute daily to work or to sell goods in the capital. There are also many more job opportunities in these villages and more shops than in the four rural villages.

Literacy is much higher in the semi-urban as compared with the rural villages: of the female heads of household, 55 percent are literate in the semi-urban versus 22% in the rural villages, and the respective percentages for males are 77% and 39%.

Similarly, the number of grades of schooling completed by females and males is more than double for the semi-urban as compared with the rural villages (see Table I-11). Income is much higher in the semi-urban (1974-75) than the rural villages. The mean total family income is 555 quetzales[1] in the rural and 1425 in the semi-urban villages. However, per capita income shows a smaller difference between the two areas: 108 and 206 quetzales for rural and semi-urban.

The occupational structure is quite different in the two areas: 96% of male heads of household are engaged in agriculture in the rural villages whereas only 28% are in the semi-urban villages. Consequently, in the rural villages, the major source of income is agriculture while it is wages in the semi-urban villages.

---

[1] 1 quetzal = 1 dollar

A higher percentage of female heads of household work for pay in the semi-uran than in the rural villages; however, the percent not working is rather high in both: 68% in rural and 62% in semi-urban villages. Note, however, that the majority of females who work in the rural villages are engaged in manufacturing at home, i.e., cottage-industries[1] (see Table I-12).

---

[1] This is prevalent especially in rural village #8 where 68% of female heads of household make palm products at home and sell them for money.

TABLE I-11

Comparison Between the Rural and the Semi-Urban Villages

Variable Name	Rural	Semi-Urban
Mean family income 1974-75 (quetzales)[1]	555	1425
Mean income per capita 1974-75 (quetzales)	108	206
Percent income from agriculture	61	21
Percent income from wages	38	74
Percent income from activities	15	17
Principal occupation of male head of household (%): agriculture	96	28
other	4	72
Principal occupation of female head of household (%): none	68	62
domestic	3	12
manufacturing[2]	16	0
merchant	6	12
specialized laborer, factory, blue and white collar	3	9
other	4	4
Community level schooling (grades)[3]	2.53	6.65
Percent literate female heads of household	22	55
Percent literate male heads of household	39	77
Female head of household's mean number of grades of schooling completed	1.00	2.67
Male head of household's mean number of grades of schooling completed	1.25	3.78
Family size (mean)	4.65	4.68

---

[1] 1 quetzal = 1 dollar

[2] Manufacturing includes cottage industries performed at home.

[3] Mean sum of female and male head of household's grades of schooling completed.

Appendix I

Rand/Rockefeller Source Files[1] Used by the Berkeley Project on Education and Nutrition

Questionnaire #	Description
R09	Time Use of Children
R10	Income and Wealth 1974
R11	Attitudes and Expectations of Women
R12	Attitudes and Expectations of Men
425	Household Census 1975

---

[1] See H. L. Corona (1978) Codebook and User's Manual: INCAP-RAND Guatemala Survey. Santa Monica: The Rand Corporation, p. 6181.

PART II

ANALYTICAL RESULTS FROM THE BERKELEY STUDY

Introduction

In the introduction to this report, we recommended that the effects of malnutrition on child growth and development and school performance need to be investigated. It was recognized that, if links could be shown between childhood nutrition, health, and educational performance, then not only the efficiency of educational investments would benefit from the improvement of nutritional and health conditions, but, also, the allocation of total resources could be made more efficiently.

In Part II, we shall present the results of four sets of analyses using INCAP and Rand data collected in the four study communities in Eastern Guatemala. In II-A, a series of models is used to find out how prenatal and early childhood nutrition, in combination with other socio-environmental factors of family and village, determine children's growth, health, verbal development, and school performance. From these longitudinal models, it appears that children's height is affected by nutritional intake and the incidence of diarrheal infection, and that the child's size at school age is a good measure of the child's nutritional and health history. These are strong and significant results and imply that the improved health and nutritional status of children can affect in important ways the children's future opportunities. It will be seen that children who have grown taller are also more likely to be further advanced in verbal development and are more likely to attend school at an earlier age.

For development planning purposes, these results are important. They enable planners to trace the effects of continuing food intake, in combination with family and environmental factors, to children's functioning at school and work in the villages. Moreover, if the intergenerational effects of

education can be measured by the impact of adult literacy on economic productivity, then school attainment can be related to economic productivity benefits. With such results, development planning can take into account benefits which derive from one sector and affect another; these are known as positive externalities of programs by economists. Better nutrition for children, leading to improved growth, stamina, verbal development, school attainment, work performance, and, utlimately, adult efficiency, combined with appropriate levels of land accessibility and school quality would enhance lives beyond the gains expected by the change in nutrition alone.

Data collected by INCAP and Rand indeed did permit the investigation of relationships between children's school and work activities, parental economic productivity and literacy, and parental literacy and perceptions about desirable family size. We were able to examine the following questions: (1) how immediate needs for work affect family decisions about the desirability of school attendance and the attainment of literacy; (2) the effects on adult economic productivity of literacy attained; (3) the effect of parental literacy on the perceptions of children's economic utility.

These three sets of important relationships are explored in Sections II-B,C,D. Using cross-sectional data as well as the findings of II-A, in II-B parental decisions on sending children to school are examined as well as how nutritional status, parental literacy, family size and structure, and family economic conditions affect educational enrollment and achievement. Using the findings of II-A in II-B, attained height at school age is used as the proxy measure of children's past nutritional and health status. Relationships between prior nutrition and health and school achievement are seen to vary by village and by type of economic activity of the family, indicating that, while the opportunity to attend school may be available to

all, children actually attend school if their parents can spare their work or value the long-run benefits of schooling in relation to the short-run costs. Moreover, controlling for family economic and attitudinal factors, it appears that height at school age does appear to explain school enrollment and achievement. Among families of comparable levels of affluence, in three of the villages, bigger, healthier children do participate in school more often than smaller children.

In Section II-C, family agricultural production is related to land holdings, type of production, and farmer's literacy. Education enabled farmers to choose the best combination of production factors, to introduce modern crops and chemical inputs, and to obtain, in the overall, higher levels of land and labor productivity. The estimation of a production function for different groups of farmers, according to their level of market integration, provides information of the influence of literacy and of an additional year of schooling on increased agricultural production.

In Section II-D, individual fertility behavior and desires are related to perceived economic need for children, family economic activity, and literacy of the heads of the household. These results allow us to see how increasing years of schooling and literacy may affect attitudes toward family size and actual fertility.

SECTION II.A
LONGITUDINAL ANALYSIS OF NUTRITION, GROWTH, VERBAL DEVELOPMENT, AND SCHOOL ACHIEVEMENT

by Alan B. Wilson

I.A.1. Introduction

This part of the report deals with 824 children who were born between January 1969, when data collection was initiated by INCAP in the four study villages, and March 1972, the last cohort to have attained age six when data collection was discontinued and who had lived in the village at any time during their first six years. Many of these children were in the village for only brief periods of time, and this sample includes children who died in infancy or who left the village while still very young. Most of the analyses are based upon that subset of 512 children who were born in one of the four villages and continued to live in the village of birth at least until age three, thus excluding the transient population. The effective sample size for the various tabulations which follow is often further reduced by unavailability of data for the set of variables over the age ranges included in the particular analysis. There is a discussion of the sources and extent of missing data in Appendix A.11.d.

The overall concern of this study is to estimate the effects of nutrition at various ages upon verbal development and school performance. This large topic is broken down into several discrete questions. (1) What are the effects of variations in diet and morbidity upon growth in stature and weight during early childhood? (2) Does a child's cumulative health at age three have short- or long-term effects upon verbal development from age three through age seven? (3) Does concurrent nutrition substantially modify verbal development during childhood? (4) Do verbal development, diet, and health affect school enrollment? (5) Among those who do attend school, do verbal development, diet, and health affect teachers' assessments of performance. This sequence of questions is depicted diagramatically in Figure A.1 in this section.

Variations in children's growth are determined by genetic factors, variations in nutrition at the cellular level, morbidity, and energy expenditure. None of these factors is measured directly, completely, or accurately in the present study. Parental anthropometry, which in turn has been affected by the parents' environments, is available as a proxy for genetic potential for growth; boys and girls are either analyzed separately or sex is included as a control variable to allow for average sex differences in growth; and, the inclusion of lagged values of children's anthropometric measures, for example, height the preceding year, presumably partly controls for prior predisposition. While the residual unexplained individual variance in growth (ranging from 10% to 30% in subsequent analyses) may be due to unmeasured genetic factors, inadequately measured morbidity, and unmeasured patterns of energy expenditure, we assume that the systematic differences between groups of children are largely due to variations in diet and morbidity -- in particular to the effects of the nutritional intervention.

Contrasts between the four villages are reviewed in Section II.A.2. While there are a number of significant demographic differences among the villages, the comparisons which are most germane to our present questions have to do with the pre-intervention nutritional status of the children in the villages. Anthropometric surveys of children conducted in 1968, before the beginning of the nutritional intervention (and before the birth of the first of the cohorts included in the sample studied here), show that there were no large or systematic differences in the average growth patterns of children in the Atole and Fresco villages.

After the initiation of the project, there are substantial differences in growth between children in the Atole and Fresco villages, beginning by age three months and increasing through age thirty-six months. At age

thirty-six months, in the Atole villages, the girls were 11.4% heavier and boys 7.3% heavier than their counterparts in the Fresco villages. These patterns of growth in weight and stature by sex, age, and treatment group are described in Section II.A.4.

While there is evidence that prior to the project the home diets of the children must have been comparable, after the introduction of the supplementation, children in the Atole villages ate considerably less at home. They more than compensated for this reduced home diet, however, by the consumption of supplementation. This suggests a substantial gross replacement of home diet by supplementation in the Atole villages. One clear impact of this replacement in the Atole villages has been to raise the proportion of their total dietary intake arising from protein sources from about 11% (the proportion in the home diets in both Atole and Fresco villages) to well over 13%.

The treatment -- the provision of a high-protein dietary supplement -- had an apparent impact on the growth of children during infancy. While the total caloric intake of the children in the Atole villages was higher than in the Fresco villages until they were about five years old, the effect of the supplementation on the amount of protein consumed was much more pronounced than its effect on total caloric intake. These dietary contrasts are detailed in Section II.A.3. During the second year of life, total caloric intake is estimated to be 17% higher in the Atole villages, and the available protein is over 46% higher.

This initial descriptive contrast raises an additional question as to whether, despite substantial replacement, the differences in total available calories in the diet accounts for the differential growth rates or whether the very large average advantage in available protein in the Atole villages

due to supplementation led to the differences. This question has been explored in two different ways.

First it was reasoned that if the limiting factor for growth in the villages was the deficiency of calories in the diet while protein available from home diet was sufficient, (a) calories from home diet and calories from supplementation should have very similar effects upon growth, and (b) supplementation in the Atole villages, while rich in available protein, and supplementation in the Fresco villages, which contained no proteins, should have similar effects on growth. The contrasts between the effects of supplementation and home diet, and between the effects of supplementation in the two treatment groups, are described in Section II.A.5.

In Section II.A.6, a more direct approach is used. The estimated available protein in the home diets of each child in kcal units is deducted from the total caloric value of the home diets. In the Atole villages, each child's supplementation intake is divided into two portions: 27% of the total being the kcal value of the protein content, the remaining 73% being calories net of protein. For each child, two new variables are reaggregated: dietary protein in kcal units, and calories net of protein, each including both home diet and supplementary sources. The effects of these two variables -- "straight" proteins and calories -- upon growth in stature and weight are estimated at several age levels.

Turning to the linkage between early physical growth and subsequent verbal development, to be discussed in Section II.A.7, there may be important interactions between biological and social factors (cf., Ricciuti, 1977). Early malnutrition and ill health, resulting in poor physical growth, may have direct effects on learning ability and energy available to explore and interact with the environment. These same factors adversely affect

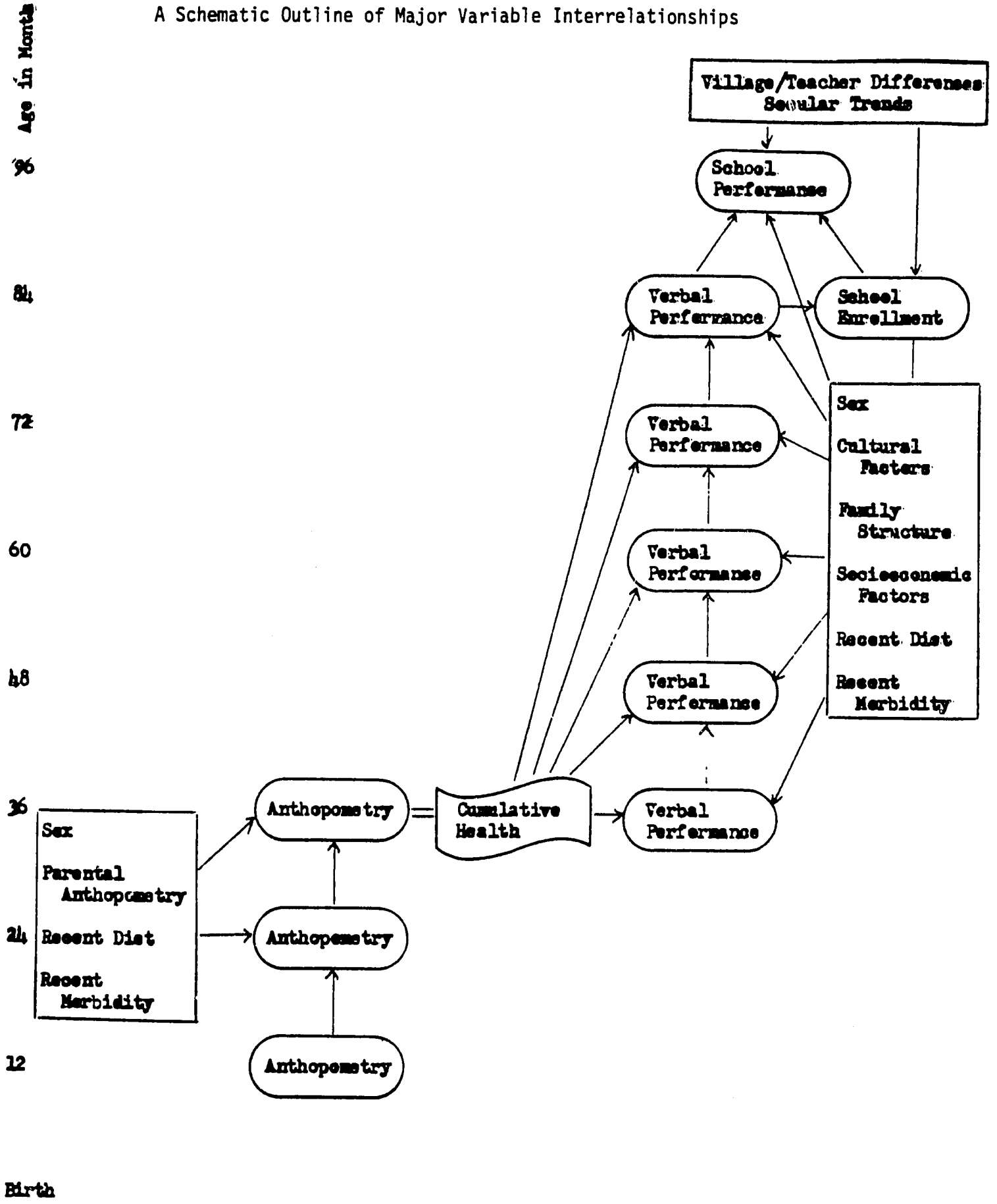
the social responses of others to the child. Similarly, the complex of poverty and illiteracy leads to poorer health care as well as minimal stimulation of learning. In the analysis undertaken here, however, we leave these interactions between biological and social sources of cognitive development as unanalyzed correlations and estimate the direct independent effects of each upon verbal development. In this analysis, we construe stature at a given age as a proxy for cumulated prior nutrition and health.

In Section II.A.8, the effects of verbal development, height, and a large number of social and economic characteristics of the family upon enrollment in school by age seven is estimated for the small group of 114 children in the sample who had reached age seven by January of 1977. Here it is plausible that health, size, and performance may have different effects for different families. For some families, the robust child may be deemed too valuable for the family enterprise to permit school enrollment; for others, the less developed child may be judged incapable of benefitting from schooling. Further analysis of factors affecting parental evaluation of the utility of child labor is undertaken in other portions of this report. In the present section, while we note large differences between villages in rates of school enrollment, we estimate the average effects, across villages and families, of a large number of characteristics.

Finally, for those children who were enrolled in school for at least one year, we examine the effects of both pre-school health and concurrent diet and morbidity during the first year of schooling, upon the teachers' assessment of performance in language and mathematics.

Figure A.1

A Schematic Outline of Major Variable Interrelationships



REFERENCES

Ricciuti, Henry. "Adverse Social and Biological Influences on Early Development." Chapter 12 in Harry McGurk (ed.) Ecological Factors in Human Development. New York: North-Holland, 1977.

II.A.2. Pre-Intervention Contrasts Between Treatment Groups and Villages Within Treatment Groups

While the four villages in this study were selected because of their common features -- isolation from urban areas, Ladino culture, agricultural base -- there are some significant contrasts between these villages which bear upon the analysis and interpretation of the data.

The two villages receiving the Atole supplementation have a far higher proportion of households who own their own land than the other two villages. This is clear in the following distribution of principal occupations of male heads of households.

TABLE A.1

PERCENTAGE DISTRIBUTION OF PRINCIPAL OCCUPATIONS OF MALE HEADS OF HOUSEHOLDS BY TREATMENT AND VILLAGE SIZE

Occupation	Fresco		Atole	
	Small	Large	Small	Large
Farmers of own land	27.9%	26.5%	50.0%	49.7%
Farmers of family land	0.0	0.6	6.2	3.6
Sub-total	27.9	27.1	56.2	53.3
Renting farmer	21.7	53.6	32.1	33.2
Agricultural laborers	28.7	4.7	3.6	6.7
Farmer of communal lands	8.5	0.6	0.9	0.0
Sub-total	58.9	58.9	36.6	39.9
All non-farm occupations	13.2	14.0	7.2	6.8
Number of Cases	129	140	112	193

This contrast shown in Table A.1 above is also manifest in the "PASTAT" factor scale (see Appendix A.11.C). This scale not only includes the father's current and usual occupations (which were themselves scored to favor land ownership), but also land owned itself, value of land planted, value of major crops, value of animals owned, and home ownership.

TABLE A.2

MEAN "PASTAT" FACTOR SCORES BY TREATMENT AND VILLAGE SIZE

Village Size	Fresco	Atole
Small	-.27 (100)	.15 (109)
Large	.03 (127)	.24 (155)
Total	-.10 (227)	.20 (264)

(Probabilities -- Treatment: .001; Size within treatment: .025*.)

While the Atole villages are clearly favored with regard to ownership of land, evidence on other village characteristics is more mixed. Three areas are discussed below: prior nutrition, prior morbidity, and other socioeconomic characteristics.

Nutrition - The lack of base-line data about the pre-intervention nutritional status in the villages, particularly the levels for earlier cohorts of children, has been one of the data gaps which has made interpretation

*Note: This is a hierarchical (nested) analysis of variance. Here, and in other tables in this section, probabilities for treatment, size within treatment, and/or interaction are reported if less than .05.

of the overall impact of the project difficult. There is, however, some evidence in the data base which suggests that the pre-1969 nutritional status was at least as favorable, perhaps more favorable, in the Atole villages.

While the home diet data for the mothers during pregnancy and lactation is spotty, for the 512 children who were born in the village and continued to live in the village through at least their third birthday, home diet estimates are available for two-thirds of the sample for the third trimester of pregnancy. Breaking these (daily) estimates down by cohort as well as treatment shows a trend in the Atole villages of declining home diets.

TABLE A.3

TOTAL KCAL VALUE OF MOTHERS HOME DIETS DURING THE THIRD TRIMESTER OF PREGNANCY BY TREATMENT AND DATE OF BIRTH OF CHILD

Date of child's birth	Fresco	Atole
1/69 - 10/69	1684 (24)	2117 (20)
11/69 - 8/70	1268 (23)	2128 (38)
9/70 - 6/71	1314 (38)	1617 (39)
7/71 - 4/72	1404 (32)	1413 (38)
5/72 - 3/73	1428 (38)	1230 (47)

The unreliability of these home diet data (which is to be discussed and estimated elsewhere) and the incompleteness of the sample of data, make even the rather substantial contrasts between the mothers' diets during the first part of the project uncertain. The suggestion, though, that the mothers in the Atole villages had enjoyed better nutrition before the intervention is congruent with their anthropometry.

TABLE A.4

MOTHERS' HEAD CIRCUMFERENCES IN CENTIMETERS
BY TREATMENT AND VILLAGE SIZE

Village Size	Fresco	Atole
Small	50.39 (100)	51.06 (99)
Large	50.81 (121)	51.39 (136)
Total	50.61 (221)	51.25 (235)

(Probabilities -- Treatment: .001;
Size within treatment: .002.)

A similar contrast is found in the somewhat smaller sample of measurements of fathers' head circumferences favoring the Atole villages, and within treatment groups, the larger villages.

TABLE A.5

FATHERS' HEAD CIRCUMFERENCES IN CENTIMETERS
BY TREATMENT AND VILLAGE SIZE

Village Size	Fresco	Atole
Small	53.25 (73)	54.26 (98)
Large	54.04 (99)	54.61 (104)
Total	53.70 (172)	54.44 (202)

(Probabilities -- Treatment: .001;
Size within treatment: .001.)

On the whole, the larger home diets of the pregnant mothers in the Atole in the early months of the project, together with the larger head

circumferences of both parents in these villages suggests a prior advantage. The declining home diets in these villages during the project period is partially off-set by their larger intake of the supplementation. Because of the high-protein content of the supplement, the mothers in the Atole villages in the later cohorts continued to enjoy a larger protein content in their total diets. The proportion of total caloric intake due to protein indeed increased among the later cohorts in the Atole villages.

TABLE A.6

KCAL VALUES OF MOTHERS' DIETS (HOME DIET AND SUPPLEMENTATION COMBINED) BY SOURCE, TREATMENT, AND COHORT DURING LAST TRIMESTER OF PREGNANCY

Date of child's birth	Calories		Protein		Total	
	Fresco	Atole	Fresco	Atole	Fresco	Atole
1/69 - 10/69	1509	1944	210	254	1720	2198
11/69 - 8/70	1158	1930	147	251	1305	2181
9/70 - 6/71	1279	1520	159	216	1439	1736
7/71 - 4/72	1337	1371	179	215	1516	1586
5/72 - 3/73	1356	1213	180	203	1537	1416

Cross-sectional anthropometric data for the study villages which were gathered in 1968, before the nutritional and health interventions were implemented, shows that the children in those villages which were to receive the Atole supplement had slightly larger head circumferences

and skinfolds, and marginally larger weights and heights, than those in the villages which were to receive the Fresco treatment when classified by age and sex (Martorell: IIA, Table 2). The contrasts between treatment groups in head circumferences by age and sex are illustrated in Figure A.2. Since these are the means of age groups in a cross-sectional study rather than growth curves in a longitudinal study, the patterns are somewhat irregular. Yet, at each age level, the children in the Atole villages were larger with the exception of males at age 12 months.

The pre-intervention contrasts in children's anthropometry, the contrasts in parental head circumferences, and the pattern over time of pregnant mothers' home diets all suggest that the home diets enjoyed by families in the Atole villages were at least as sufficient and probably somewhat more adequate than those in the Fresco villages.

Morbidity

While mothers in the Atole villages, on the average, apparently enjoyed larger diets prior to the project period, as shown in Tables A.3 and A.4, they suffered much higher rates of infant mortality. The number of pregnancies of the female heads of household, as recorded in the Rockefeller/Rand survey of 1975, is tallied below in Table A.7. There is, of course, a very large variation between women within villages. However, the average difference of one child is highly significant.

Figure A.2

Mean Head Circumference of Children in 1968
by Treatment Group, Sex, and Age

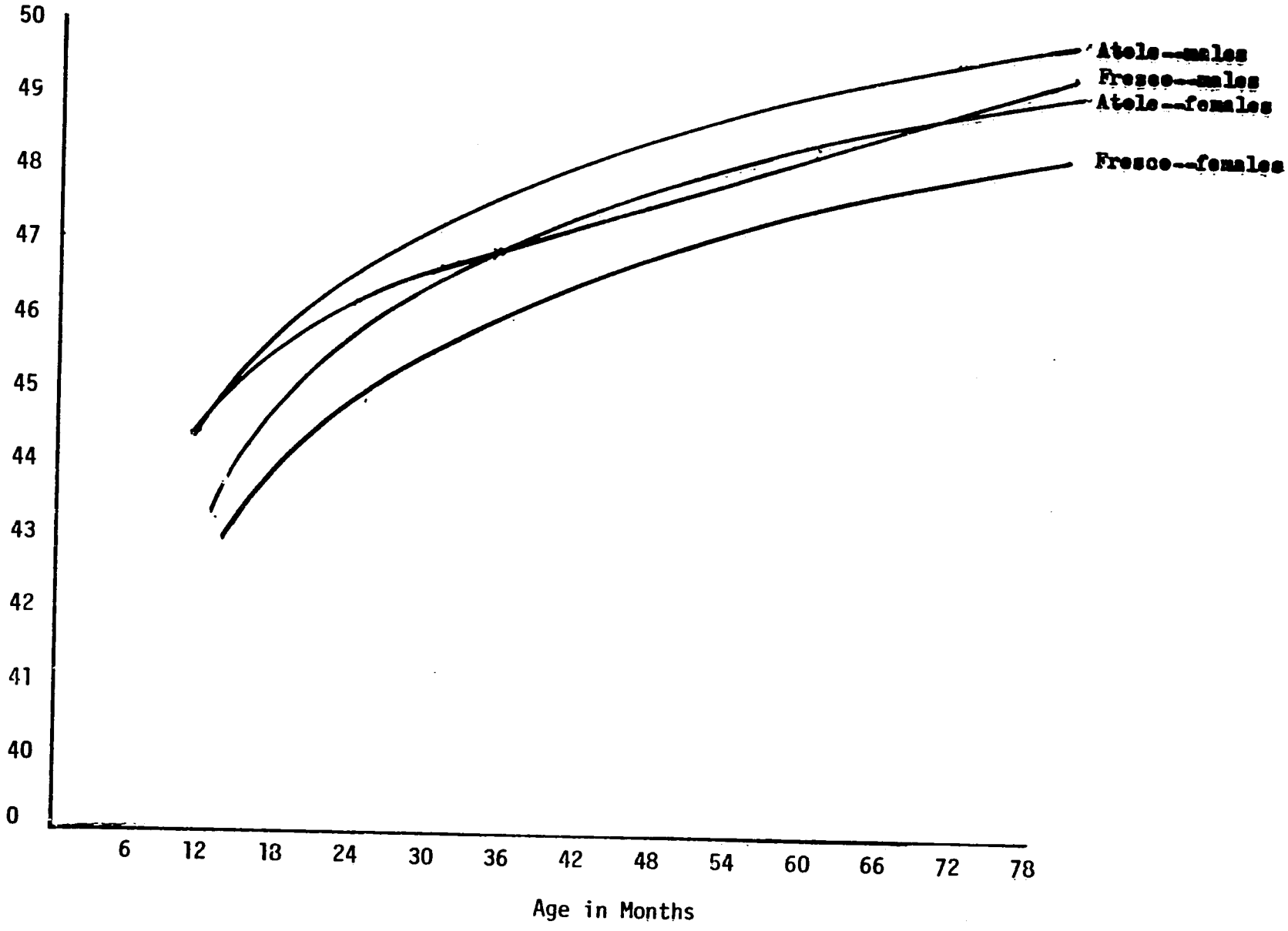


TABLE A.7

NUMBER OF PREGNANCIES OF FEMALE HEAD OF HOUSEHOLD BY TREATMENT AND VILLAGE SIZE

Village Size	Fresco	Atole
Small	5.09 (100)	6.37 (110)
Large	6.12 (127)	6.82 (158)
Total	5.67 (227)	6.63 (268)

(Probabilities -- Treatment: .001; Size within treatment: .019.)

However, despite this higher average rate of pregnancies, the actual family sizes at the time of the birth of a child during the study years shows insignificant difference between treatment groups.

TABLE A.8

FAMILY SIZE AT THE BIRTH OF CHILD IN STUDY

Village Size	Fresco	Atole
Small	5.00 (104)	5.40 (112)
Large	5.79 (135)	5.76 (161)
Total	5.45 (239)	5.62 (273)

(Probabilities -- Size within treatment: .003.)

While the discrepancy between Tables A.7 and A.8 could be due to different rates of mobility of members of the household, a close study of household composition and mobility shows that it is the number of children who died -- usually in infancy and mostly before the initiation of the project -- who account for the discrepancy.

TABLE A.9

NUMBERS OF DECEASED ELDER SIBLINGS OF CHILD AT CHILD'S BIRTH BY TREATMENT AND VILLAGE SIZE

Village Size	Fresco	Atole
Small	.60 (104)	1.02 (110)
Large	.59 (133)	1.28 (160)
Total	.59 (237)	1.17 (270)

(Probabilities -- Treatment: .001.)

The introduction of health services in the villages, which INCAP has reported to have drastically reduced infant mortality, had its largest benefit in the Atole villages which initially had the highest rates. By the time these children reached age six the mean number of deceased elder siblings in the Fresco and Atole villages were .63 and 1.20 respectively.

The numbers of deceased elder siblings, especially in the Atole villages, probably does reflect inadequate public health services and sanitation. The variable also reflects variation between families within villages in the nurture of children, as we shall see below when we examine factors affecting home diet.

Other socioeconomic contrasts - Male heads of households in the Fresco villages included a higher proportion who were literate.

TABLE A.10

MEAN LITERACY SCORES OF MALE HEADS OF HOUSEHOLDS BY TREATMENT AND VILLAGE SIZE (AVERAGE OF Z-SCORES OF READING AND YEARS OF SCHOOLING: "ZPARD")

Village Size	Fresco	Atole
Small	.60 (91)	-.24 (109)
Large	.04 (118)	-.12 (148)
Total	.28 (209)	-.17 (257)

(Probabilities -- Treatment: .001; Size within treatment: .029; Interaction: .001.)

In the factor scale, "CONSUMP", which is composed of items describing the physical qualities of housing, enumerating certain possessions in the household, and including household cash salaries, there is a clear interaction between treatment and village size.

TABLE A.11

MEAN "CONSUMP" FACTOR SCORES BY TREATMENT AND VILLAGE SIZE

Village Size	Fresco	Atole
Small	-.60 (100)	+.26 (110)
Large	+.43 (127)	-.18 (158)
Total	-.02 (227)	-.00 (268)

(Probabilities -- Size within treatment: .003; Interaction: .001.)

A rather large number of demographic characteristics of the villages, some presumably prior to and others during the course of the project period, did not reveal any substantial differences between villages. These included the proportion of nuclear families, the distribution of children over cohorts, the numbers of younger siblings born by age 6, migration, mother's health, the age of the mothers at the birth of the child, and others.

The decline over time in the total kcal value of mothers' diets in the Atole villages, while retaining their relative advantage in consumption of protein through the use of supplementation, shown in Table A.6, suggests an increasing pattern of replacement. The likelihood of initial dietary advantage, the possibility of replacement, and the predominant role of home diet in the nutritional intake of the mothers, underscore the necessity of including home diets in any subsequent effort to estimate nutritional effects.

The existence of significant village contrasts with regard to variables which may have significant causal impact--even as proxies for unmeasured or unknown factors, as the numbers of deceased elder siblings may reflect differences in hygiene or nutrition--require that the analysis be done either within villages or, preferably, since there is substantial variation within villages, these variables be included in analyses as "controls."

Reference:

Martorell, R., no date, "Article IIA: Food Supplementation and Growth from Birth to Seven Years of Age." Table 2, "Anthropometric Characteristics of the Study Villages in 1968." (Mimeographed.)

II.A.3. Children's Diets

This section consists of a single tabulation of estimated kcal values of daily diets of the group of 512 children in the study sample who were born in the village and still living in the village at age 3. The reason for the selection of this sub-sample of the total data set is to exclude transients, whose values might distort point estimates, while using all available data for the more limited stable population in estimating each mean.

Because psychometry, school enrollment, school test scores, and, at the later ages, anthropometry, are measured at annual intervals, and these are the major dependent variables in the subsequent analysis, the dietary data were also aggregated to an annual level, then reduced to an average daily level. E.g., the total kcal value of supplementation was aggregated from age 12 months to age 24 months, then divided by 365.25 to give a daily average--attributed to 18 months.

For the home diet data, all available measures were used to generate these point estimates. The 18 month estimates, for example, are the average of the measures at 18 months with the average of the sum of the 15 month and 21 month measures.

Age of diet estimate	Measures averaged
18 mo.	(18 mo. + (15 mo. + 21 mo.)/2)/2
30 mo.	(30 mo. + (27 mo. + 33 mo.)/2)/2
42 mo.	(42 mo. + (36 mo. + 48 mo.)/2)/2
54 mo.	54 mo.
66 mo.	(60 mo. + 72 mo.)/2
78 mo.	(72 mo. + 84 mo.)/2

Care was taken in selecting the measures to be averaged not to include the same measure in two successive estimates lest a spurious correlation be generated leading to an overestimate of the stability and reliability of home diets in longitudinal analyses. (The exception, here, is in the 78 mo. estimate which uses the 72 mo. measure. However, since there were so few cases available for that age level, this estimate is not used elsewhere.) The only home diet measure which was not used was the 24 month measure. There was no balancing 12 month home diet measure to include it in estimating the 18 month consumption, and the 36 month measure was required for estimation of 42 month consumption.

The units for the measurement of protein in both home diets and supplementation were transformed from grams to kcal units. The kcal value of the protein was then subtracted (at the individual level) from the total kcal value to generate the variable labelled "calories" in this table and in subsequent analyses.

The attrition to the numbers of cases for whom estimates are available among those 512 children after age 66 months is due to the fact that we have aggregated cohorts born from January, 1969 to March, 1972. The later born children had not reached ages 60, 72, and 78 months when data collection was discontinued in March, 1977.

In examining the tabulation we can see that (1) the home diets of the children in the Fresco villages is consistently larger than in the Atole villages--though not much, and contains about the same proportion of protein; (2) children in the Atole villages consumed more supplementation, and, of course, were the only ones to obtain additional protein thereby; (3) the total dietary kcal was higher in the Atole villages during the earlier ages (18 to 42 months), but was exceeded by diets in the Fresco villages at later ages (66 and 78 months).

At all ages, due to the consumption of Atole, both the absolute value and the percent of total diet which is derived from proteins is higher in the Atole villages.

The values in the table which seem to be severely out of line are those at age 66 months in the home diets (and hence total diets) in the Atole villages. This may be a cohort effect among the 85 children who had reached age 72 months by March, 1977.

In any event, mainly because of the attrition to the sample at the older age levels, most of the analysis which entails the use of these home diet estimates is based upon the first four which seem, in the aggregate, to be orderly.

The analysis of school enrollment and attainment will, of necessity, be based upon the earlier cohorts who have reached age 6 or 7 by 1978.

Selected columns from Table A.12 are shown in graphical form in Figures A.3 and A.4. While the home diets of children in the Atole villages prior to the initiation of the project were probably larger than those in the Fresco villages, after the initiation of the intervention, they declined to about 93% of the total energy value of those in the Fresco villages. This suggests an energy replacement of home diet by supplementation of somewhat more than 7% of the home diet.

However, as Figure A.3 illustrates, the total energy intake in the Atole villages after intervention is only slightly higher -- about 4% higher on the average -- than that in the Fresco villages. This could well be similar to pre-intervention differences which would suggest that there was virtually no real differential supplementation in total energy consumption.

The large contrast in nutritional intake between the Atole and Fresco villages is illustrated in Figure A.4 which shows that the consumption of

protein in the Atole villages runs 30% higher than in the Fresco villages. This advantage, which is even larger during the earliest months of life, is due to the high protein content of the Atole supplement.

TABLE A.12

CHILDREN'S DIETS IN KCAL UNITS BY AGE AND SOURCE

A. HOME DIET

	Calories		Protein		Total		% Protein		Number of Cases	
AGE	ATOLE	FRESCO	ATOLE	FRESCO	ATOLE	FRESCO	ATOLE	FRESCO	ATOLE	FRESCO
18	544	550	58	65	602	615	9.6	10.5	229	226
30	736	786	89	99	825	885	10.8	11.1	229	223
42	871	900	108	116	980	1016	11.0	11.4	234	228
54	980	1024	124	132	1104	1156	11.2	11.4	195	183
66	996	1142	128	146	1124	1289	11.4	11.4	85	72
78	1169	1262	145	161	1314	1424	11.1	11.3	30	26

B. SUPPLEMENTATION

	Calories		Protein		Total		% Protein		Number of Cases	
AGE	ATOLE	FRESCO	ATOLE	FRESCO	ATOLE	FRESCO	ATOLE	FRESCO	ATOLE	FRESCO
18	99	13	37	0	136	13	27%	0%	229	226
30	122	32	46	0	168	32	27	0	229	223
42	119	50	45	0	164	51	27	0	234	228
54	106	68	41	0	148	68	27	0	195	183
66	105	82	39	0	145	82	27	0	85	72
78	117	102	44	0	161	102	27	0	30	26

C. TOTAL DIET

	Calories		Protein		Total		% Protein		Number of Cases	
AGE	ATOLE	FRESCO	ATOLE	FRESCO	ATOLE	FRESCO	ATOLE	FRESCO	ATOLE	FRESCO
18	641	563	95	65	735	628	12.9	10.4	229	226
30	857	819	136	99	993	917	13.8	10.9	229	223
42	990	950	153	116	1143	1066	13.5	11.1	234	228
54	1086	1092	165	132	1252	1224	13.4	11.0	195	183
66	1101	1224	167	146	1269	1370	13.3	10.8	85	72
78	1286	1364	190	161	1476	1525	13.0	10.7	30	26

Figure A.3

KCAL Values of Supplementation and of Total Diet
by Treatment Group and Age

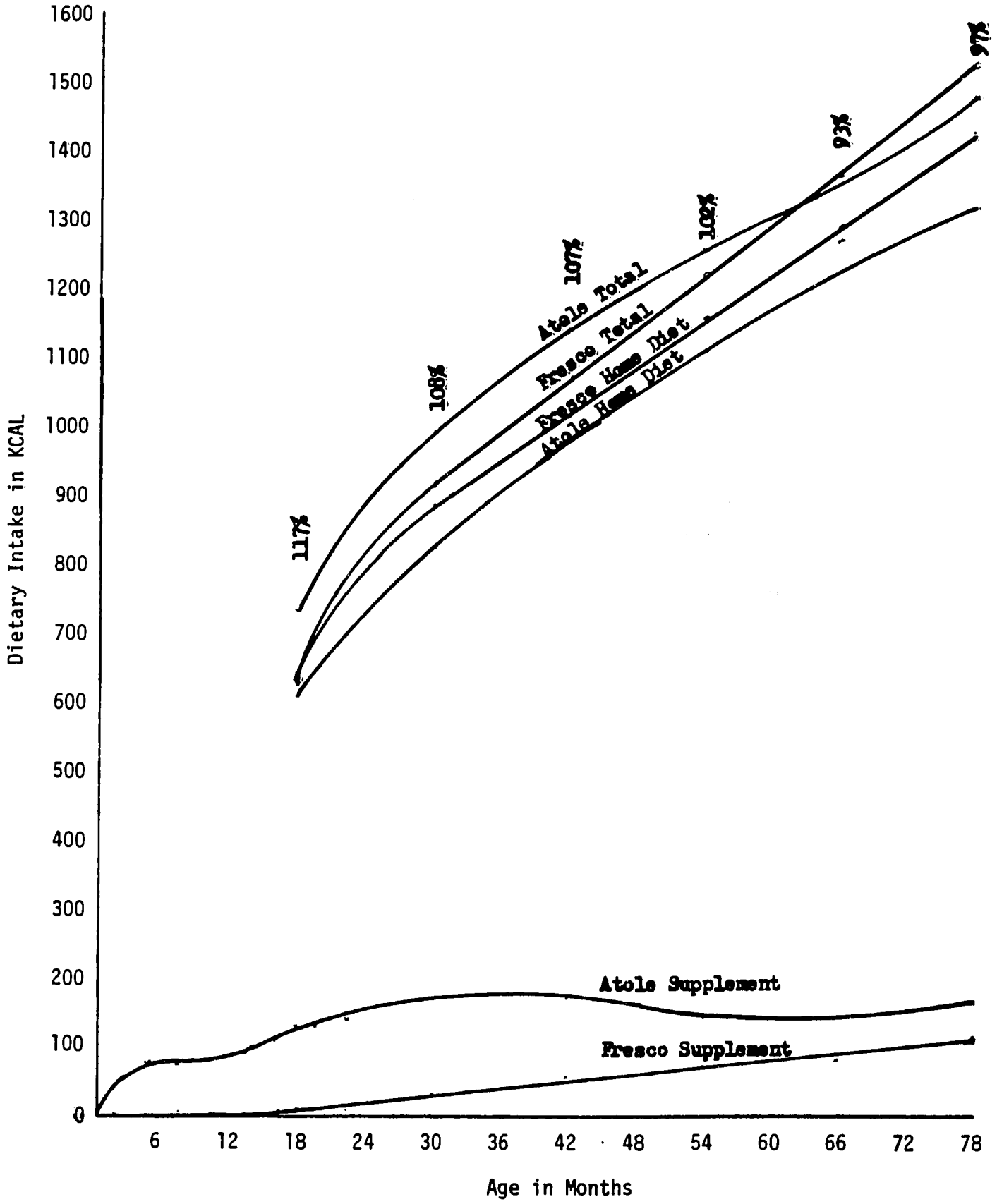
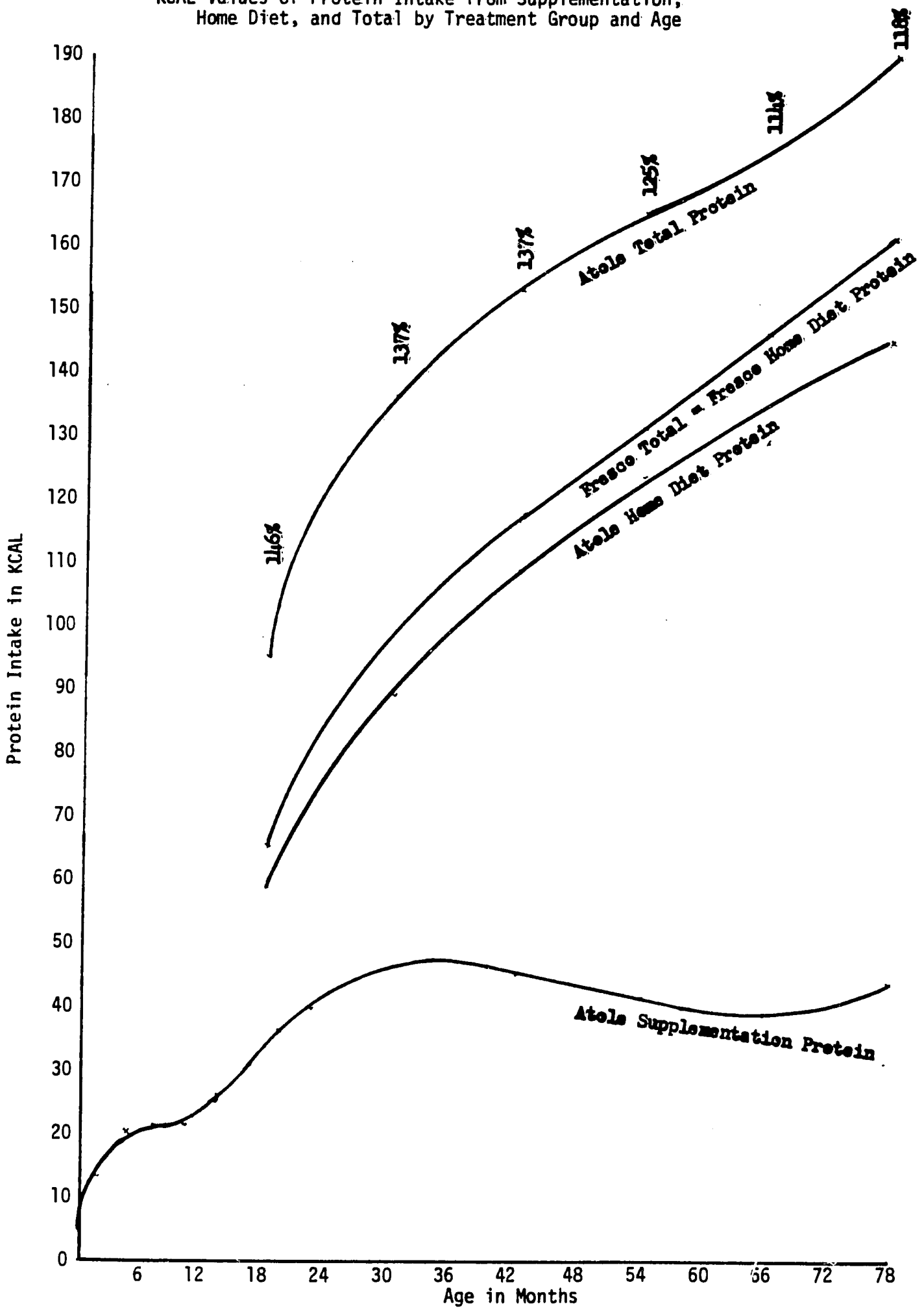


Figure A.4
KCAL Values of Protein Intake from Supplementation, Home Diet, and Total by Treatment Group and Age



II.A.4. Physical Growth by Sex and Treatment

This section consists of eight figures displaying growth patterns (weight and height) by sex and treatment for the ages 0-36 months, and then, on a smaller scale, for ages 24-84 months. The percentile growth charts prepared for the U.S. by the National Center for Health Statistics (Monthly Vital Statistics Report (HRA) 76-1120, Vol. 25, No. 3 Supplement, June 22, 1976) are used to provide a comparative background for this display. Also, the two tabulations of means (of heights and weights) by sex, treatment, and age which were used to plot values on the charts are included as Tables A.14 and A.15.

Once again, the subset of 512 children who were born in the villages, and present at age 3, were used as a data base to assure that, at least through age 5, despite some missing data, the estimates are based on roughly the same group of children -- not a shifting population of transients.

These charts clearly portray the deficit in growth patterns of the entire study sample as compared with the U.S. Figures A.7 and A.8 show the striking deacceleration in growth in weight which occurs between the ages of 3 and 9 months. After age 24 months, the average rate of growth -- the slopes of the lines -- is not too dissimilar from the patterns for the U.S., but at a lower level.

Each of the charts shows that by age 12 months, the children in the Atole villages have gained a clear advantage in both weight and height over children in the Fresco villages. For most contrasts, there is a growing

gain so that the difference in intercepts increases with increasing age. The exception, shown in Figure A.12, is that the boys in the Fresco villages show some recovery in weight between ages 3 and 7, moving from below to about the U.S. 5th percentile, and converging slightly with the boys in the Atole villages who lie close to the 10th U.S. percentile.

An interesting substantive contrast which emerges first in these descriptive figures, and will reappear in subsequent analyses, is a sex x treatment interaction. Comparing Figures A.5 with A.6, and A.7 with A.8, it is apparent that the divergence between the Atole and Fresco groups is substantially larger among girls than among boys. This is also true of the later contrasts, but it is not so vivid because of the smaller scale of the later charts.

In the Fresco villages, the consumption of supplementation is a negligible portion of the children's total diets -- they must subsist on what they are provided with at home. This leads to a competition for scarce resources -- a competition which is, in part, governed by cultural values. The values of the parents as wel as the expectations, assertiveness, docility, and strength of the children themselves advantage boys in this competition.

In the Atole villages, on the other hand, the competition is attenuated by the presence of a "free good" -- the rich supplementation which is available to all.

My interpretation here goes speculatively beyond the data (though later we will find that there is less sex difference in home diets in the Atole then in the Fresco villages). While speculative, this interpretation suggests a possible latent consequence of the nutritional intervention experiment on the sex role expectations of girls in the Atole villages.

The recovery in weight of the Fresco village boys (shown in Figure A.12), but not of the girls (shown in Figure A.11), is both consistent with the above interpretation and also the increased average calorie intake of older children in the Fresco villages shown in the tabulation of children's diets.

If we contrast the home diets at age 18 months by sex and treatment, we find an analogous interaction. This is shown in Table A.13, below. In the Atole villages, the girls are eating almost (98.5%) as much as the boys at home; in the Fresco villages, they are eating considerably less (89.0%) than the boys. While the sex treatment interaction effect on home diets is less at later ages -- girls tending to get about 90% of the boys' home diets in both treatment groups -- we shall see later, in Section A.5 and A.6, that this early diet is particularly crucial for growth.

Table A.13

Total Home Diets at Age 18 Months by Sex and Treatment

Sex	Treatment	
	Atole	Fresco
Male	606	645
Female	597	574
Female as % of Male	98.5%	89.0%

TABLE A.14

MEAN HEIGHT BY AGE, SEX, AND TREATMENT AMONG CHILDREN BORN IN THE VILLAGES AND PRESENT AT AGE 36 MONTHS

	GIRLS				BOYS			
	Fresco		Atole		Fresco		Atole	
Age	Mean	N	Mean	N	Mean	N	Mean	N
0	47.10	5	47.54	7	48.97	15	47.81	15
.5	48.52	45	48.64	69	49.48	69	49.22	65
3	55.96	81	56.08	98	57.43	102	57.49	94
6	61.22	77	61.65	95	62.66	114	63.48	106
9	64.55	79	65.30	103	65.82	118	66.79	114
12	67.31	84	68.13	95	68.65	119	69.73	112
15	69.24	80	70.58	98	71.03	117	72.08	113
18	71.56	83	73.13	95	73.10	122	74.25	111
21	73.71	86	75.42	100	75.21	123	76.44	106
24	75.70	86	77.43	106	77.10	124	78.53	117
30	79.16	83	82.13	104	80.97	125	83.02	108
36	83.06	84	86.32	98	85.21	118	87.03	111
42	87.13	85	90.05	103	88.65	124	90.84	114
48	90.48	84	93.78	98	92.36	117	94.41	110
60	97.03	69	100.16	86	98.40	93	100.32	83
72	103.46	41	105.27	58	103.95	61	105.95	52
84	107.78	20	111.05	26	109.40	26	111.50	26

TABLE A.15

MEAN WEIGHTS BY AGE, SEX, AND TREATMENT AMONG CHILDREN BORN IN THE VILLAGES AND PRESENT AT AGE 36 MONTHS

	GIRLS				BOYS			
	Fresco		Atole		Fresco		Atole	
Age	Mean	N	Mean	N	Mean	N	Mean	N
0	2.98	69	3.09	88	2.98	97	3.19	98
.5	3.15	45	3.34	69	3.20	69	3.47	66
1.5*	3.96	90	4.08	108	4.12	117	4.44	114
3	5.04	81	5.21	98	5.31	101	5.64	94
4.5*	5.70	93	5.94	108	6.00	123	6.35	119
6	6.26	77	6.56	96	6.64	113	7.02	107
7.5*	6.62	88	6.93	110	6.94	127	7.35	119
9	6.97	79	7.18	103	7.25	118	7.68	113
10.5*	7.07	95	7.50	111	7.52	129	7.91	121
12	7.22	85	7.71	99	7.69	121	8.11	112
13.5*	7.45	93	7.96	108	7.96	128	8.33	121
15	7.65	80	8.19	99	8.22	117	8.57	115
16.5*	7.81	93	8.45	108	8.34	129	8.86	120
18	8.07	83	8.64	97	8.57	123	9.07	114
19.5*	8.19	93	9.07	107	8.75	133	9.34	119
21	8.40	88	9.34	101	8.97	123	9.54	107
22.5*	8.63	87	9.54	107	9.30	124	9.88	119
24	8.90	87	9.84	107	9.59	124	10.19	119
27.0*	9.37	95	10.47	114	10.08	133	10.83	122
30	9.86	83	11.01	105	10.58	125	11.40	109
33.0*	10.32	94	11.52	112	11.14	133	11.93	122
36	10.81	91	12.04	106	11.64	132	12.49	117
39.0*	11.32	96	12.45	114	12.01	133	12.87	123
42	11.83	85	12.99	103	12.49	124	13.33	114
45.0*	12.18	94	13.38	110	12.97	131	13.78	123
48	12.56	89	13.78	103	13.43	121	14.19	117
54.0*	13.34	94	14.70	109	14.17	127	14.99	121
60	14.13	69	15.42	86	14.84	93	15.61	84
66.0*	14.87	70	16.28	86	15.63	96	16.37	86
72	15.55	41	16.88	59	16.39	61	17.29	52
78.0*	16.44	41	17.91	61	17.35	64	18.40	54
84	16.99	20	18.81	26	18.38	26	19.49	26

*Interpolated values

Figure A.5

Girls Length by Age Percentiles: Ages Birth-36 Months

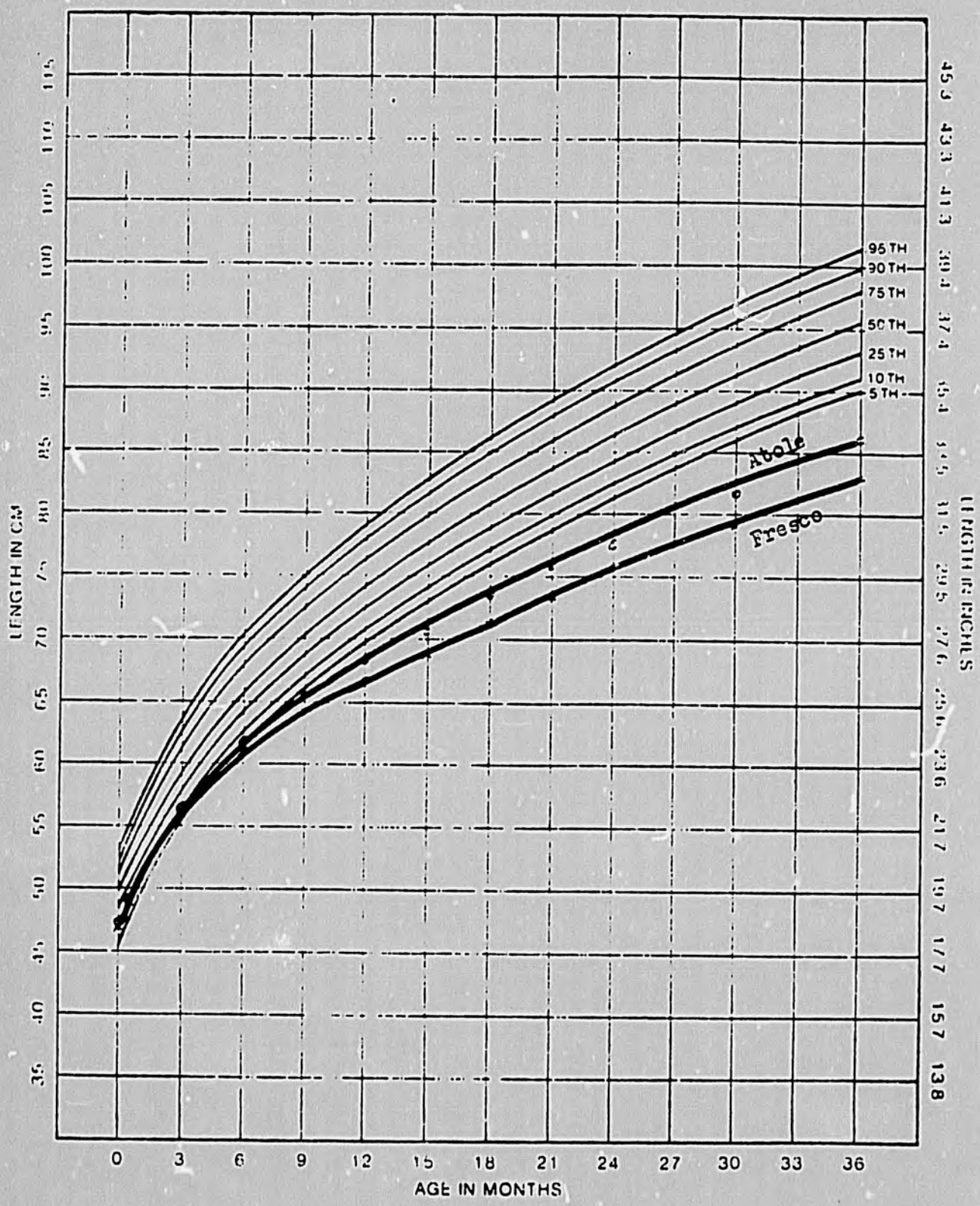


Figure A.6

Boys Length by Age Percentiles: Ages Birth-36 Months

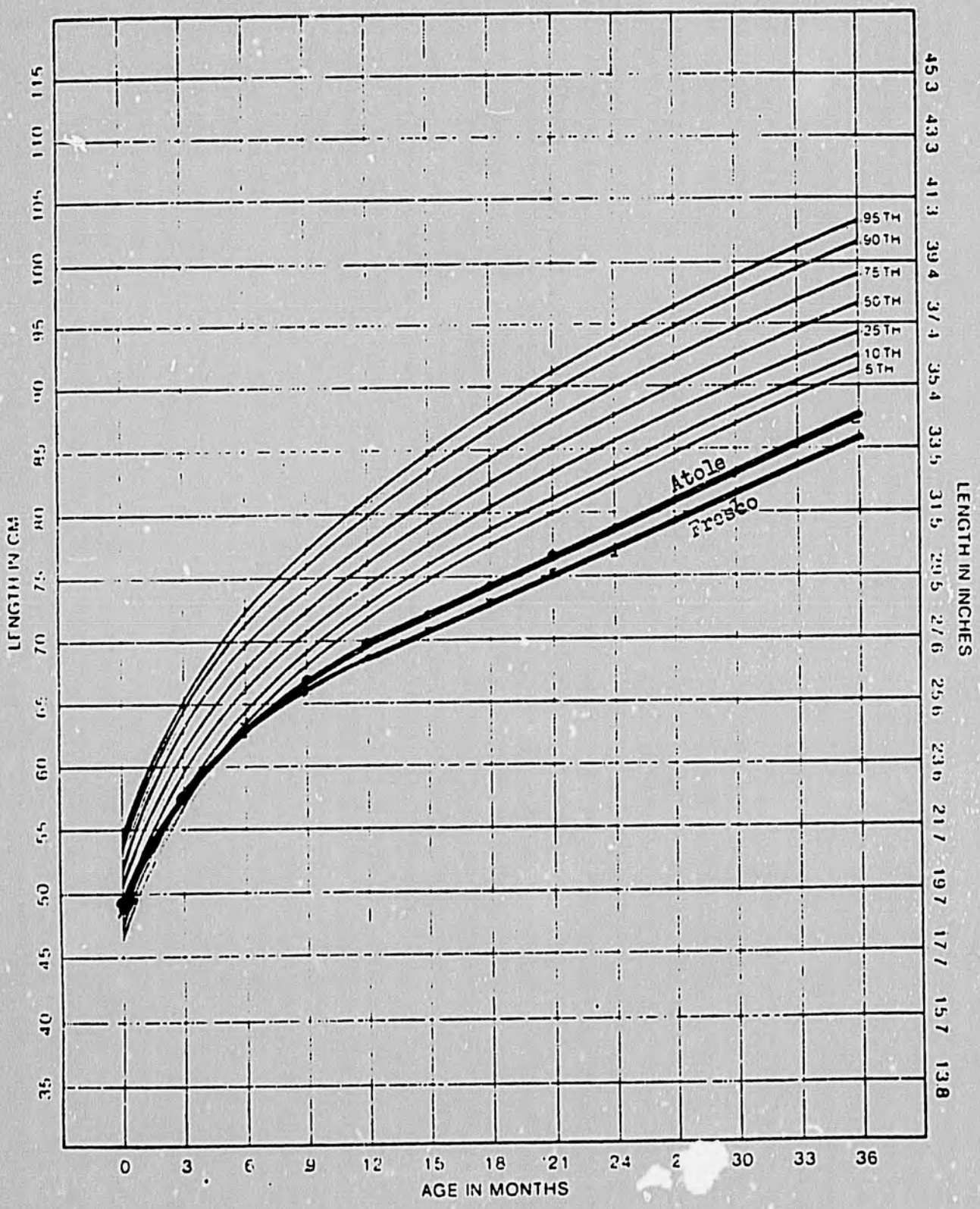


Figure A.7

Girls Weight by Age Percentiles: Ages Birth-36 Months

Figure A.8

Boys Weight by Age Percentiles: Ages Birth-36 Months

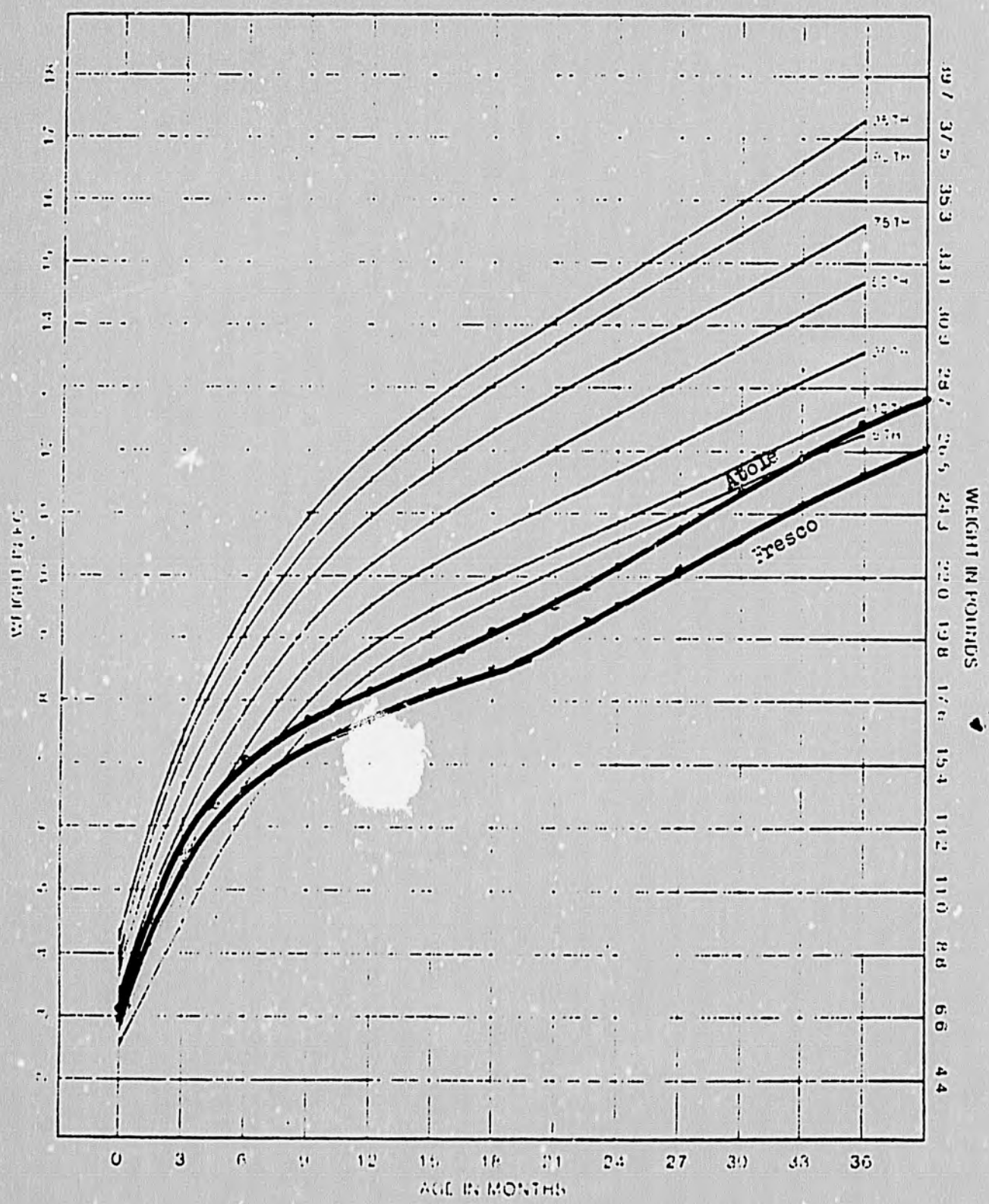


Figure A.9

Girls Stature by Age Percentiles: Ages 2 to 18 Years

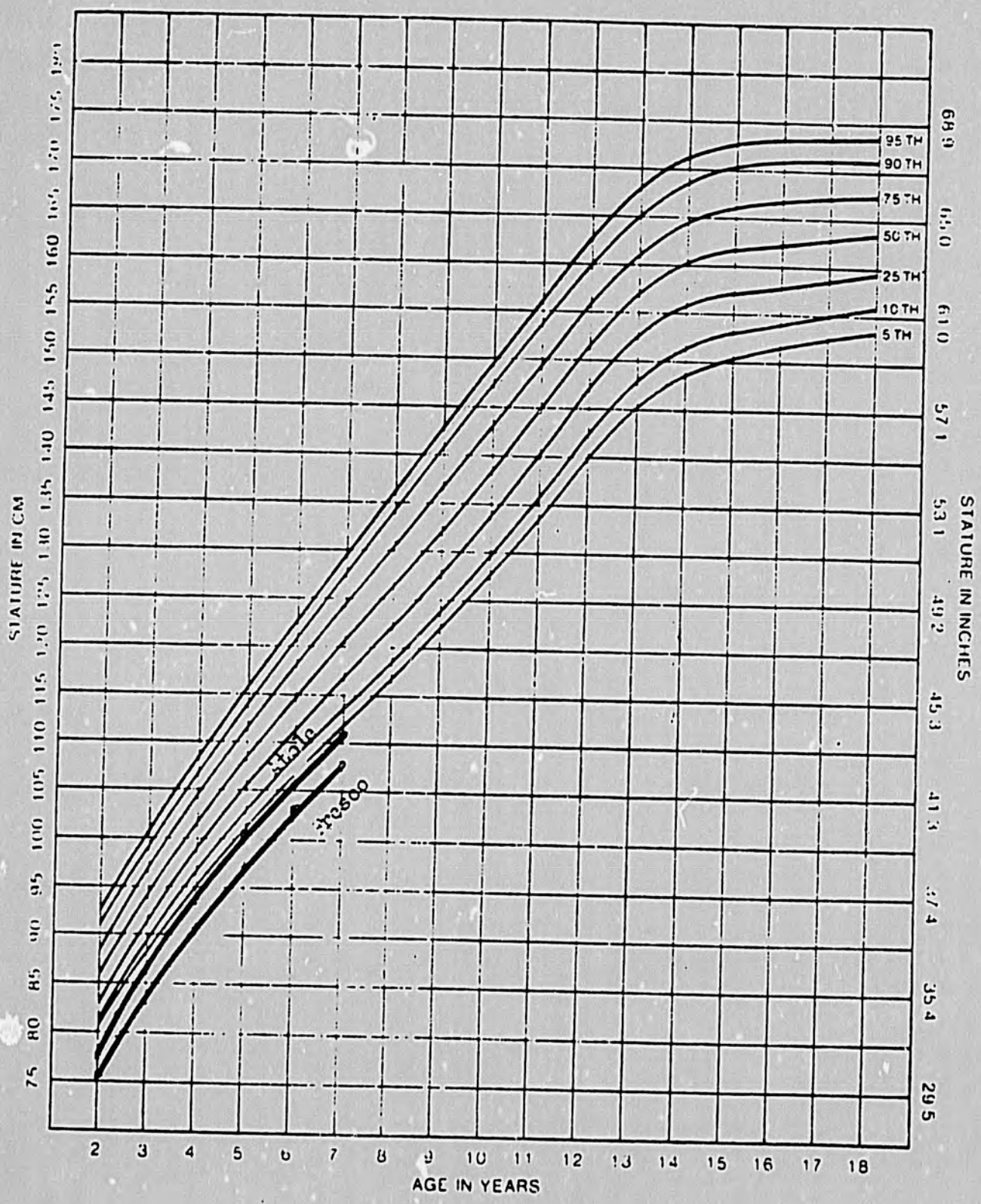


Figure A.10

Boys Stature by Age Percentiles: Ages 2 to 18 Years

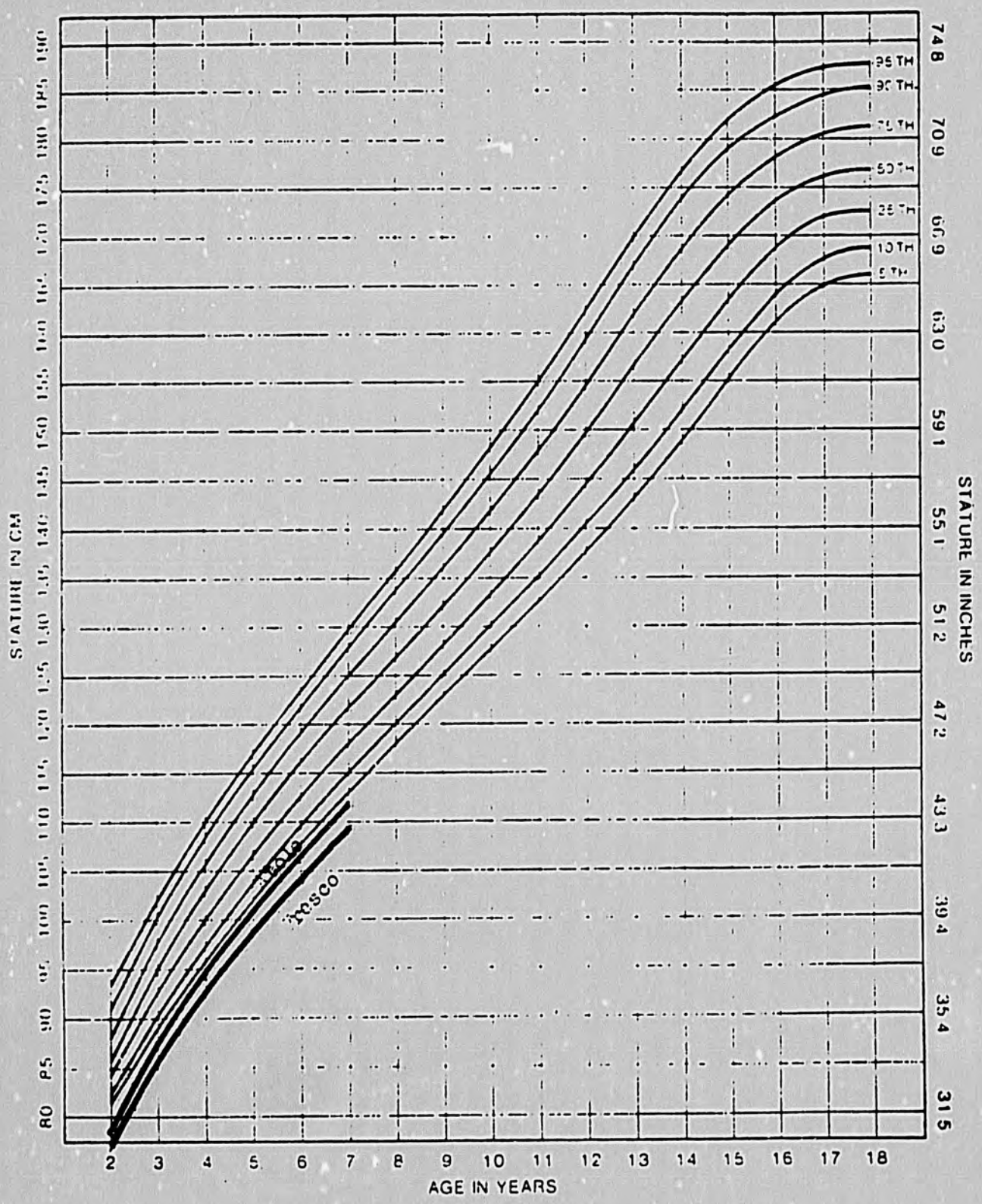


Figure A.11

Girls Weight by Age Percentiles: Ages 2 to 18 Years

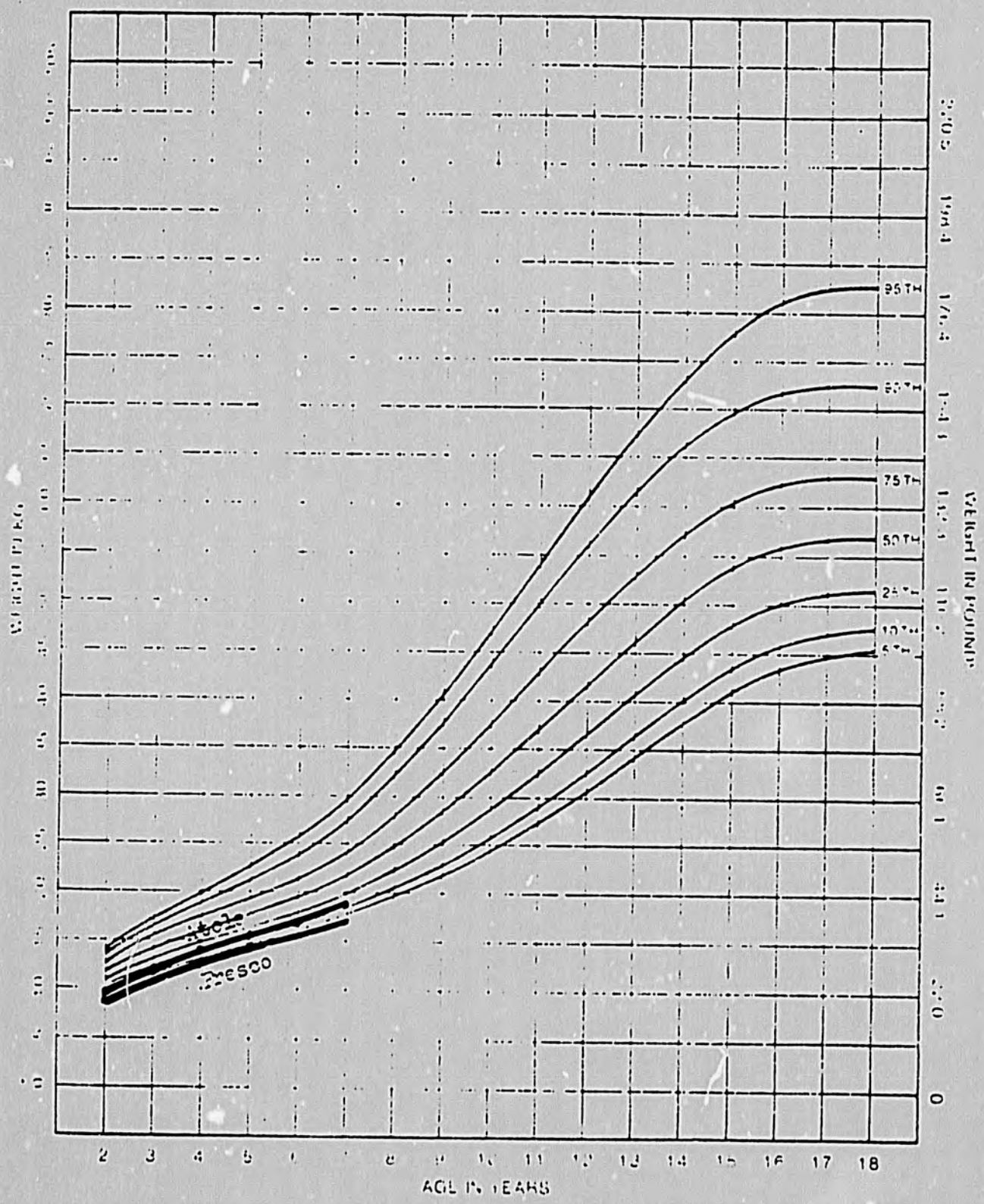
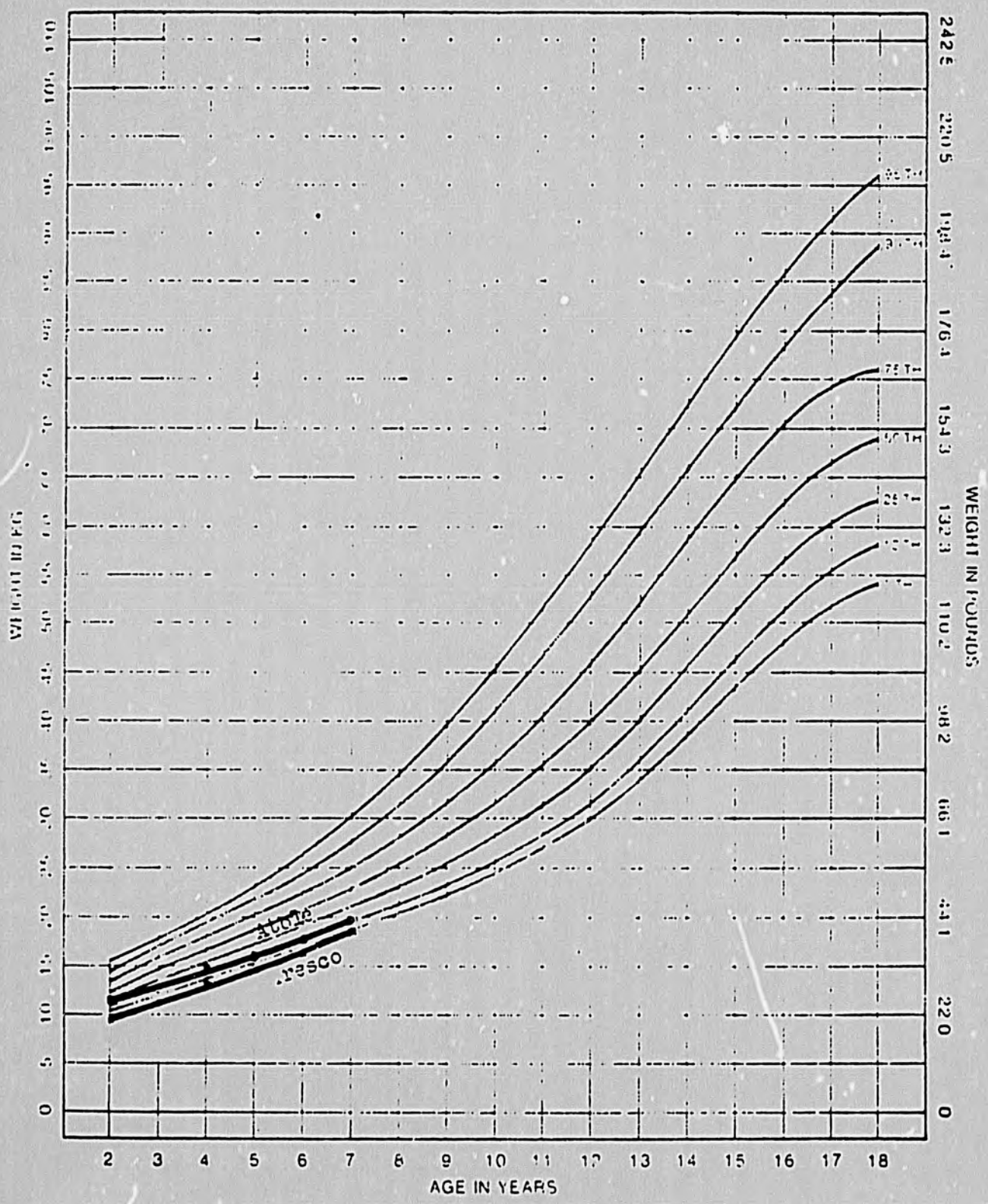
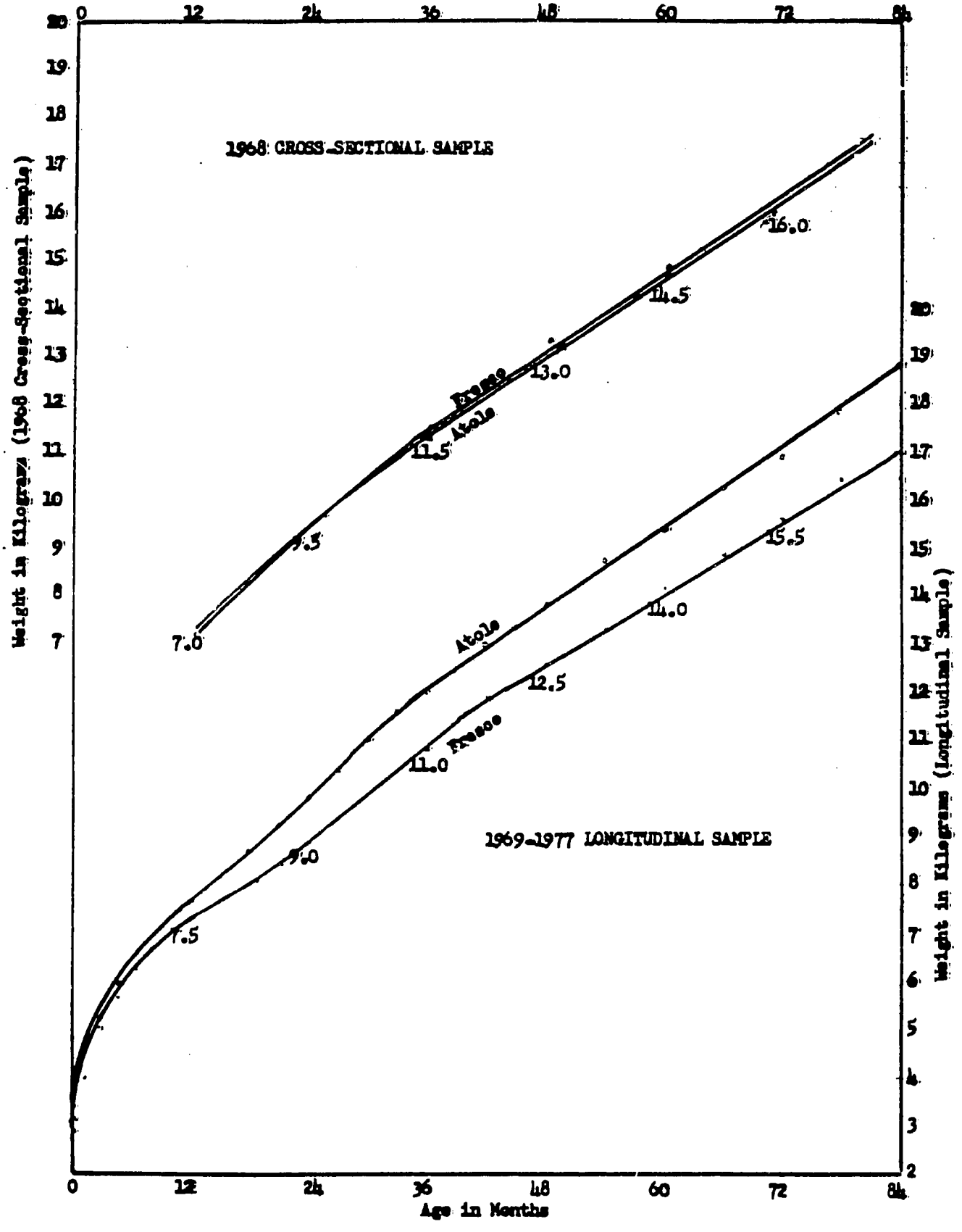


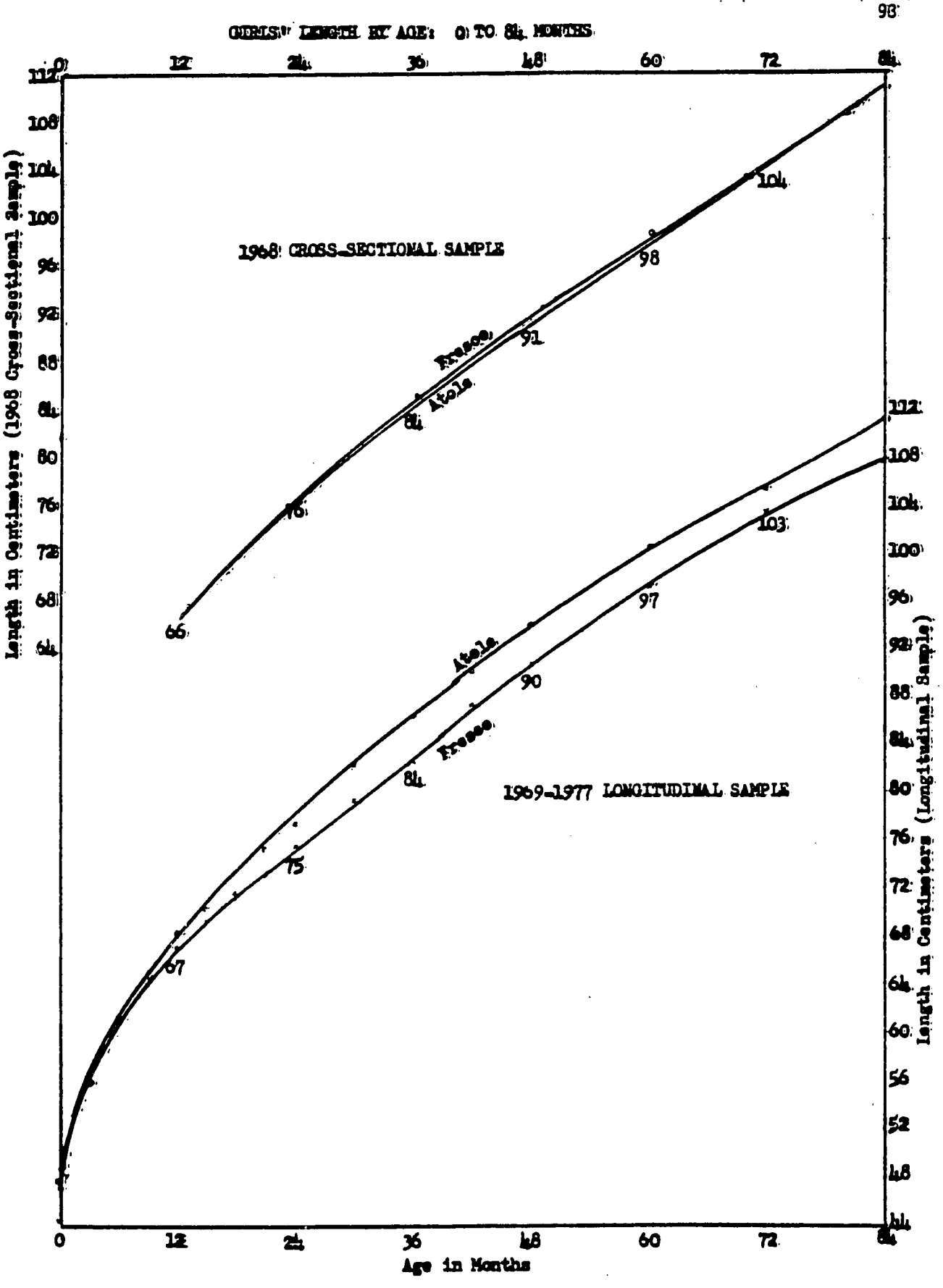
Figure A.12

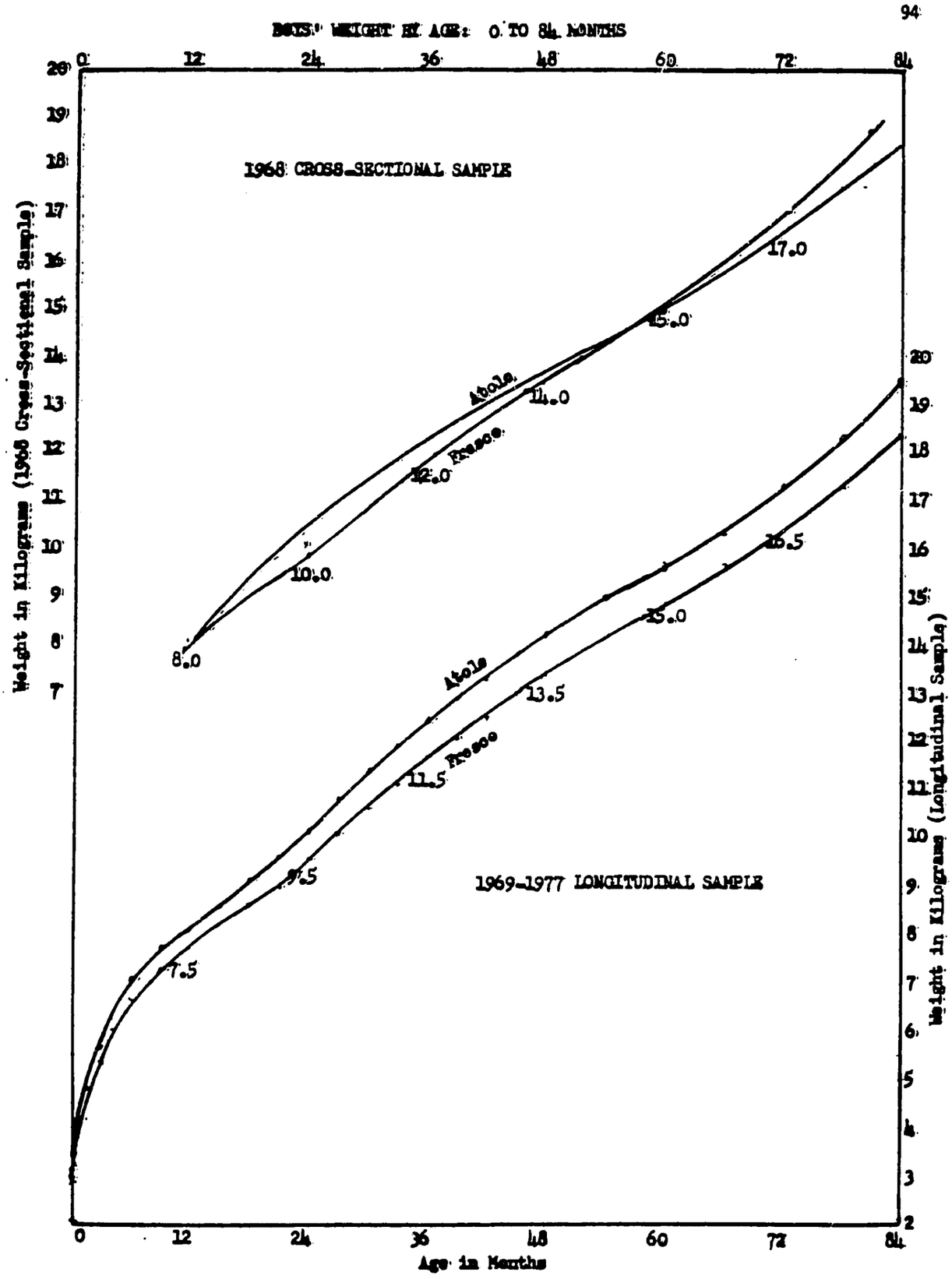
Boys Weight by Age Percentiles: Ages 2 to 18 Years

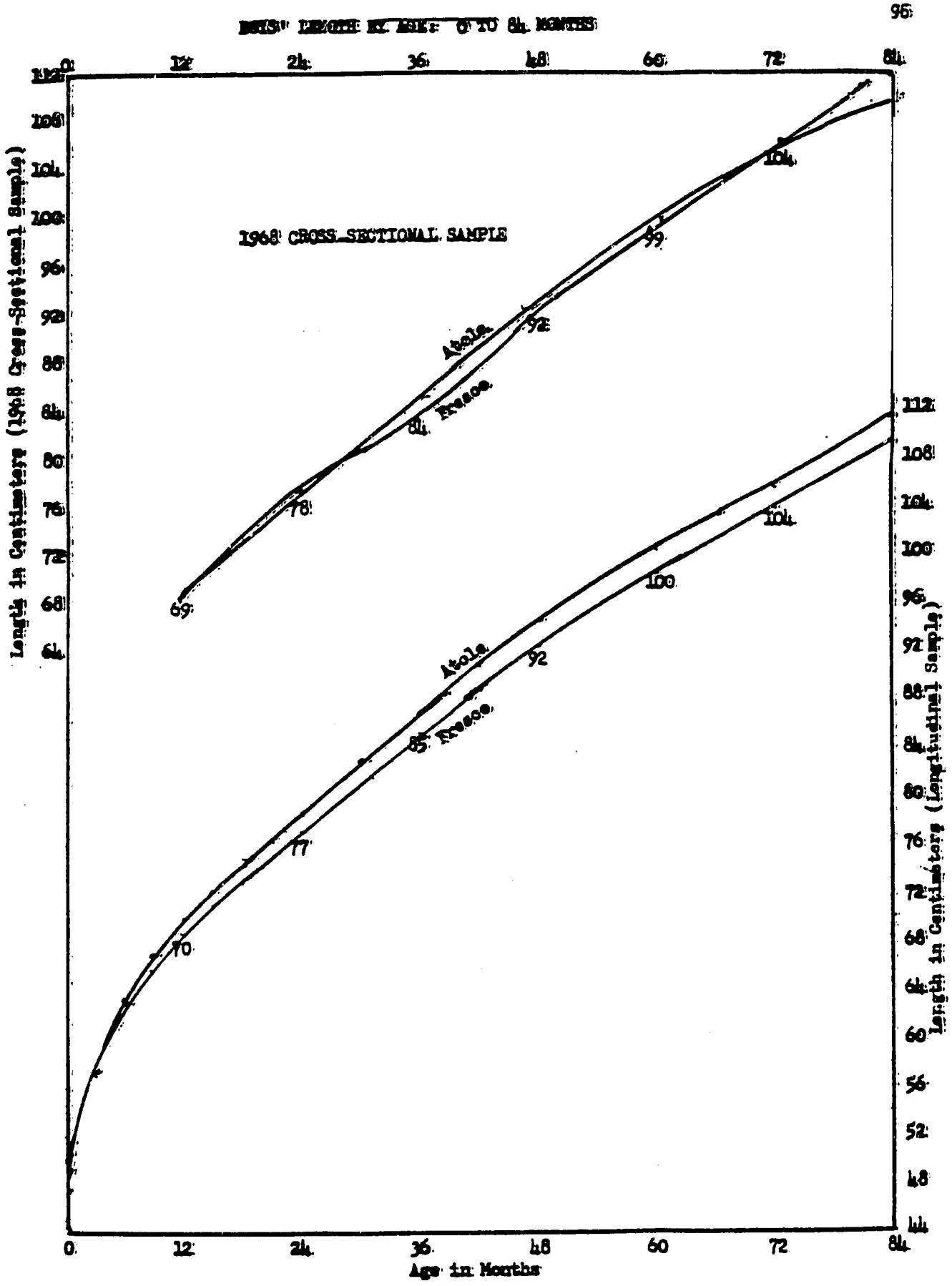


GIRLS: WEIGHT BY AGE: 0 TO 84 MONTHS




GIRLS' LENGTH BY AGE: 0 TO 84 MONTHS
1968 CROSS-SECTIONAL SAMPLE
1969-1977 LONGITUDINAL SAMPLE
Freese
Atole
Length in Centimeters (1968 Cross-Sectional Sample)
Length in Centimeters (Longitudinal Sample)
Age in Months
0
12
24
36
48
60
72
84
112
108
104
100
96
92
88
84
80
76
72
68
64
60
56
52
48
44
66
76
84
91
98
104
67
75
84
90
97
103


BOYS' WEIGHT BY AGE: 0 TO 84 MONTHS
1968 CROSS-SECTIONAL SAMPLE
1969-1977 LONGITUDINAL SAMPLE
Atole
Fresco
8.0
10.0
12.0
14.0
15.0
17.0
7.5
9.5
11.5
13.5
15.0
16.5
Weight in Kilograms (1968 Cross-Sectional Sample)
Weight in Kilograms (Longitudinal Sample)
Age in Months
0
12
24
36
48
60
72
84


BOYS' LENGTH BY AGE: 0 TO 84 MONTHS
1968 CROSS-SECTIONAL SAMPLE
Atola
Freeo
Length in Centimeters (1968 Cross-Sectional Sample)
Length in Centimeters (Longitudinal Sample)
Age in Months
0
12
24
36
48
60
72
84
69
78
84
92
99
104
70
77
85
92
100
104

II.A.5. Home Diet, Supplementation, and Physical Growth

A longitudinal model depicting hypothesized causal relations among four developmental sequences: home diet and supplementation at ages 30, 42, and 54, weight and height at ages 24, 36, 48, and 60, and 23 other exogenous variables is presented in Figures A.13 and A.14, and, equivalently, the structural equation is given in Equation A.1. (See Appendix A.11.a for a brief discussion of the estimation of path models of this sort and a list of references.)

In Figure A.13, the main structural relations between the endogenous variables -- the diet and growth sequences -- is portrayed. The vertical arrows represent the "stabilities" in each sequence, e.g., those who are taller at age 24 months can be expected to be taller at age 36 months, etc. The curved arrows represent the interrelationship between sequences. Prior weight is hypothesized to influence the amount of food an individual is wont to consume. In turn, dietary consumption -- both home diet and supplementation -- are hypothesized to affect growth in weight and in stature at the next time period.

The variables taken as exogenous, in addition to the age 24 month starting points for weight and stature, were introduced in places where a potential effect seemed theoretically plausible. More controversially exogenous variables were excluded from certain equations where there was a conceptual rationale for the exclusion, and prior exploratory ordinary least squares regressions seemed to show that they would make a negligible contribution to the prediction of the dependent variable. The exclusions -- i.e., the coefficients in the system of equations which were set at zero -- were also checked a posteriori by examining the partial derivatives of these coefficients to see whether relaxing the assumption -- changing the specification -- would result in a better fit.

There are a number of constraints on the estimation of this model aside from theoretical plausibility. (1) The model as specified here fully used the 140K capacity of the computer at Berkeley. The introduction of any additional unknowns would have to be offset by the elimination of variables, or unknowns, specified in the model. (The magnitude of this model also made convergence a slow and costly operation). (2) If possible reciprocal effects -- whether replacement or reinforcement -- between home diet and supplementation is to be estimable, some variable(s) having a substantial impact on home diet, say, must be excludable from the determination of supplementation (except through home diet), and vice versa. (3) With a developmental sequence such as is shown in Figure A.14 where there is a single fallible indicator of the "true" value of the variable, the terminal error terms -- in this case ε_1 and ε_3, are not identifiable. Some side-constraint must be imposed on their values in order to provide an estimate.

Since the measures of home diet have been assumed to be the most fallible measures in the INCAP data, making some plausible estimate of these error terms is of central importance both for this model and for use in subsequent models in which the present estimate could then be taken as a fixed value. The constraint which was imposed was that the <u>standardized</u> values of the error terms should be the same. Since the variance in home diets increases with increasing age, equality constraints on the unstandardized error variance would imply increasing accuracy as the child gets older. While this is possible, it is purely speculative. The opposite presumption could also be rationalized. It seemed more parsimonious and conservative to assume equality in the proportion of variance which should be ascribed to error regardless of age.

An additional reason for analyzing the standardized covariance matrix -- i.e., correlation matrix -- was the desire to constrain the effects of father's and mother's heights upon the growth in stature of the child to be proportionately the same, despite the greater variance in father's height.

The estimation of the model, in standardized form, led to estimates of error variances of .23 for home diet, .03 for measurements of weight, and .00 for the measurements of supplementation and of stature. The overall fit of the model was "good." (In general, a value of x^2 less than the number of degress of freedom, together with a high probability -- close to 1.00 -- that given the specifications a random covariance matrix would not provide as good a fit, was taken as indicative of a "good" fit. In this case, x^2/df was .73 and p=1.00.)

Since this is the most comprehensive model including the determination of home diets to be estimated, these error estimates of 23% of the variance in home diet and 3% of the variance in weights are used in the estimation of subsequent models in this section (A). Also, it is later assumed that the errors in measurement of supplementation and of height are negligible.

Whether the errors of measurement of home diets are in fact proportional at different age levels (and in the different villages), as was assumed, is, of course, open to question and further inquiry. If taken as an indicator of total consumption other than supplementation, the fact that modally mothers continued to lactate until their children were about 24 months old suggests that there is a systematic underestimation and that there is another unknown but possibly large component of error at age 18 months.

This standardized solution was converted to unstandardized form, and, with the error terms and coefficients of the effects of parents' heights upon children's heights fixed, iterated to convergence.

Table A.16 ("Diet and Growth, Age 24-60 Months") shows the unstandardized coefficients of the structural equations. In this and subsequent tables of coefficients, an asterisk denotes the absolute value of a coefficient is more than twice its standard error; a non-zero value indicates a coefficient less than its standard error; and, a zero indicates a coefficient less than its standard error. (Blanks occur where variables have been excluded -- presumed to be zero -- in a particular equation.)

The strongest determinants of each of the endogenous variables is its lagged value showing a strong stability over time: those who eat more, or are larger, at an early age are apt to eat more, or be larger, at later ages. The overall determination of each variable, shown in the row marked R^2 near the bottom of the table, suggests that measurement error was not too grossly underestimated.

Other than prior home diet, the strongest predictors of the level of home diet are sex (the boys eating more than the girls) and the number of elder siblings who were deceased. This variable must be acting as a proxy for some important aspects of child nurturance in the home which is not being tapped by the literacy and SES scales. In the earlier section on village contrasts, it was pointed out that the frequency of infant mortality prior to the project was higher in the Atole villages. However, a replication of this model in the Atole villages alone shows this variable to be a strong negative predictor of home diet within this treatment group (i.e., it is not a proxy for village characteristics, or for replacement). There is also a trend, over and above sex differences, for the larger individuals to consume more.

The strongest predictor of supplementation is treatment group. As was evident in the tabulation of children's diets, those in the Atole

villages consumed far more of the supplementation offered them than in the Fresco villages. Within villages, the mothers' attendance at the center is the strongest predictor of the child's consumption. Once again, as in the case of home diets, boys tend to consume more of the supplementation than do the girls. Children from the materially better homes are less likely to use the supplementation than those from the poorer homes.

There is no clear evidence in this, or the subsequent tables to be presented below, of a replacement of home diet by supplementation within villages. We have already noted in prior sections, A.2 and A.3, that in the aggregate, children's home diets in the Atole villages are not much larger than those in the Fresco villages, and, presumably, similar to the total energy value of the diets they might have consumed in the absence of any treatment. In the presence of the treatment, however, while their home diets are diminished, on the average, at an individual level within villages, those who eat more at home are likely to consume more supplementation. The most significant coefficient is the figure .08 representing a positive relationship between home diet and supplementation at age 18 months in the following table (Table A.17). (Table A.17 is a replication of this entire model shifted down one year -- showing growth from ages 12 to 48 months rather than 24 to 60 months.)

The most interesting finding in Tables A.16 and A.17 is the significant effect which supplementation, but not home diet, has upon growth in both height and weight. In Table A.17, we can see that the supplementation has its largest effect on growth in stature at the earliest ages (12-24 months), and, an additional incremental effect the following year (24-36 months) but is not thereafter significant.

This is congruent with the growth curves by sex and treatment earlier examined which showed the disparity in growth between the Atole and Fresco

villages to be well established by age 36 months with average growth curves which are nearly parallel thereafter.

Both of these sources of dietary consumption are measured as average daily diets in kcal units. If it were simply the calorie value of the supplementation which was having an effect upon growth, the effect of home diet (which provides by far the largest portion of the diet) should be approximately identical. The possible attenuating effect of measurement error in the home diet estimates has been allowed for in the model. As the tallies of children's diets showed, while the supplementation of the children in the Atole village did lead to an aggregate advantage in total caloric consumption during their early years through about age four, the largest and most significant effect of the supplementation was to increase the proportion of the diet from protein sources.

We shall approach this is two ways: first, by re-running this same large model in the Atole and Fresco villages separately. Here we would expect the effect of supplementation to be strong in the Atole villages, but similar to or less than the effect of home diet in the Fresco villages. Then, in the next section, rather than contrast home diet with supplementation, we shall contrast the effect of total protein in the diet with total calories from non-protein sources.

Table A.18 shows the estimated structural coefficients in the Atole villages alone for 127 cases from age 12 to 48 months. As before, and as expected, supplementation shows a strong positive effect on growth -- especially at the early ages. It is interesting to note in this table that the sex difference in home diets in the Atole villages is considerably attenuated and is no longer statistically significant. This is consistent with the earlier suggestion that the presence of a "free good" in the form of supplementation reduces the competition for food at home, thus lessening the disadvantage of the girls.

Table A.19 is, unfortunately, an intermediate solution of the same model for the Fresco village children. The iterative process of minimization of the function was not complete when data process was discontinued (September 19th). However, the direction in which coefficients are being modified to improve the fit is clear. The effect of supplementation in the Fresco villages upon growth is no longer significant and is approximately of the same magnitude as the effect of home diet.

While the evidence in this section is weakened by the incompleteness of the solution shown in Table A.19, the corroborative evidence from the model developed in the next section will be more compelling.

List A.1
Variables in Diet and Growth Model
Age 24-60

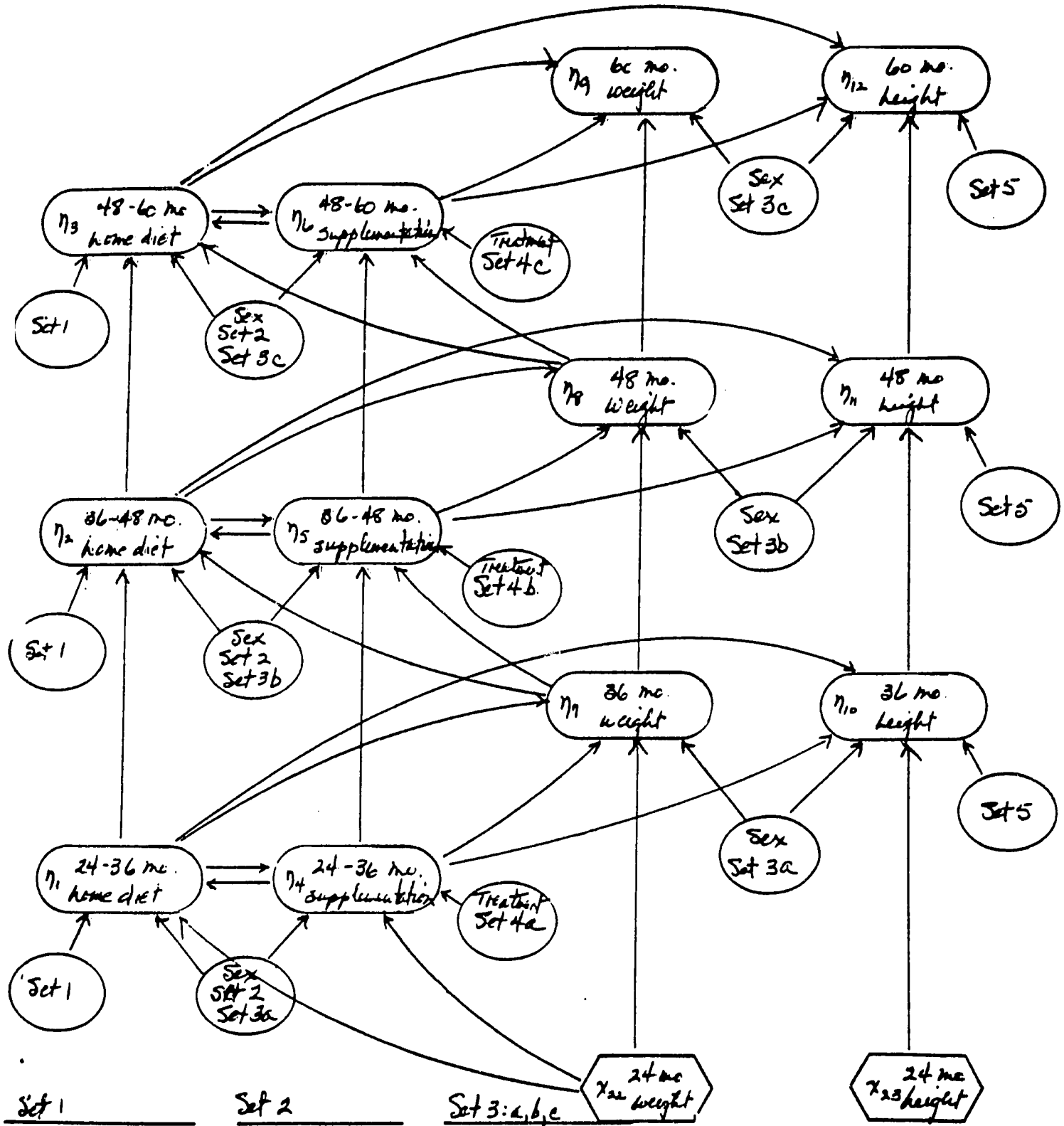
Symbol: Observed Variables	Symbol: Constructs	Acronym	Age	Variable Description
y_1	η_1	BCHCAL2	30	Estimated total home diet in kcal units
y_2	η_2	ACHCAL3	42	Estimated total home diet in kcal units
y_3	η_3	RCHHDC11	54	Estimated total home diet in kcal units
y_4	η_4	CHSP18	30	Average daily total supplementation in kcal units
y_5	η_5	CHSP19	42	Average daily total supplementation in kcal units
y_6	η_6	CHSP13	54	Average daily total supplementation in kcal units
y_7	η_7	WT12	36	Weight in kilograms
y_8	η_8	WT14	48	Weight in kilograms
y_9	η_9	WT15	60	Weight in kilograms
y_{10}	η_{10}	HT12	36	Height in centimeters
y_{11}	η_{11}	HT14	48	Height in centimeters
y_{12}	η_{12}	HT15	60	Height in centimeters
x_1	ξ_1	SEX		Sex
x_2	ξ_2	ZMARD		Mother's literacy and schooling
x_3	ξ_3	ZPARD		Father's literacy and schooling
x_4	ξ_4	MODVOC		Mother's modernity and vocabulary
x_5	ξ_5	CONSUMP		Housing, possessions, and salary
x_6	ξ_6	PASTAT		Father's occupation, ownership of land, animals
x_7	ξ_7	ZMAOCC		Mother's occupation

List A.1 continued

Symbol				
Observed Variables	Constructs	Acronym	Age	Variable Description
x_8	ξ_8	YOUNG3	36	Number of younger siblings at age
x_9	ξ_9	FMSZO	0	Number of household members at age
x_{10}	ξ_{10}	TREAT		Treatment: Atole vs. Fresco villages
x_{11}	ξ_{11}	MAHT		Mother's height in centimeters
x_{12}	ξ_{12}	PAHT		Father's height in centimeters
x_{13}	ξ_{13}	MAATT3	24-36	Mother's attendances at supplementation center: proportion
x_{14}	ξ_{14}	MAATT4	36-48	Mother's attendances at supplementation center: proportion
x_{15}	ξ_{15}	MAATT5	48-60	Mother's attendances at supplementation center: proportion
x_{16}	ξ_{16}	CHDIAR20	24-36	Child's incidence of diarrhea: proportion
x_{17}	ξ_{17}	CHDIAR21	36-48	Child's incidence of diarrhea: proportion
x_{18}	ξ_{18}	CHDIAR14	48-60	Child's incidence of diarrhea: proportion
x_{19}	ξ_{19}	MORBOTH2	24-36	Child's incidence of ailments other than diarrhea
x_{20}	ξ_{20}	MORBOTH3	36-48	Child's incidence of ailments other than diarrhea
x_{21}	ξ_{21}	MORBOTH4	48-60	Child's incidence of ailments other than diarrhea
x_{22}	ξ_{22}	WT10	24	Weight in kilograms
x_{23}	ξ_{23}	HT10	24	Height in centimeters
x_{24}	ξ_{24}	DSTRUC3	36	Family structure: nuclear vs. all other
x_{25}	ξ_{25}	DIED3	36	Number of elder siblings deceased at age

Figure A.13

Diet and Growth Model -- Outline of Structural Relations



Set 1

Father Status
Mother Occupation
Younger Siblings
Family Size at birth

Set 2

Mother Literacy
Father Literacy
Modernity
Affluence
Family Structure
Deceased Elder Sibs

Set 3: a, b, c

a. 24-36 mo diarrhea
24-36 mo. morbidity
b. 36-48 mo. diarrhea
36-48 mo. morbidity
c. 48-60 mo diarrhea
48-60 mo morbidity

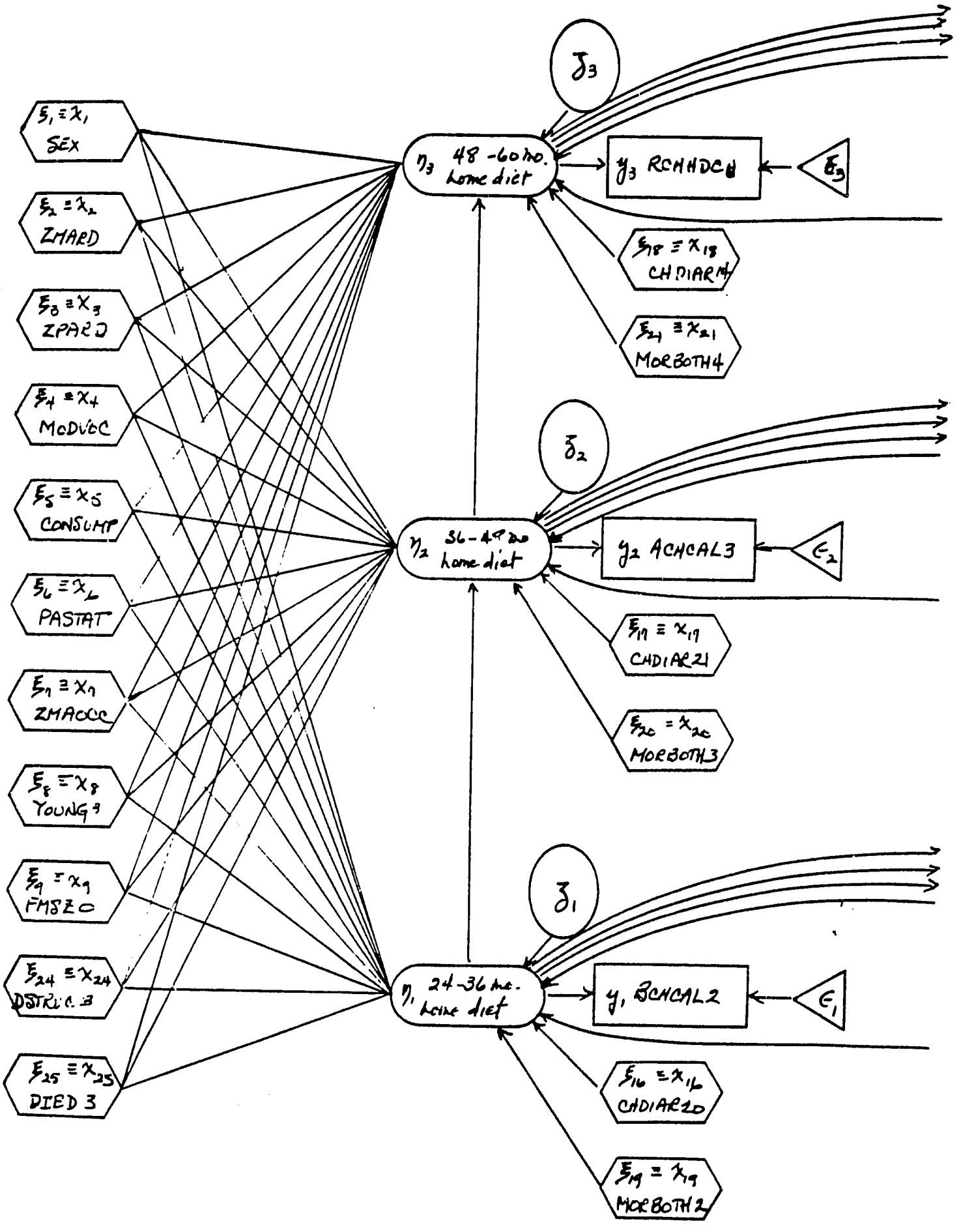
Set 4: a, b, c

a 24-36 mo. Mothers Attend.
b 36-48 mo. Mothers Attend
c 48-60 mo Mothers Attend

Set 5

Mothers height
Fathers height

Figure A.14
Diet and Growth Model -- Detail of a Segment



Equation A.1

Structural Equation for Diet and Growth Model

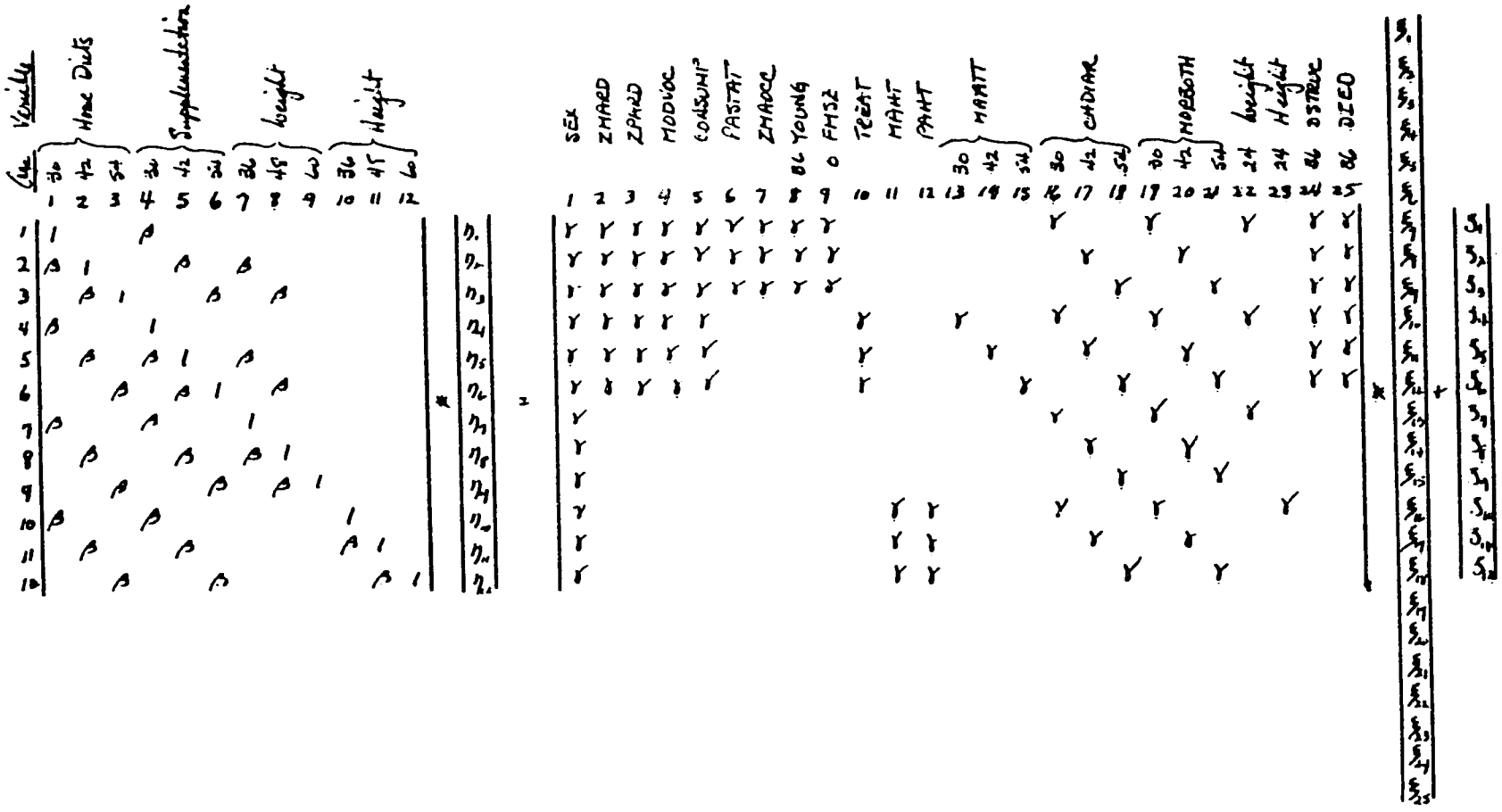


Table A.16

DIET AND GROWTH - AGES 24-60 MONTHS
Unstandardized Structural Coefficients
All Villages combined
χ^2=414.19; df=567; χ^2/df=.73; P=1.00
N=188

		HOME DIET			SUPPLEMENTATION			WEIGHT			HEIGHT		
	AGES	30	42	54	30	42	54	36	48	60	36	48	60
y_i	Mean	637	985	1139	98	103	113	11.8	13.6	15.1	85.7	93.1	99.5
	Standard Deviation	240.4	296.7	369.3	104.4	100.2	98.0	1.3	1.3	1.5	3.8	4.0	4.1
$-\beta_{ij}$	Lagged value		1.05*	.71*		.88*	.79*		.96*	1.05*		.94*	.99*
	Home diet				0	0	.02	0	0	0	0	.0004	.0005
	Supplementation	0	.27	0				0	0	.0012*	.0032*	.0006	0
	Weight		6.73*	0		0	0						
γ_{ik}	SEX	68.39*	0	66.17	-13.87	12.74*	-15.31	.16	0	0	0	0	-.63
	ZHARD	29.73	0	25.77	-11.17	0	0						
	ZPARD	0	0	0	8.52	0	0						
	MODYOC	39.11	0	0	9.88	5.13	0						
	CONSUMP	0	-32.15	0	0	0	-9.81*						
	PASFAT	-23.77	-17.86	0									
	ZHAOCC	49.29	0	27.50									
	YOUNG3	-40.64	0	0									
	FHSZO	11.19	16.17	14.39									
	TREAT				108.50*	0	0				.05*	.03*	0
	MAHT										.05*	.03*	0
	PAHT												
	HAATT				154.84*	36.46*	39.92*						
	DIARRHEA	0	0	663.48	0	0	0	0	0	-1.66	-2.10	-4.10	0
	MORBIDITY	0	0	0	0	-16.74	0	0	0	0	0	0	0
	HITO	0			8.31			.96*					
	HITO										.87*		
	DSTRUC3	-108.76	0	0	30.66	0	-29.69						
	DIED3	-40.52*	0	0	3.95	2.88	0						
	$R^2 = 1 - \psi_{ii}$	.29	.80	.51	.57	.84	.75	.75	.88	.89	.91	.89	.90
	$F_{k, N-k-1}$	4.78	44.06	10.90	17.78	67.06	36.26	88.51	104.22	105.93	72.41	79.13	79.79
	p <	.001	.001	.001	.001	.001	.001	.001	.001	.001	.001	.001	.001

TABLE A.17

DIET AND GROWTH - AGES 12-48 MONTHS
Unstandardized Structural Coefficients
All Villages Combined
χ^2=610.03; df=567; χ^2/df=1.08; P=.10
N=243

		HOME DIET			SUPPLEMENTATION			WEIGHT			HEIGHT		
	AGES	18	30	42	18	30	42	24	36	48	24	36	48
μ_i	Mean	611	865	1009	79	107	114	9.7	11.9	13.6	77.4	85.8	93.2
	Standard Deviation	227.5	266.9	299.5	93.6	108.5	113.7	1.1	1.3	1.4	3.7	3.9	4.0
β_{ij}	Lagged values												
	Home Diet		.69*	1.01*	.08*	.69*	.94*	0	1.01*	1.01*	0	.90*	.97*
	Supplementation	-.32	0	0		-.02	-.02*	.0028*	.0002	0	.0072*	.0006	.0004
	Weight		0	0		0	0		0	0		.0031*	0
γ_{ik}	SEX	17.65	76.44*	0	0	0	13.63*	0	0	0	0	.28	-.20
	ZHARD	31.41	0	0	0	0	-4.45						
	ZPARD	29.19	0	19.15	5.56	0	0						
	MODVOC	40.42*	21.88	0	0	0	0						
	CONSUMP	0	0	0	-6.60	-8.90*	0						
	PASTAT	0	-30.28	-32.30									
	ZMAOCC	34.89*	37.73*	0									
	YOUNG3	0	0	0									
	FMSZO	0	0	15.39									
	MAHT										.05*	.03*	0
	PAHT										.04*	.03*	0
	MAATT				102.43*	92.85*	38.98*						
	DIARRHEA	0	0	-454.02	0	0	0	0	-.65	0	-1.86	-2.61	-3.17
	MORBIDITY	112.62	-136.53	73.11	0	-17.83	-19.97	0	0	-.14	0	0	0
	WT6	0			0			.80*					
	HT6										.95*		
	DSTRUC3	0	0	-101.44	25.55	0	0						
	DIED3	-14.39	-16.57	0	0	0	0						
	TREAT				128.11*	33.18*	-17.31*						
	$R^2 = 1 - \psi^*$	.18	.61	.75	.55	.75	.79	.66	.80	.88	.71	.82	.89
	$F_{k, N-k-1}$	3.41	14.12	43.50	21.11	48.06	61.27	71.93	149.56	275.14	71.51	130.76	230.98
	p <	.001	.001	.001	.001	.001	.001	.001	.001	.001	.001	.001	.001

TABLE A.18

DIET AND GROWTH FROM 12-48 MONTHS
Unstandardized Coefficients
Atole Villages
χ^2 = 473.71; df=533; χ^2/df=.89; P=.97
N=177

		HOME DIET			SUPPLEMENTATION			WEIGHT			HEIGHT		
	AGES	18	30	42	18	30	42	24	36	48	24	36	48
$-\beta_{ij}$	Lagged values												
	Home diet		.69*	1.06*		.66*	.96*		1.10*	1.06*		.90*	1.00*
	Supplementation	.00*	.32	0	-.15	0	-.06	0	0	0	0	0	0
	Weight		0	0		0	0	.0023*	.0006	0	.0068*	.0028*	0
γ_{ib}	SEX	0	45.77	0	0	0	27.30*	0	0	0	-.72*	-.18	0
	ZMARD	38.84*	0	-24.83	0	0	0						
	ZPARD	31.03	0	0	19.23	0	0						
	MODVOC	0	20.79	0	23.05*	0	0						
	CONSUMP	49.86*	0	30.63	0	-19.31*	0						
	PASTAT	-29.91	-19.74	-44.66									
	ZMAOCC	-16.90	45.54*	-22.31									
	YOUNG3	0	-67.79	0									
	FMSZO	0	0	20.11									
	MAHT										.05*	.03*	0
	PAHT										.05*	.03*	0
	MAATT												
	DIARRHEA	397.86*	0	0	127.12	0	0	0	0	0	0	-3.26	-2.35
	MORBIDITY	55.73	0	122.40	28.93	0	0	0	0	-.19	.89	0	0
	Height (12 mo.)	0			-11.96			.78*					
	Height (12 mo.)										.93*		
	DSTRUC3	0	0	0	0	0	0						
	DIED3	-34.42*	0	0	0	0	0						
	$R^2 = (1 - \psi)$	.31	.65	.72	.13	.60	.75	.64	.79	.88	.76	.81	.89
	$F_{k, N-k-1}$	3.59	12.72	17.41	1.31	12.11	24.51	35.14	93.41	140.58	46.48	65.75	114.52
	$p <$	.001	.001	.001	.30	.001	.001	.001	.001	.001	.001	.001	.001
y:	Mean	595	795	963	108	192	168	10.1	12.3	14.0	89.4	96.9	94.3
	Standard Deviation	300.1	222.6	377.0	16.4	113.4	158.8	1.0	1.2	1.4	3.3	8.5	8.7

TABLE A.19

DIET AND GROWTH FROM 12-48 MONTHS
Unstandardized Coefficients
Fresco Villages
X^2=1148.88; df=532; X^2/df=2.16; p=.00
N=116

		HOME DIET			SUPPLEMENTATION			WEIGHT			HEIGHT		
	AGES	18	30	42	18	30	42	24	36	48	24	36	48
μ_i	Mean	650	919	1060	14	35	56	9.3	11.4	13.2	76.4	84.6	92.0
	Standard Deviation	249.1	276.6	315.7	12.7	27.3	36.4	1.2	1.3	1.4	3.7	3.9	3.9
β_{ij}	Lagged values		.69	1.01		.70	.94		1.01	1.01		.90	.97
	Home Diet				.06	.02	.00	.0008	.0006	.0000	.0008	.0013	.0006
	Supplementation	-.30	.11	-.07				.0011	.0006	.0001	.0011	.0008	.0000
	Weight		.86	11.48		4.23	-2.07						
γ_{ik}	SEX	.03	75.38	-.03	0		13.50	0	0	0		.28	-.20
	ZWARD	31.26					-4.44						
	ZPARD	29.07		19.15	5.56								
	MODVOC	40.50	21.79										
	CONSUMP				-6.60	-8.87							
	PASTAT		-30.24	-32.28									
	ZMAOCC	35.00	37.71										
	YOUNG3												
	TMSZO			15.40									
	MWIT										.05	.03	0
	PWIT										.05	.03	0
	MWTT					92.66	38.98						
	DIARRHEA			-455.68						-.0015	1.86	-2.61	-3.17
	MORBIDITY	112.10	-137.00	73.00		-17.79	-19.95			-.14			
	HT6							.78					
	HT6										.95		
	DSTRUC3			-102.75	25.41								
	DIED3	-14.43	-16.56										
	$R^2 = 1 - \Psi^*$	.14	.60	.76	.14	.54	.70	.63	.79	.87	.69	.80	.88
	$F_{k, N-k-1}$	1.26	7.11	19.08	1.30	8.51	16.58	30.50	66.51	122.98	29.74	53.94	98.55
	p <	.50	.001	.001	.80	.001	.001	.001	.001	.001	.001	.001	.001

*Significance of individual coefficients not calculated.

II.A.6. Calories, Protein, and Physical Growth

Since in the large model presented in Section A.5 family characteristics were presumed to affect growth through their effects upon a child's diet, but not directly, we here simplify the model, taking weight and stature to be the endogenous variables and the various components of diet to be exogenous. The other factors directly affecting growth in the prior model -- sex, diarrhea, other morbidity, and parents' heights -- are again included.

In the present model, instead of "home diet" and "supplementation," protein (measured in kcal units) and calories (from non-protein sources) are used by adding a set of identities to the simultaneous equation system. (See Equation A.2 in Appendix A.11.b.)

E.g.:

η_7 ("calories" at age 30) = ξ_1 (home diet calories) + X_2 (supplementation calories). These identities are shown in the variable description list for η_7 through η_{12}, and, in matrix form, the equations appear in the lower half -- the last six rows -- of the structural equation. The model is shown here as a path diagram in Figure A.15.

Except for the reaggregation of the dietary data, the model is similar to the right hand side of Figure A.13 in Section A.5. The errors in measurement of the home diet data (whether calorie or protein) are set at 23% of their variance from the earlier estimate; and the errors of measurement of weight are set at 3% of their variance.

This model also was estimated at two age ranges: 24 to 60 months and 12 to 48 months. The estimated structural coefficients are shown in Tables A.24 and A.25.

These two tables are analogous to the right hand side of Tables A.16 and A.17 in Section A.5, above, the major change being the reaggregation of the dietary data.

Another difference between the model portrayed in Figure A.15 and the right hand size of Figure A.13 is the omission of lagged values of weight affecting dietary intake. These were omitted so that calories and protein could be estimated additively from their home diet and supplementation sources, with the home diet component measured with error. E.g.,

$$\eta_7 = \xi_1 + \xi_2 \quad \text{and} \quad x_1 = \xi_1 + \delta_1 \ , \quad x_2 = \xi_2 .$$

An alternative would be to use the entire model specified in Figures A.13 and A.14 and Equation A.1, replacing home diet and supplementation with calories and protein, respectively.

Once again we find the lagged values of the measures of weight and height to have the largest impact, as is to be expected. These lagged values from these various sources are tallied in Table A.20 below.

TABLE A.20

Coefficients of Lagged Values

Variable	Age	Section 5		Section 6	
		Table A.16	Table A.17	Table A.24	Table A.25
Weight	24		.80		.82
	36	.96	1.01	.98	1.01
	48	.96	1.01	.96	1.01
	60	1.05		1.04	
Height	24		.95		.97
	36	.87	.90	.89	.90
	48	.94	.97	.95	.97
	60	.99		.98	

The first column in each case consists of the smaller sub-sample who had reached age 60 months (and been measured at that age). The differences between estimates at a given age -- comparing figures in each row -- are small. For example, from the four estimates of the regression of height at age 48 months on height at age 36 months, we would estimate one centimeter at age 36 to be associated with between .94 and .97 centimeters at age 48.

The coefficients from Tables A.24 and A.25 of this section, representing the effects of calories and protein on growth in weight and stature are tallied together in Table A.21 below. Once again, the figures from Table A.24, covering the ages 12 to 48 months, are based upon a somewhat larger sample size (243 as contrasted with 189 for Table A.25).

The largest effect shown in these models is the effect of protein on growth in stature between ages 12 and 24. The coefficient implies, other things being constant, each increment of a kcal due to protein would result, in the average, of .0213 centimeters. From the tabulation of children's diets earlier, we see that at age 18 months, the children in the Atole villages enjoyed a 30 kcal advantage in the protein content of their diet. If other factors were equal, this protein advantage would lead to 30 x .0213 = .639 cm. advantage in growth in stature between ages 12 and 24.

TABLE A.21

Unstandardized Structural Coefficients
From Tables A.24 and A.25 Showing the Effect of
Calories and Protein on Weight and Height by Age

Variable	Age	Calories		Protein	
		Table A.24	Table A.25	Table A.24	Table A.25
Weight	24	---a	-.0001	---a	+.0099*
	36	0	+.0004	0	0
	48	-.0002	0	0	0
	60	-.0004*	---a	+.0032*	---a
Height	24	---a	-.0020*	---a	+.0213*
	36	0	0	+.0075	+.0068
	48	0	+.0006	0	0
	60	0	---a	+.0035*	---a

a not estimated in model.

TABLE A.22

Growth in Stature From Age 12 to 24 Months
By Sex and Treatment

Sex	Atole	Fresco	Difference
Boys	8.80	8.45	.35
Girls	9.30	8.39	.91
Average	9.05	8.42	.63

In Table A.22, growth in stature is tallied by sex and treatment. While there is an enormous interaction with sex -- the girls in the Atole villages enjoying the greatest advantage, as has been discussed before -- the average difference between the Atole and Fresco villages in growth is .63 cm. This is almost exactly the advantage in growth which would be estimated from the model.

The negative coefficient attached to the effect of increased calories is more difficult to construe. There are several possibilities which require further analyses to sort out. If the error of measurement of home diets has been underestimated, the negative effect of calories is overestimated. In Table A.28, where home diet measurement error is set at 40% (rather than 23%), the negative effect of calories disappears while protein continues to have a significant positive effect. Secondly, the contribution of the mother's nursing has been ignored because it was not measurable. Increased calorie consumption during the second year may be associated with early weaning. Third, there may actually be an adverse effect of increased calorie consumption in the absence of proportionate protein consumption. That is, growth may be a function of the proportion of protein in the diet as well as the absolute amounts of protein and calories separately. Finally, diet should have been treated as a jointly endogenous variable as it was in Section A.5, above. Heavier children require more calories in order to maintain their body weight. Including an effect of weight on calorie consumption (in the model shown in Figure A.15) could attenuate the negative estimate of calories on growth.

The coefficient with respect to the effect of protein on gain in weight between ages 12 and 24 similarly implies a difference, on the average, between the Atole and Fresco villages of 30 x .0099 = .297 kg. In Table A.23, showing the gains in weight between ages 12 and 24 by sex and treatment, we once again find that the estimated effect of the protein almost accounts for the Atole advantage.

TABLE A.23

Growth in Weight From Age 12 to 24 Months
By Sex and Treatment

Sex	Atole	Fresco	Difference
Boys	2.08	1.90	.18
Girls	2.13	1.68	.45
Average	2.105	1.79	.315

This significant effect of protein on growth declines in subsequent years. While it reemerges as statistically significant between the ages of 48 and 60 months, it is of smaller magnitude. This is congruent with the growth curves showing a large treatment difference during the first two years and a continuing, but not widely increasing, advantage during later years.

Another interesting and important finding in the solutions presented in Tables A.24 and A.25 of this section is the consistent negative effect of the incidence of diarrhea upon growth. Although these coefficients do not reach statistical significance (i.e., twice their standard errors), the effect is substantial and seems to continue through later ages.

To explore this further, as far as the data would allow, a series of first differences in stature (i.e., growth during a one year interval) were calculated and the model re-estimated taking one year's growth at a time. This permitted the inclusion of more cases in each estimate since anthropometric measurements were only required at two consecutive anniversaries instead of four as in Tables A.24 and A.25. The path diagram and equations for these cross-sectional models are shown in Figure A.16 and Equation A.4. The structural coefficients are shown in Table A.30.

This sequence of six growth periods shows a consistently significant negative effect of diarrhea. Only during the children's sixth year does the effect disappear (become less than its standard error). While the range of incidence of diarrhea is from 0% to 100% of the days of the year, the average is about 10%. The coefficients of -4 shown for the 4th, 5th, and 7th years of life suggest that a reduction in the average incidence from 10% to 0% would increase average growth by 4 x .10 = .4 cm. per year.

Whereas the effect of the dietary intervention was highly significant during the first two or three years, in later years it seems that it is less the consumption but the ability to utilize nutrients efficiently that seems to make a substantial difference.

Two additional estimates of the model portrayed in Figure A.15 for ages 12 to 48 months were estimated fixing the estimated error for home diet components of calories and protein at 10% and at 40% respectively. The effect of reducing the estimated error (from 23% to 10%) is to magnify the estimated effect of protein on growth. The effect of increasing the estimated error (from 23% to 40%) is to diminish the estimated effect of protein on growth. Even with this exaggerated estimate of error, however, the coefficient is substantial and significant. These two alternative estimates are shown in Tables A.27 and A.28, respectively.

Figure A.15

Calories, Protein, and Growth Model

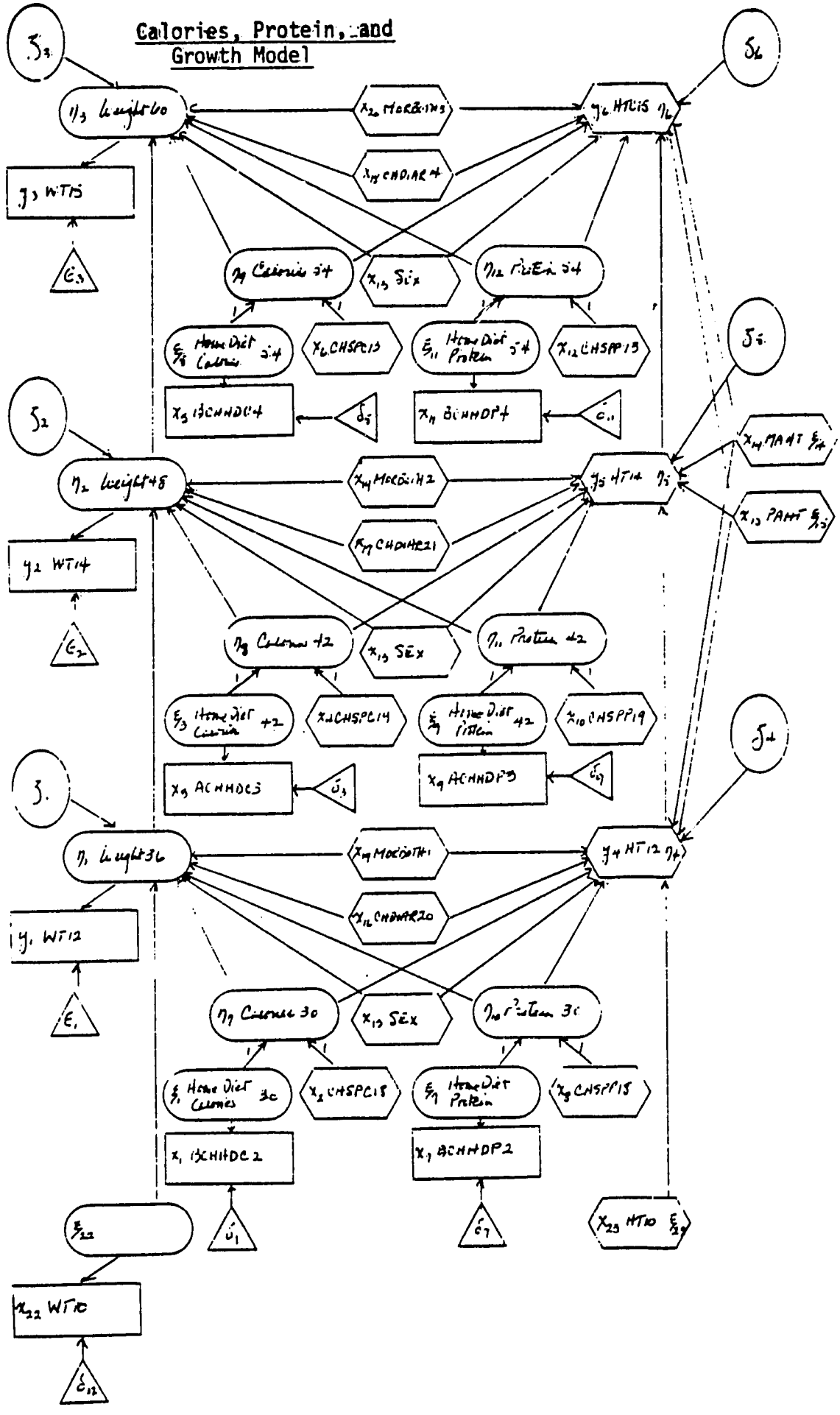


TABLE A.24

CALORIES, PROTEIN, AND GROWTH
Age 24-60 Months
Unstandardized Coefficients
All Villages Combined
x^2=251.66; df=393; x^2/df=.64; P=1.00
N=189

		WEIGHT			HEIGHT		
		36	48	60	36	48	60
y_i	Mean	11.8	13.6	15.1	85.7	93.1	99.5
	Standard deviation	1.3	1.3	1.5	3.8	4.0	4.1
$-\beta_{ij}$	Lagged values		.96*	1.04*		.95*	.98*
	Calories	0	-.0002	-.0004*	0	0	0
	Protein	0	0	.0032*	.0075	0	.0035*
γ_{ik}	SEX	.15	0	0	0	0	-.58*
	MAHT				.05*	.03*	0
	PAHT				.04*	.03*	0
	DIARRHEA	0	0	-1.65	-2.22	-3.93	-4.14
	MORBIDITY	0	0	-.17	0	0	0
	WT10 (24 months)	.98*					
	HT10 (24 months)				.89*		
	$R^2 = 1 - \psi$*	.76	.88	.89	.81	.89	.90
	F	98.31	217.69	243.49	95.67	184.30	197.44
	p <	.001	.001	.001	.001	.001	.001

TABLE A.25

Calories, Protein, and Growth
Ages 12-48 Months
All Villages Combined

x^2=401.35; d.f.=393; x^2/d.f.=1.02; p=.37
N=256

Home Diet Error at 23%

		WEIGHT			HEIGHT		
		24	36	48	24	36	48
y_i	Mean	9.70	11.88	13.62	77.36	85.80	93.10
	Standard deviation	1.15	1.30	1.41	3.64	3.88	4.03
$-\beta_{ij}$	Lagged values		.98*	1.01*		.90*	.98*
	Calories	-.0005	0	0	-.0022*	0	.0005
	Protein	.0081*	0	0	.0232*	.0070	0
γ_{ik}	SEX	.11	0	0	-.31	0	-.27
	MAHT				.05*	.03*	0
	PAHT				.05*	.03*	0
	DIARRHEA	-.49	0	0	-2.20	-2.67	-3.12
	MORBIDITY	0	0	-.16	0	0	0
	WT6 (12 months)	.82*					
	HT6 (12 months)				.97*		
	$R^2 = 1 - \psi^*$	.64	.79	.88	.70	.82	.89
	$F_{k, N-k-1}$	72.39	158.12	293.72	71.09	141.32	255.27
	p <	.001	.001	.001	.001	.001	.001

TABLE A.26
CALORIES, PROTEIN, DIARRHEA, AND GROWTH -- REDUCED FORM
Ages 12-48 Months
All Villages Combined
N=256

		WEIGHT			HEIGHT		
		24	36	48	24	36	48
Calories	18 mo.	-.0005	-.0005	-.0005	-.0022	-.0020	-.0020
	30 mo.	0	.0003	.0003	0	.0005	.0005
	42 mo.	0	0	-.0001	0	0	.0005
Protein	18 mo.	.0081	.0080	.0081	.0232	.0210	.0206
	30 mo.	0	-.0002	-.0002	0	.0070	.0068
	42 mo.	0	0	.0011	0	0	.0009
Diarrhea	12-24 mo.	-.4924	-.4847	-.4905	-2.2006	-1.9868	-1.9480
	24-36 mo.	0	-.4296	-.4346	0	-2.6694	-2.6172
	36-48 mo.	0	0	.1323	0	0	-3.1169

TABLE A.27

Calories, Protein, and Growth
Ages 12-48 Months
All Villages Combined

x^2=394.27; d.f.=393; x^2/d.f.=1.00; p=.47
N=256

Home Diet Error Set at 10%

		WEIGHT			HEIGHT		
		24	36	48	24	36	48
y_i	Mean	9.70	11.88	13.62	77.36	85.80	93.10
	Standard Deviation	1.15	1.30	1.41	3.64	3.88	4.03
$-\beta_{ij}$	Lagged values		.98*	1.01*		.90*	.98*
	Calories	-.0013*	0	0	-.0044*	0	0
	Protein	.0128*	0	0	.0357*	.0095	0
γ_{ik}	SEX	.14	.10	0	0	0	-.26
	MAHT				.05*	.03*	0
	PAHT				.05*	.03*	0
	DIARRHEA	-.64	0	0	-2.57*	-2.70	-3.13
	MORBIDITY	0	0	-.16	0	0	0
	WT6 (12 months)	.80*					
	HT6 (12 months)				.96*		
	$R^2 = 1 - \Psi$*	.66	.79	.88	.72	.82	.89
	$F_{k, n-k-1}$	78.93	155.74	292.64	77.95	140.94	254.21
	p <	.001	.001	.001	.001	.001	.001

TABLE A.28
Calories, Protein, and Growth
All Villages Combined

Home Diet Error Set at 40%

x^2=409.40; d.f.=393; x^2/d.f.=1.04; p=.27
N=256

		WEIGHT			HEIGHT		
		24	36	48	24	36	48
y_i	Mean	9.70	11.88	13.62	77.36	85.80	93.10
	Standard deviation	1.15	1.30	1.41	3.64	3.88	4.03
$-\beta_{ij}$	Lagged values		.98*	1.01*		.90*	.98*
	Calories	0	.0003	0	0	.0008	.0004
	Protein	.0052*	0	0	.0139*	.0060	0
γ_{ik}	SEX	0	0	0	-.37	0	-.26
	MAHT				.05*	.03*	0
	PAHT				.05*	.03*	0
	DIARRHEA	0	0	0	-1.89	-2.63	-3.13
	MORBIDITY	0	0	-.15	0	0	0
	WT6 (12 months)	.82*					
	HT6 (12 months)				.98*		
	$R^2 = 1 - \psi^*$	.64	.79	.88	.70	.82	.89
	$F_{k, N-k-1}$	73.28	158.65	294.90	71.41	143.26	256.57
	p <	.001	.001	.001	.001	.001	.001

TABLE A.29
Calories, Protein, and Growth
Fresco Villages Only
x^2=229.14; d.f.=309; x^2/d.f.=.74; p=1.00
N=122

		WEIGHT			HEIGHT		
		24	36	48	24	36	48
y_i	Mean	9.27	11.41	13.16	76.23	84.54	91.85
	Standard deviation	1.20	1.27	1.35	3.73	3.93	3.98
$-\beta_{ij}$	Lagged values		.87*	.98*		.88*	.96*
	Calories	-.0007	0	0	.0019	.0015	0
	Protein	.0072	0	0	-.0199	0	.0127*
γ_{ik}	SEX	.35*	.26*	0	0	.79*	-.45
	MAHT				.05*	.03*	0
	PAHT				.05*	.03*	0
	DIARRHEA	-.50	0	0	0	-2.20	-4.00
	MORBIDITY	.47	0	0	-1.27	.80	-.76
	WT6 (12 months)	.80*					
	HT6 (12 months)				.96*		
	$R^2 = 1 - \psi$*	.63	.76	.85	.65	.81	.88
	$F_{k, N-k-1}$	32.29	62.00	111.83	26.18	59.43	108.29
	p <	.001	.001	.001	.001	.001	.001

FIGURE A.16

Growth in Stature - Cross Sectional Model

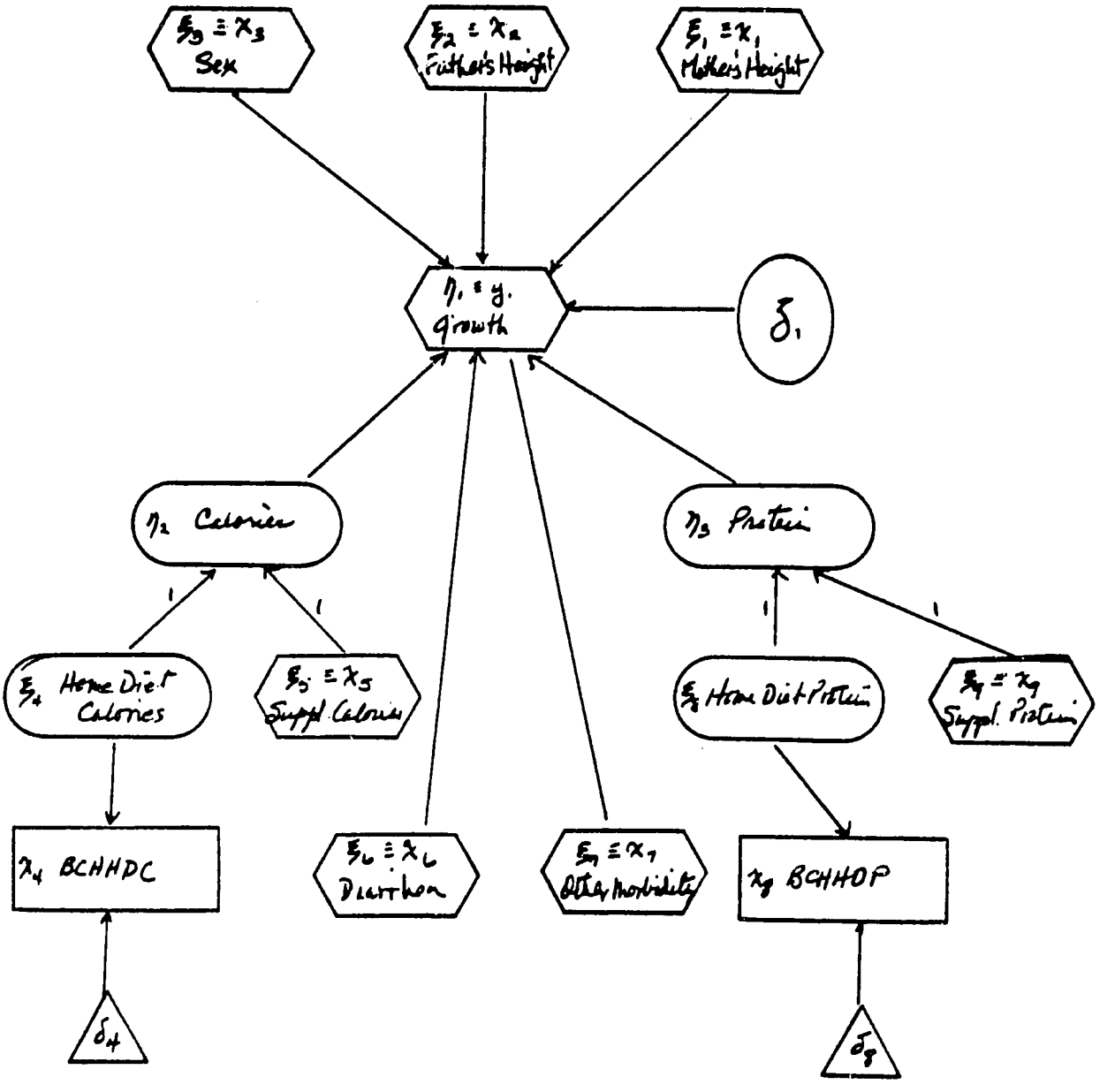


TABLE A.30

GROWTH IN STATURE -- CROSS SECTIONAL MODELS: UNSTANDARDIZED COEFFICIENTS

		12-24	24-36	36-48	48-60	60-72	72-84
y_i	Mean	8.66	8.45	7.29	6.42	5.56	5.09
	Standard deviation	2.03	1.67	1.33	1.33	1.27	1.51
$-\beta_{ij}$	Calories	-.0040*	0	0	0	0	0
	Protein	.0333*	.0062	0	0	0	0
γ_{ik}	Mother's height	.0155*	.0155*	.0155*	.0155*	.0155*	.0155*
	Father's height	.0145*	.0145*	.0145*	.0145*	.0145*	.0145*
	Sex	0	0	-.3571*	-.3571*	0	0
	Diarrhea	-2.2632*	-1.8612	-4.0850*	-4.0851*	0	-4.0851*
	Other morbidity	0	0	0	0	0	0
	$R^2 = 1 - \psi$*	.06	.04	.06	.07	.03	.11
	$F_{k, n-k-1}$	2.81	1.74	2.52	2.32	0.53	0.49
	$p <$	.01	.30	.025	.05	---	---
	N	294	295	328	222	121	37
	χ^2	4.24	.20	1.26	9.53	1.60	2.23
	df	49	49	49	49	49	49
	χ^2/df	.09	.00	.03	.19	.03	.05
	P	1.00	1.00	1.00	1.00	1.00	1.00

II.A.7. Physical Growth, Current Nutrition, and Verbal Development

A central focus of the initial design of the INCAP study was upon the effect which supplementation -- the alleviation of calorie/protein deficiency might have upon the development of the child. Particular attention has been paid to cognitive development. Because of this interest, a comprehensive and eclectic battery of development and psychometric scales was periodically administered to children in the villages.

In a series of publications and reports, members of the INCAP study team have shown that there were significant correlations between supplementation and a number of psychometric tests, specific to certain age levels by sex, when controlling for indicators of family background.

In this section we will, on the one hand, narrow the scope of analysis by focusing attention upon two closely related indicators of verbal development; on the other hand, we hope to extend prior analyses by explicitly taking into account the developmental character of verbal performance (using the longitudinal aspects of the data base), including home diet as the major component of nutritional intake, taking account of and estimating the errors of measurement in both the dependent variables -- the verbal tests -- and the measures of home diet, and including a considerably more comprehensive array of control variables which might have prior and independent effects upon verbal performance.

The initial model to be estimated is shown as a path diagram in Figure A.17. This model is a combination of a constrained factor analytic model and a structural equation (regression) model. The right hand side of the "picture" represents the factor analytic model. The hypothetical true

value of verbal facility at each of five age levels is represented by two fallible indicators, "naming" and "recognition." (These tests are analogous to the widely used Peabody Picture Vocabulary, assessing the ability to correctly name an object shown in a picure and to select a picture to go with a name.)

The variance of each of these tests is partitioned into a communality, the hypothetical verbal factor, and a uniqueness, represented by the ε_i shown in cirles in the diagram. This uniqueness, in turn, includes both measurement error and specificity -- the systematic content tapped by one performance, e.g., naming, which is not captured by the other, e.g., recognition.

The model permits both occasion correlations between errors -- e.g., tester effects in making consecutive assessments of related performances; and retest correlations between uniqueness -- e.g., correlations over time between the specific aptitudes tapped by each test. These correlations are represented by the double-headed curved arrows at the extreme right hand edge of the diagram.

While the right hand side of the diagram represents the measurement model, the central and left hand portions represent the simultaneous structural equation system. The vertical arrows represent what might be called the "stability" of verbal development: those who perform better at one age are likely to perform better at the next. This is, of course, due to the cumulative nature of vocabulary acquisition, the advantage which fluency at one point provides for further subsequent language acquisition, and possibly more stable latent factors leading to individual differences. Congruent with much prior research (cf. Benjamin Bloom, Stability and Change in Human Characteristics), we would expect these "stabilities" to progressively increase with age.

Finally, down the left edge of the diagram, is an array of independent variables which may plausibly affect verbal performance either directly as, say, by education or exposure, or as proxies for unmeasured factors affecting development or socialization. The first of these variables, height at age 36 months, we take to be a long-term cumulative proxy for prior nutrition and absence of morbidity. There is no known genetic linkage between height and cognitive development. Sex is included, of course, because of the ubiquity of both earlier development and, especially in the case of verbal fluency, superior development of females over males. The possible effect of family size upon verbal development has been the subject of antithetical hypotheses and contradictory findings. On the one hand is the argument, especially favored by economists, that the limitation of family size reflects a more intensive investment by the parents in the socialization of each child. On the other hand, the presence of a large number of siblings and members of an extended family provide a denser social environment permitting -- requiring -- more verbal communication. The existence of contradictory evidence in large-scale studies (cf. Velandia, Grandon, and Page, 1978) suggests that both of these abstract interpretations may be culturally dependent. That is, family size may not be universally related to the educative experiences of the child in the same way.

Most of the other variables in the array are fairly traditional indicators of familial advantage which, separately or collectively, may serve as proxies for the quality of the home environment. The variable with the accronym "MODVOC" is an index composed of a mother's vocabulary test together with a modification of the Brown-Inkeles OM Modernity scale. "ZMARD" and "ZPARD" are indices of mother's and father's literacy, respectively. These three are fairly direct indicators of verbal exposure. The variable "DIED3" refers

to the number of stillborn and deceased elder siblings of the child. As was mentioned in an earlier section, this variable may serve as a proxy for some otherwise unmeasured aspects of maternal care -- whether health practices or other aspects of care cannot be known.

This entire model which has been described in some detail above with reference to the path diagram shown in Figure A.17 is also represented by the set of matrix equations shown as Equations A.5 and A.6 in Appendix A.11.b. As before, the unknown elements of the model were estimated using the LISREL maximum likelihood program.

One technical difference between this model and the estimation of models presented in prior (and subsequent) sections should be noted. The covariance matrix which was "fitted" to the model, in this case, was based upon "pairwise" deletion of missing data. That is, each variance and covariance is based upon all the cases for which data were available for the variable or pair of variables under consideration. While this maximal use of data is congruent with the emphasis of the maximum-likelihood estimation of population, rather than sample, parameters, it does run the hazard of producing internally inconsistent statistics and a non-Gramian matrix which has no inverse. In this instance, a very good fit was obtained to the matrix generated by pairwise deletion -- indeed, better than an alternative solution using "listwise" deletion for the far fewer cases for whom data were present for all variables. This procedure was not used systematically throughout the analysis, however, because of the frequency with which it failed.

The estimates of the measurement model are shown in Table A.33 in standardized form -- i.e., as correlations rather than covariances. Two things are of interest in this solution. First is the very low correlations between the uniquenesses of the tests -- the off-diagonal elements in the

bottom matrix presented. Indeed, the occasion correlations (immeditaely below the diagonal) are estimated at approximately zero. The second is the increasing relative importance of the "naming" test, over the "recognition" test in the estimation of vocabulary with increasing age. This is shown both on the diagonal element estimates of the proportion of variance which is unique, and in the relative factor weights in the bottom row. The factor scores were normalized by ascribing a weight of one to the "naming" test. The relative weight of the "recognition" test shows a monotonic decline with age. This is likely due to the greater discriminatory power of the test calling for active rather than passive vocabulary with increasing age.

The estimates of the structural relationships for this model are shown, again in standardized from -- as "normalized" partial regression coefficients in Table A.34. The strongest relationships are, as would be expected, the lagged values of the verbal factors, referred to above as "stabilities." These coefficients do, as anticipated, increase with age.

The most important substantive result of this estimation lies in the large and highly significant predictive effect of height at age 36 months on verbal performance, even when allowing for the concommitant and somewhat confounded effects of a large number of background variables. Not only does height, which we take to be a proxy for cumulative health from conception to age 36 months, affect performance at age three, but it also predicts positive growth -- learning -- from year to year up through age seven. This is very clear when we look at the first row of the reduced form estimates in Table A.35. One standard deviation difference in height at age three is associated with about one-third of a standard deviation difference in level of performance at that age. By age seven, it is associated with almost one-half of a standard deviation difference.

These reduced form coefficients for height can be reproduced almost

exactly from the structural form coefficients for height and for the "stabilities" of verbal performance alone. For example:

At age 48 months $.420 \cong (.35)(.80) + .14$;

At age 60 months $.430 \cong (.42)(.83) + .08$;

At age 72 months $.410 \cong (.43)(.87) + .04$;

At age 84 months $.480 \cong (.41)(.85) + .13$.

From Table A.15 of Section A.4, above, we obtain the following tally of heights at age 36 months by sex and treatment group.

Table A.31

Mean Heights by Treatment and Sex at Age 36 Months

Sex	Treatment		Difference
	Fresco	Atole	
Boys	85.21	87.03	1.82
Girls	83.06	86.32	3.26
Mean	84.14	86.68	2.54

Once again, we see the interaction between sex and treatment which has been noted before. The average difference in height at age 36 months between Atole and Fresco villages is 2.54 cm. which is .61 standard deviations of the measures of height at that age for the group as a whole. Using the coefficients of the first two rows of the reduced form in Table A.35 to estimate the verbal factor differences in standard deviation units, we have the results in Table A.32.

Table A.32

Observed and Predicted Verbal Factor Differences Between Treatment Groups

Age	Predicted Difference	Observed Diff. in Standard Units		
		Naming	Recognition	Weighted Average
36	.21	.25	.02	.12
48	.28	.42	.28	.35
60	.32	.28	.32	.30
72	.34	.34	.30	.33
84	.41	.47	.32	.41

We have been using height by age as a proxy for cumulative prior nutrition and absence of morbidity. We do not presume that the linkage between height and verbal development is direct or proximate; rather, that the healthier children were more exploratory, active, and expressive, and, consequently, elicited a more favorable and responsive social environment (cf., Chavez, Martinez, and Yashine, 1975). The associations reported here do not distinguish between possible physiological direct effects upon capacity to learn and indirect effects through modified social environments. While it is theoretically important to sort out and specify these paths, from a policy point of view the fact that early growth affects verbal development is important. (The extent to which height, per se, among equally healthy children, may generate positive social responses could inflate the coefficients.)

Among the other independent variables which were included, primarily as controls, in the model, sex shows the usual earlier verbal development and continued superiority of girls. There are collinearities among the remaining variables which make substantive interpretation of individual statistics quite tenuous. For example, the three variables dealing with

parental literacy -- "MODVOC," "ZMARD," and "ZPARD" -- all have positive coefficients in the reduced form equations, but only one coefficient reaches statistical significance in the structural equation model, and that very possibly by chance since there are 15 such coefficients. Because of the intercorrelation among these variables and their conceptual interrelationships, if estimation of their joint effect were of primary interest, they should probably be blocked together or treated as a common factor. Similarly, while the coefficients for family size are here consistently positive, which is consistent with the "social density" theory, the coefficients for nuclear families (STRUC3 in the bottom row), which tend to be larger, on the average, are negative.

Since our interest in this study is primarily on the effects of nutrition and health, we have not made a real effort to sort out the independent effects of these more sociological variables but, rather, rest content to treat them as controls.

In the second model of this section, we look at the effects of diet and morbidity during the intervening annual periods from age 36 to age 60 months when holding stature at age 24 months and other background variables constant. This model explicitly addressed the policy issue of the value of nutritional intervention during later years of childhood as contrasted with the first two years of life when using verbal development as a criterion.

The right hand side of the model is identical to the first three years of the model as depicted in Figure A.17 above. The model again includes the vertical arrows -- the "stabilities" or lagged values of verbal performance. On the left hand side, however, height is measured at age 24 months rather than 36 months. The major change is the insertion of four causal variables at each age level: calories consumed during the prior year, protein consumed during the prior year, incidence of diarrhea, and

incidence of ailments other than diarrhea during the prior year. Calories and protein are calculated by adding the values, in kcal units, of home diet and supplementation calories and protein, respectively, and allowing for a 23% measurement error in each of the measures of home diet.

Figure A.18 below, illustrates one cross-section of the model as a path diagram. Equation A.7 and A.8 (in Appendix A.11.b) present the longitudinal model in matrix equation form.

Table A.36, entitled "Stature, Diet, and Verbal Development: Ages 36, 48, and 60" presents the estimates of the structural relationships. This time they are presented as unstandardized coefficients.

The coefficients of determination, R^2, are almost identical to those previously estimated which suggests immediately that the alterations to the model do not affect the predictability of verbal performance. Once again, the "stabilities" have the largest explanatory effect, and height at age 24 months has a substantial and significant effect.

The only other statistically significant finding in this estimated model is the negative effect of morbidity between the ages of 24 and 36 months upon verbal performance at age 36 months. The estimates of nutritional intake during these periods, on the other hand, do not show any relationship to subsequent performance.

On the whole, this analysis points to the predominant importance of nutrition and health during early childhood and the negligible differentiating effect of subsequent nutritional disparities. This is congruent with the prima facie evidence from the physical growth charts presented in Section A.4 which showed almost parallel growth slopes after age 36 months in the Atole and Fresco villages; and, also, the analysis of Section A.6 which showed that only the earliest estimation of dietary intake -- between ages 12 and 24 months -- affect physical growth significantly.

Before leaving the topic, we should report a series of cross-sectional analyses of verbal performance factors at each age level. These models are described in List A.4, listing the variables for one cross-section, Figure A.18 which depicts the path model, and Equation A.9 which presents the model in matrix form.

The major substantive difference between these cross-sections and the prior structural model is the omission of the series of lagged values and the simultaneous estimation of the equations. These are more analogous to ordinary least squares regression equations. (Indeed, except for the inclusion of a fixed error term in the measurement of home diet, these could be estimated by ordinary regression.)

These regressions differ from the prior model in two other respects. First, each regression is based on all cases available at that age level, while the prior structural model was based on listwise deletion of cases missing values on any variable at any age level. Second, it was possible to reintroduce several background variables which had to be eliminated from the prior model because of the size of the data matrices.

Substantively, the most apparent difference between these cross-sectional results and the structural estimates from Table A.37 is the reemergence of the predictive importance of the proportion of protein in the diet at later ages for verbal performance. The main contributory factor to differences in the proportion of protein in the diet is the consumption of atole. But, as we described in Section A.5, the consumption of supplementation has its own very high "stability" and is also strongly affected by the mother's attendance at the centers and the availability of atole rather than fresco. That is, the same children who enjoyed a relatively high protein diet at an early age and, hence, enjoyed cumulatively better nutrition at ages two and three,

continue to enjoy the supplementation at a later age. Our previous analysis has shown that early growth does affect verbal performance at an early age. The omission of the lagged values of verbal performance would create a spurious association between later consumption and performance.

$$\begin{tikzcd} & V_1 \arrow[r] & V_2 \arrow[r] & V_3 \\ PP_1 \arrow[ur] \arrow[r] & PP_2 \arrow[r] & PP_3 & \end{tikzcd}$$

If the true causal model here, as sketched above, where V represents verbal performance and PP the proportion of protein in the diet, one would find an association between PP_3 and V_3 even though, by hypothesis, there is no causal connection.

REFERENCES

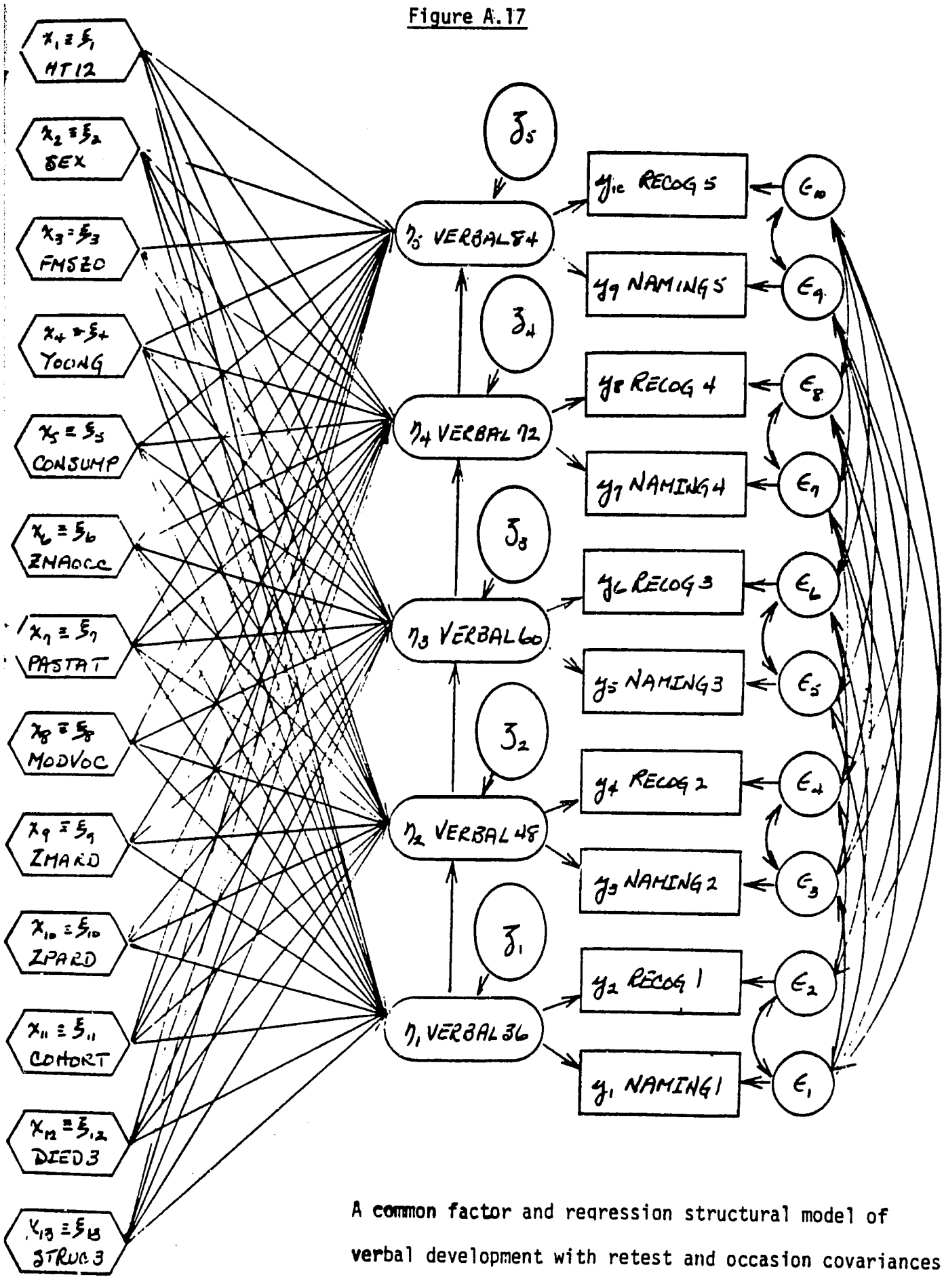
Chavez, A.; Martinez, C.; and Yashine, T. "Nutrition, Behavioral Development, and Mother-Child Interaction in Young Rural Children." Federation Proceedings 34 (1975): 1574-82.

Greene, Lawrence S., ed., Malnutrition, Behavior, and Social Organization. New York: Academic, 1977.

Page, Ellis B., and Grandon, Gary M. "Family Configuration and Mental Ability: Two Theories Contrasted With U.S. Data." American Educational Research Journal 16 (Summer 1979): 257-72.

Velandia; Wilson; Grandon, Gary M.; and Page, Ellis B. "Family Size, Birth Order, and Intelligence in a Large South American Sample." American Educational Research Journal 15 (Summer 1978): 399-416.

Figure A.17



A common factor and regression structural model of verbal development with retest and occasion covariances among test uniqueness.

TABLE A.33

VERBAL DEVELOPMENT - CORRELATIONS AMONG SCORES

Factor Scores		36	48	60	72	84
	36	1.00				
	48	.84	1.00			
	60	.76	.88	1.00		
	72	.71	.82	.92	1.00	
	84	.68	.79	.88	.95	1.00
"Naming"						
	36	1.00	(452)	(336)	(222)	(110)
	48	.63	1.00	(354)	(235)	(117)
	60	.65	.71	1.00	(239)	(119)
	72	.65	.65	.77	1.00	(119)
	84	.65	.65	.73	.87	1.00
"Recognition"						
	36	1.00	(453)	(338)	(223)	(111)
	48	.58	1.00	(355)	(235)	(117)
	60	.52	.56	1.00	(240)	(120)
	72	.49	.58	.64	1.00	(119)
	84	.38	.46	.59	.65	1.00

Test Uniqueness Covariances		36		48		60		72		84	
		Nam.	Rec.	Nam.	Rec.	Nam.	Rec.	Nam.	Rec.	Nam.	Rec.
36	Nam.	.30									
	Rec.	-.00	.30								
48	Nam.	.03		.25							
	Rec.		-.01	-.00	.28						
60	Nam.	.06		.02		.18					
	Rec.		-.02		-.03	-.00	.34				
72	Nam.	.04		-.02		-.01		.11			
	Rec.		-.04		.03		.06	-.00	.41		
84	Nam.	.07		-.01		-.04		.02		.07	
	Rec.		-.10		-.02		.07		.11	-.00	.45
Relative Factor Weights		1.00	1.34	1.00	1.02	1.00	.80	1.00	.66	1.00	.62

TABLE A.34

Verbal Development, Ages 36 to 84 Months, Standardized Coefficients
All Villages Combined
x^2=42.78; df=166; x^2/df=.26; p=1.00
N=105 (minimum)

		36	48	60	72	84
y_i	Standard Deviations	3.41	4.22	4.61	4.30	3.91
$-\beta_{ij}$	Lagged Values	--	.80*	.83*	.87*	.85*
γ_{ik}	HT12	.35*	.14	.08	0	.13*
	SEX	-.18	0	0	-.09	0
	FMSZO	0	0	0	0	.15*
	YOUNG3	0	0	-.13	.07	0
	CONSUMP	.24*	0	0	0	.06
	ZMAOCC	0	0	0	0	-.05
	PASTAT	-.12	0	0	0	-.10*
	MODVOC	0	0	0	0	0
	ZMARD	0	.09	0	0	.05
	ZPARD	0	0	0	0	.11*
	COHORT	0	0	0	0	0
	DIED3	0	0	-.10	0	.17*
	STRUC3	0	-.11	0	-.14*	-.16*
	$R^2=1-\psi^*$	.25	.74	.82	.88	1.00
	F_k, N-k-1	2.31	18.34	27.45	48.00	--
	p <	.05	.001	.001	.001	.001

TABLE A.35

VERBAL DEVELOPMENT -- REDUCED FORM -- STANDARDIZED

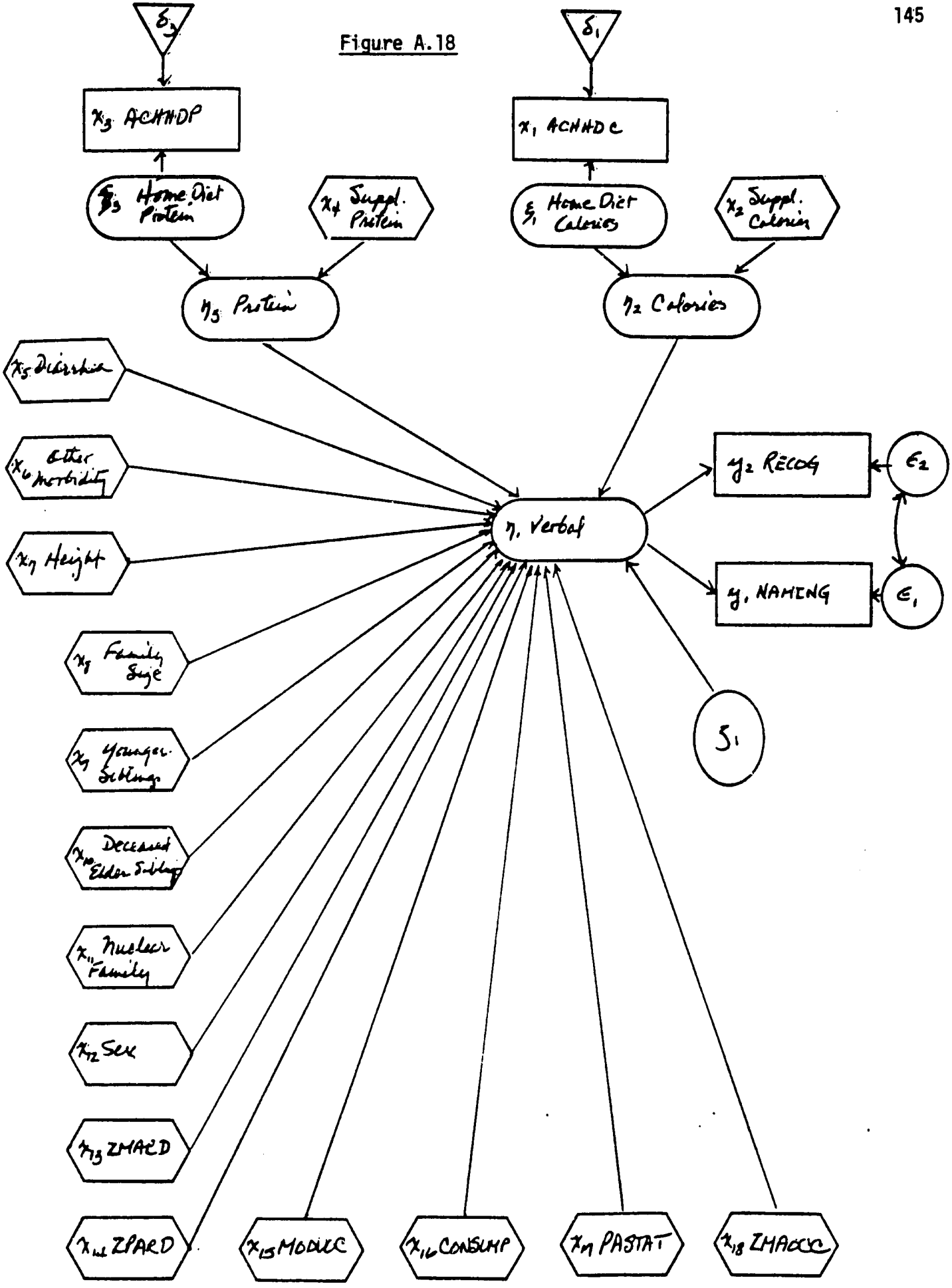
Acronym Symbol	36	48	60	72	84
HT12	.35	.42	.43	.41	.48
SEX	-.18	-.22	-.14	-.22	-.16
FMSZO	.09	.13	.17	.15	.28
YOUNG3	-.06	.02	-.11	-.02	.00
CONSUMP	.24	.11	.16	.16	.20
ZMAOCS	-.02	-.03	-.07	-.02	-.07
PASTAT	-.12	-.07	-.09	-.09	-.18
MODVOC	.05	.02	.09	.13	.14
ZMARD	.06	.14	.10	.15	.18
ZPARD	.07	.03	.00	.02	.10
COHORT	.01	.03	.03	.03	.01
DIED3	-.06	-.02	-.11	-.15	.05
STRUC3	-.01	-.12	-.11	-.23	-.36

TABLE A.36

STATURE, DIET, AND VERBAL DEVELOPMENT
Ages 36, 48, 60
Unstandardized coefficients
x^2=201.23; df=458; x^2/df=.44; P=1.00
N=243

		36	48	60
η_i	Standard Deviations	3.42	4.04	4.03
$-\beta_{ij}$	Lagged variables		.96*	.82*
	Calories	0	0	0
	Protein	-.0112	0	0
γ_{ik}	SEX	-1.2430	0	0
	ZMARD	.4767	.4306	0
	MODVOC	.2209	0	0
	CONSUMP	0	-.3492	0
	FMSZO	.1119	.1836	0
	CHDIAR20, 21, 14	0	0	0
	MORBOTH2, 3, 4	-2.1833*	0	0
	DSTRUC3	-1.8415	0	0
	HT10 (24 mo.)	.3473*	.0943	.1563*
	$R^2 = 1 - \psi$*	.25	.73	.84
	F	6.91	57.96	109.83
	P <	.001	.001	.001

Figure A.18



STATURE, DIET, AND VERBAL FACTOR SCORES -- CROSS SECTIONS

TABLE A.37
STATURE, DIET, AND VERBAL FACTOR SCORES -- CROSS SECTIONS
UNSTANDARDIZED COEFFICIENTS

		36	48	60	72	84
η_1	Standard deviations	3.24	3.99	4.50	4.22	3.96
$-\beta_{ij}$	Calories	0	-.0051*	-.0050*	-.0058*	-.0040*
	Protein	0	.0244*	.0247*	.0285*	.0178
γ_{ik}	Diarrhea	0	0	7.23	0	34.52
	Other morbidity	-1.37	0	-2.16*	-1.22	-1.93
	Height	.30*	.34*	.29*	.10	.33*
	FMSZO	.12	.19	.24	0	.49
	YOUNG3	-.66	-.55	-1.91*	0	-1.52*
	DIED3	-.34*	-.30	-.52*	-.35	-.43
	DSTRUC3	-2.85*	-2.55*	0	-1.28	0
	SEX	-1.09	-1.04*	-.79	-1.24	0
	ZMARD	.34	.88	.89*	.60	.66
	ZPARD	0	0	0	0	0
	MODVOC	0	.44	.71*	1.36*	0
	CONSUMP	.55	0	0	0	0
	PASTAT	-.32	0	0	0	0
	ZMAOCC	0	0	0	0	-.47
	$R^2 = 1 - \psi^*$	.26	.22	.26	.26	.47
	$F_{16,\ N-k-1}$	7.14	5.89	5.35	2.71	2.05
	$p <$	.001	.001	.001	.005	.05
	N	354	367	262	142	54
	χ^2	25.63	20.89	21.34	19.31	25.13
	χ^2/df	.13	.11	.11	.10	.13
	p	1.00	1.00	1.00	1.00	1.00

II.A.8. School Enrollment by Age 7

Originally it had been planned to analyze enrollment in school through the use of a series of dichotomous discriminant functions, each predicting the likelihood of enrollment by a given age. This plan had to be abandoned because of limitations of the data. Enrollment data were only available for those entering in January and completing a full year as evidenced by completion of the year-end examination. There were too few children in the sample who had reached age seven by January of 1977, and too small a proportion who entered at age six to make these analyses feasible.

The predictive equation which was estimated, then, is limited to the one group of 114 children who were over age seven on January 1978 -- the last year for which data were collected.

The model which is portrayed in Figure A.19 of this section (and Equation A.10 in Appendix A.11.b) is similar to an ordinary least squares regression with a dichotomous, one or zero valued, dependent variable except for the inclusion of the error terms in home diet measurement and the factor score structure of the verbal test. The values of these coefficients -- the error terms and factor coefficients -- were taken from the prior solutions reported in earlier sections as constants rather than as parameters to be re-estimated.

The coefficients estimated for this model are presented in Table A.38. While the model fit the data well, as indicated by the values of X^2 and R^2, only three of the coefficients are individually statistically significant. The verbal factor, discussed in the previous section, is positively related to enrollment as might be anticipated. The negative coefficient of the factor, COHORT, indicates a positive secular trend within this

sub-sample. That is, the children who were born earlier are less likely to have enrolled by age seven than those who were born later.

The most striking coefficient is that attached to the dummy variable representing the large Atole village. Far fewer children in this village enrolled in school than in the others. This large village contrast was not anticipated and is not explainable from the data at hand for this cohort. Further analyses of economic differences between villages in subsequent sections, and their implications for the value of children's labor, suggest reasons for these contrasts.

Diet during the immediately preceding year does not appear to have any effect upon the likelihood of enrolling in school when controlling for verbal development and other factors. Once again it is arguable that the verbal development depends upon early diet and morbidity, but the within village variation between age six and seven does not have any differentiating effect.

Figure A.19

School Enrollment by Age 7 Among Children Over 7 in January, 1978

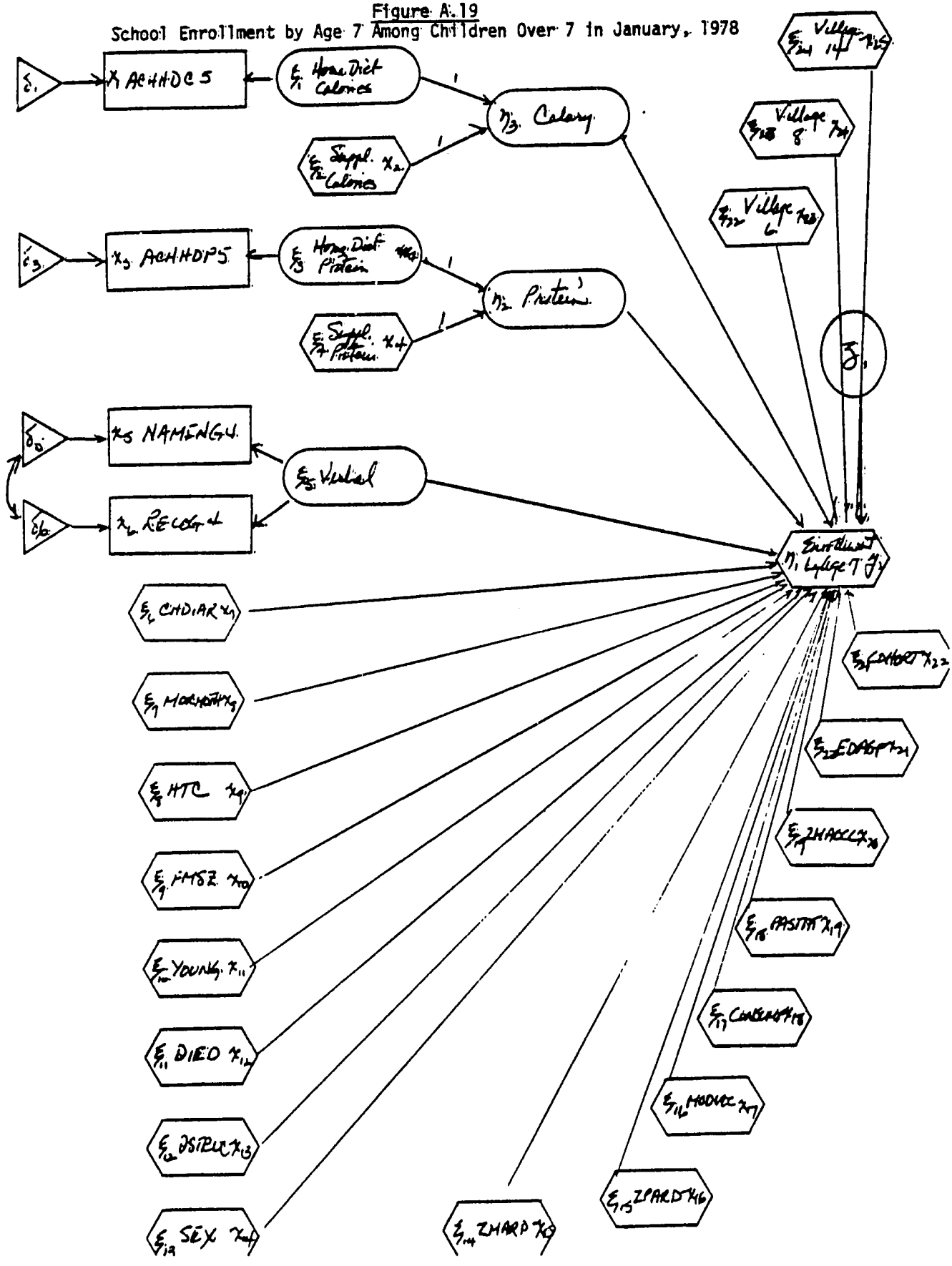


Table A.38
School Enrollment by Age 7
Among Children Over Age 7 on January, 1978.
All Villages Combined.
χ^2=18.97; df=304; χ^2/df=.06; P=1.00
N=114

		Enrollment
y_i	Proportion	.46
	Standard Deviation	.50
$-\beta_{ij}$	Protein	-.0030
	Calories	.0003
γ_{ik}	Verbal factor	.0275*
	CHDIAR15	0
	MORBOTH4	0
	HTC16	.0135
	FMSZ6	0
	YOUNG6	.0566
	DIED6	0
	DSTRUC6	0
	SEX	-.0909
	ZMARD	0
	ZPARD	.0656
	MODVOC	.0684
	CONSUMP	.0583
	PASTAT	0
	ZMAOCC	.0506
	EDASP	0
	COHORT	-.0252*
	D6 Atole-Large	-.3127*
	D8 Fresco-Small	0
	D14 Atole-Small	0
	$R^2 = 1 - \psi^*$	.42
	$F_{22, 91}$	2.98
	$p <$	.001

II.A.9. Teachers' Assessment of School Achievement

While at one time nationally administered objective tests were recorded each year for children enrolled in school, in recent years, this practice has been modified so that the recorded scores are at least partially modified by subjective teacher judgment. This means that a variety of social and psychological factors may play directly upon these assessments independently from whatever effect they may have upon measured academic performance. The educational literature abounds with accounts of the impact of docility, dress, language, and family background upon teachers' perceptions of "good students." There is also a well-documented tendency for teachers to "normalize" their estimates of performance, that is, to assess performance relatively to the distribution of achievement in a class. The "good" student in one school might appear mediocre in another where average standards are higher.

Because of this relativization of grades, and possibly idiosyncratic bases of judgment used by different teachers, the model developed here, like the preceding model of school enrollment, includes dummy variables representing the different schools. It is quite possible that there are different factors among the common variables in the model which affect different teachers -- i.e., that there are interactions between these variables and teachers.

The small number of cases in this sub-sample of enrolled children (N=83) effectively precludes separate, village by village, analyses of factors affecting teacher judgment. Moreover, without some knowledge of teacher characteristics, such interactions, if they exist, would not be readily interpretable. The model estimated, then, is a cross-sectional model aggregating the four villages. It is portrayed in Figure A.20 (and in

Equation A.10 in Appendix A.11.b). Note that the language and mathematics assessments, which are intercorrelated at r = .87, are taken as indices of a global assessment factor. Otherwise, the model is similar to that in the prior section on enrollment.

The estimates of this model are presented in Table A.39. Peculiarly, teachers' assessments in the two Atole villages are substantially lower than in the Fresco villages. This is true both in the raw scores used and of the partial coefficients while holding many variables constant.

Of greater substantive importance in the present analysis is the effect of verbal development upon teachers' assessments. Within schools, the standardized regression of assessment on verbal proficiency is .41. This is the most substantial of the independent predictors.

While not reaching statistical significance in this small sample, the effect of concurrent nutritional intakes -- protein and calories -- upon teachers' assessments is quite substantial. When transformed to standardized form, these coefficients are .34 and .27 respectively. These substantial, albeit non-significant, relations make it imprudent to conclude that concurrent dietary adequacy is irrelevant to classroom performance even when holding prior physical and verbal growth constant.

The other variables which show significant relationships in this model are of less interest in the present context. Family size and maternal modernity are negatively related to teacher assessment. Neither of these has substantial zero order relationship's to assessment, and the emergence of significant partials may be due to their confounding with the child's verbal performance. The significant effect of paternal status -- landholding farmers -- may well be a sociologically biasing factor affecting teacher judgment.

The clearest thread tying this analysis to the foregoing sections is that the child's verbal development, which is strongly affected by his or her early nutrition and health, has a substantial impact, on the average, upon teachers' assessments of performance in school. Current diet may also make some difference, though this is less substantial and we can have less confidence in this relationship.

Figure A.20
Teacher Assessment Model

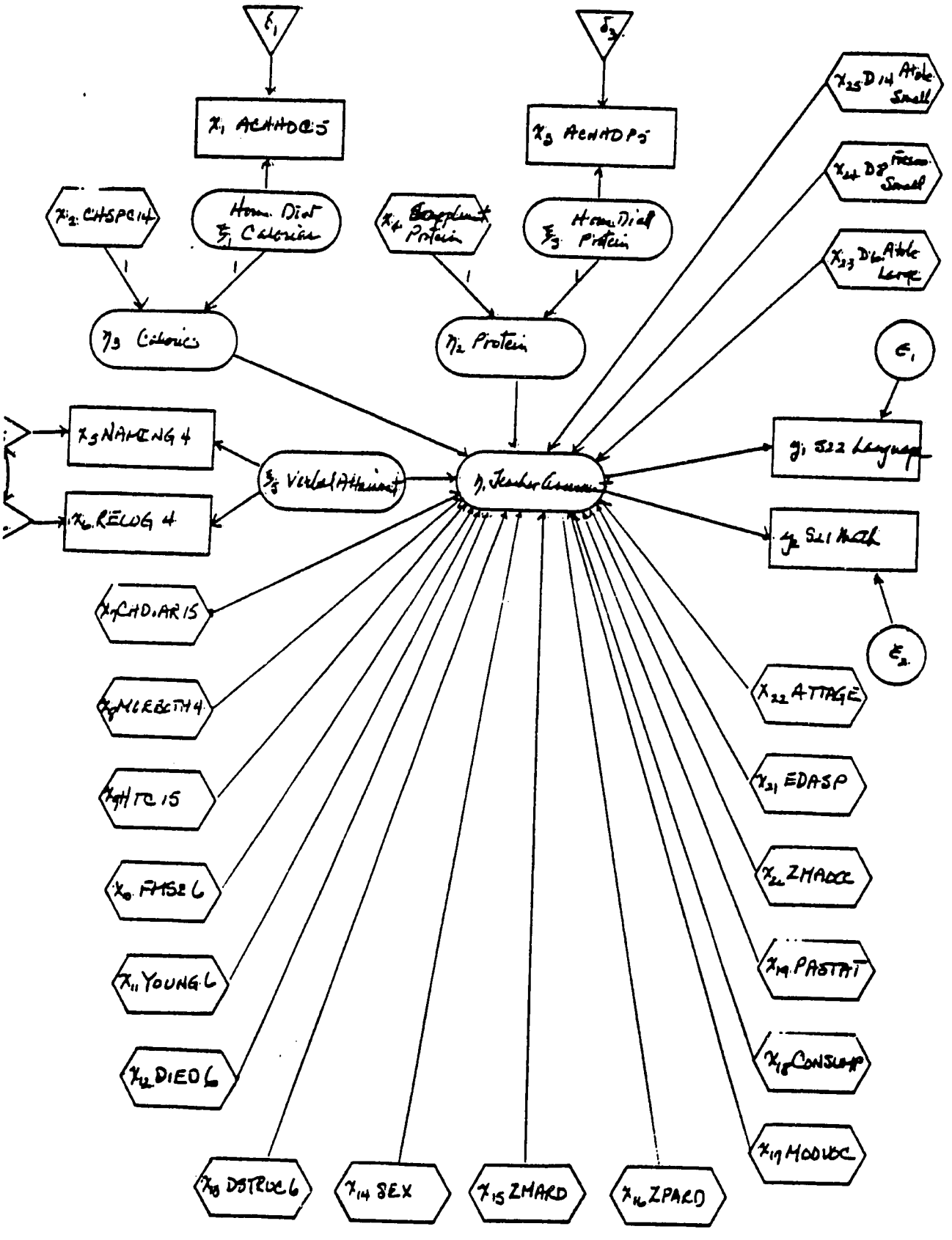


TABLE A.39

TEACHER ASSESSMENTS IN LANGUAGE AND MATHEMATICS
at End of First Completed Year of Schooling.
All Villages Combined.

χ^2=361.72; df=328; χ^2/df=1.10; P=.10
N=83

		S22 Language	S21 Mathematics	Composite Factor
y_i	Mean	57.51	55.89	
	Standard Deviation	20.72	19.87	19.83
$-\beta_{ij}$	Protein 2			.21
	Calories 3			.02
γ_{ik}	Verbal factor 5			1.87*
	CHDIAR15			0
	MORBOTH4			11.93
	HTC15			0
	FMSZ6			-3.34*
	YOUNG6			0
	DIED6			0
	DSTRUC6			8.00
	SEX			-4.97
	ZMARD			-2.91
	ZPARD			0
	MODVOC			-5.20*
	CONSUMP			3.76
	PASTAT			5.75*
	ZMAOCC			0
	EDASP			2.00
	ATTAGE			.35
	D6 Atole-Large			-21.56*
	D8 Fresco-Small			0
	D14 Atole-Small			-21.83*
	$R^2 = 1 - \psi$*			.59
	$F_{20, 62}$			4.38
	p <			.001
λ	Relative factor weights	1.00	.97	
ε	Unique variance	57.68	47.05	
ε/s_i^2	Proportion of variance unique	.13	.12	

TABLE A.40
ZERO ORDER CORRELATIONS AMONG A SUB-SET OF VARIABLES
IN THE TEACHER ASSESSMENT MODEL
N=83

	S21	S22	NAMING4	RECOG4	ACHHDC5	ACHHDP5	CHSPC14	CHSPP14
S21	1.00							
S22	.87	1.00						
NAMING4	.19	.22	1.00					
RECOG4	.08	.16	.66	1.00				
ACHHDC5	.28	.27	-.03	-.15	1.00			
ACHHDP5	.20	.24	-.12	-.16	.73	1.00		
CHSPC14	-.05	-.14	.11	-.08	-.11	-.20	1.00	
CHSPP14	-.23	-.29	.16	-.01	-.27	-.26	.67	1.00
CHDIAR15	-.08	-.07	.01	-.12	.02	-.07	-.06	-.04
MORBOTH4	.20	.26	.06	-.02	.01	.13	-.04	-.13
HTC15	.16	.10	.15	.11	.09	.13	.16	.16
FMSZ6	-.11	-.06	-.02	.07	-.07	-.03	.14	.19
YOUNG6	-.14	-.09	-.08	-.16	-.09	-.01	-.01	.09
DIED6	-.07	-.18	.06	.06	-.16	-.24	.18	.23
DSTRUC6	.05	-.01	-.04	-.15	-.03	-.01	-.05	.08
SEX	-.09	-.11	-.09	-.03	.19	.20	-.27	-.29
ZMARD	.07	.06	.23	.14	.22	.16	.02	-.07
ZPARD	.07	.05	-.00	-.02	.21	.06	.05	-.17
MODVOC	.07	.06	.33	.23	.27	.12	.25	.06
CONSUMP	.15	.17	.26	.19	.03	-.05	.06	.04
PASTAT	.15	.16	-.09	-.05	.02	-.05	.13	.11
ZMAOCC	-.15	-.15	-.07	-.05	.23	-.00	-.20	-.16
EDASP	.23	.30	.30	.24	.15	.09	.08	-.04
D6	-.06	-.28	.17	.09	-.06	-.21	.06	.35
D8	-.02	-.06	-.36	-.25	.35	.08	-.06	-.37
D14	-.26	-.08	.09	.09	-.27	-.01	.15	.53
ATTAGE	-.04	-.07	-.21	-.14	-.18	-.07	-.00	.15

II.A.10. Summary

Rather than a discursive reiteration of the already brief text of the preceding sections, this summary consists of a series of substantive propositions which are organized into three groups: (1) findings reported with a high degree of confidence, (2) findings reported with a lesser confidence, and (3) anomalous findings.

The classification here is based not only on levels of statistical significance of individual regression coefficients but also on a combination of apparent robustness, evidenced by internal consistency across models, interpretability, and, to a lesser extent, congruence with findings in other research.

A few statistically significant coefficients which appear in various tables in this part of the report are not discussed in either the text or the summary. These are primarily coefficients of social and economic indicators which were included as controls but were not themselves of central interest. Because they were included eclectically and, to some extent, redundantly, the magnitudes of the partial regression coefficients vary from table to table and hence fail on the criterion of robustness. As indicated in various places in this report, however, the focus of this section is primarily on the biological rather than social backgrounds of achievement.

I. Findings in Section A reported with high confidence:

1) The availability of free high-protein supplementation in the Atole villages led to considerable replacement of home diet by supplementation and a consequent increase in the proportion of protein in the diets of children in those villages.

2) While children's anthropometry by age before the project were similar, after the introduction of supplementation, growth in weight and stature was greater in the Atole villages.
3) Differences between Atole and Fresco villages in the growth of children in both weight and stature was greater for girls than for boys.
4) Within the Atole villages, children who consumed more supplement before age 36 months gained more in both weight and stature; this was not true in the Fresco villages.
5) Incidence of diarrhea has a consistently negative effect upon physical growth -- particularly stature.
6) Children are more likely to consume supplement if their mothers attend the supplementation centers.
7) Children from more affluent homes (CONSUMP scale) are less likely to consume supplementation.
8) Boys eat more at home than girls, even allowing for initial weight differences.
9) Parental literacy is associated with the provision of better home diets.
10) Taller children perform better on verbal performance tests.
11) Children from more affluent homes (CONSUMP scale) perform better on verbal tests.
12) Girls perform better than boys on verbal tests.
13) Children with high verbal proficiency are more likely to enroll in school.
14) Children with high verbal proficiency are more likely to perform better in school.

15) Verbal performance, stature, weight, home diet, and supplementation show high stability over time. They show "simplex" patterns of intercorrelations.

16) Concurrent nutritional intakes affect the school performance of children of similar stature.

II. Findings reported with somewhat less confidence:

17) The sex-treatment interaction noted in 3, above, is due to the relatively more ample home diets of girls, at early ages, in the Atole villages, which they enjoy because of the presence of a free nutritional supplement which, in turn, reduces competition for food at home.

18) The effect of the Atole supplement upon physical growth and lack of effect of Fresco supplement upon physical growth, noted in 4, above, is due to the high protein content of the Atole.

19) Parental literacy, particularly maternal literacy, favorably affects verbal development.

20) Verbal performance is enhanced by the presence of more elder persons in the family; it is depressed by the presence of younger siblings in the family.

21) The incidence of morbidity (other than diarrhea) depresses verbal performance.

22) Families in which children have died provide a poorer home diet to surviving children.

III. Anomalous finding:

23) The negative coefficients of the regressions of growth in weight and stature upon caloric consumption, holding protein constant, do not have any clear biological interpretation. There are several possible misspecifications of the models in Section II.A.6 where these results were found. There, the lack of information on intake of nursing infants, including their home diet, may have been the cause of this anomalous result. This unexplained coefficient reduces confidence in the estimate of the independent effect of protein, reported as 17, above.

Appendix A.11.a

A Note on Path Models and Their Estimation

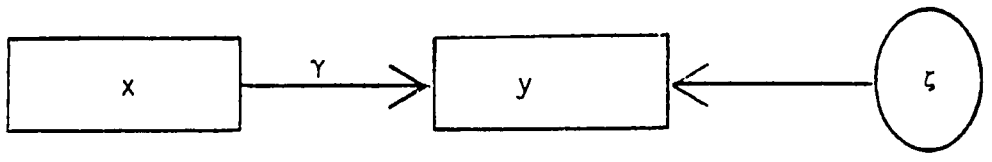
A common simple regression model could be formulated as

$$y = \gamma x + \zeta,$$

$$\text{assuming } E(x) = E(x\zeta) = 0.$$

This regression model could be sketched as a path diagram using rectangles to contain the observed variables, x and y, and an oval for the "shocks" or unexplained variance in the equation.

Figure A.21



Multiplying the equation by x and taking expectations, we find:

$$\sigma_{xy} = \gamma\sigma_x^2 = E(x\zeta),$$

$$\gamma = \sigma_{xy}/\sigma_x^2 .$$

Multiplying by y we have:

$$\sigma_y^2 = \gamma\sigma_{xy} + E(y\zeta),$$

$$\sigma_\zeta^2 = \sigma_y^2 - \gamma\sigma_{xy} .$$

While the foregoing traditional least squares estimates of the conditional expectation of y given x provide unbiased prediction, they do not necessarily correspond to the causal mechanisms which generated the data. This "least squares bias" in estimating structural relations is particularly clear where there is substantial error of measurement.

In many applications, it is presumed that measurement error is negligible -- at least when compared with the unexplained variance, σ_{ζ}^{2}. This stance is not tenable, however, when dealing with psychometric variables, indices of attitude, most self-reported recall information, or small samples of behavior. In the present study, the measurement of home diet, based upon a mother's report of the child's consumption for three days, can be presumed to be an unreliable estimator of the child's average consumption for a year.

Where the two observed variables, x and y, are taken to be unreliable indicators of latent or "true" unobserved variables, say ξ and η , respectively, the model can then be specified

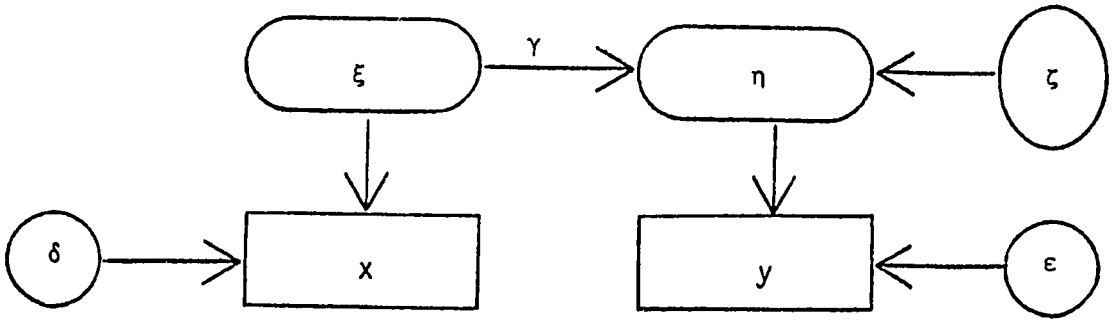
$$x = \xi + \delta,$$

$$y = \eta + \varepsilon,$$

$$\eta = \gamma\xi + \zeta,$$

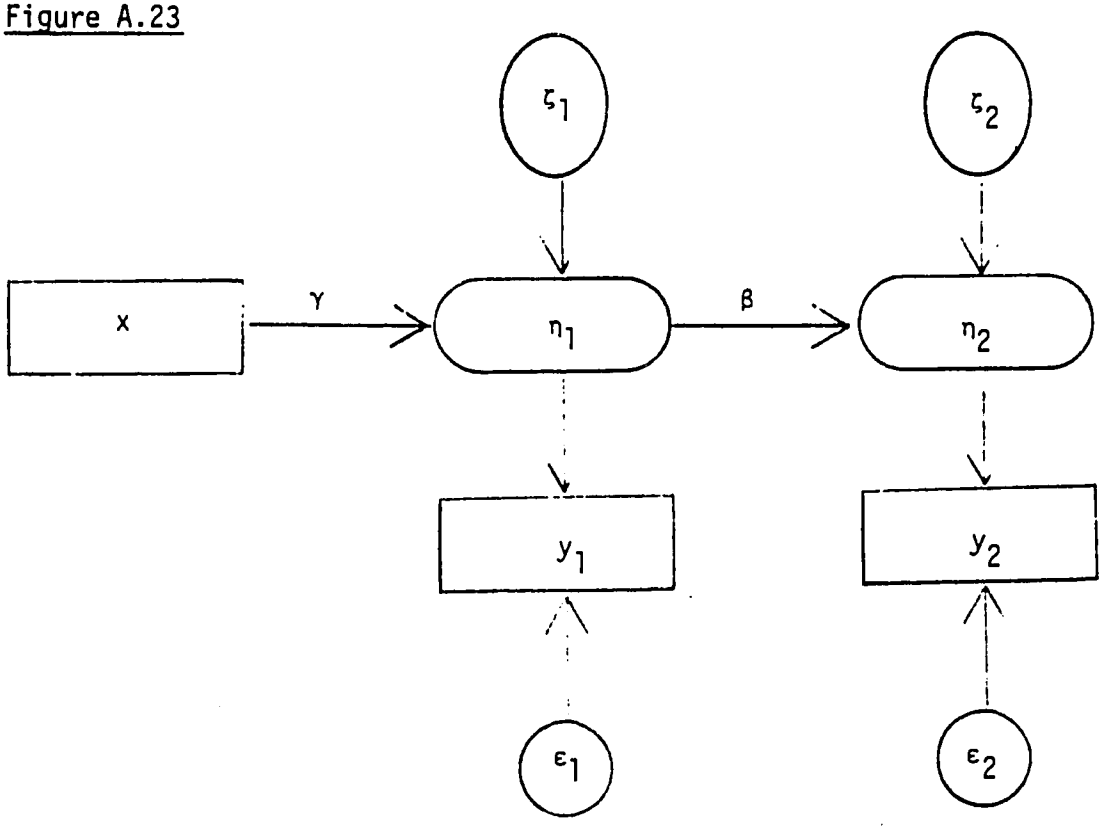
where δ and ε are errors of measurement. Using ovals to contain the unobserved variables, this can be sketched

Figure A.22



Three kinds of additional information would permit an estimate of the structural coefficient, γ : (1) an independent estimate of the unreliability of x, σ_δ^2; (2) an exogenous variable, measured without error, which has a substantial effect on x; or (3) multiple fallible indicators of the "true" unobserved variables.

Figure A.23



In the causal chain portrayed above, we have introduced the additional variable, x, which is presumably measured without error. Our measurement model equations are:

$$y_1 = \eta_1 + \varepsilon_1, \text{ and}$$
$$y_2 = \eta_2 + \varepsilon_2 .$$

Substituting the two measurement equations into the structural equation, we have

$$(y - \varepsilon) = \gamma(x - \delta) + \zeta .$$

Again, multiplying through by x and taking expectations, we get

$$\gamma_{xy} - E(x\varepsilon) = \gamma(\sigma_x^2 - E(x\delta)) + E(x\zeta) .$$

Assuming

$$E(x\varepsilon) = E(x\zeta) = 0,$$

then

$$\sigma_{xy} = \gamma(\sigma_x^2 - \sigma_\delta^2)$$

and

$$\gamma = \sigma_{xy}/(\sigma_x^2 - \sigma_\delta^2) .$$

We can see that γ will be larger than the ordinary least squres regression coefficient, σ_{xy}/σ_x^2, depending on the degree of unreliability of the independent variable, x. However, the value of γ is not identifiable without further information.

Note also that the structural coefficient is not biased by the error of measurement of the dependent variable, σ_ε^2. If we know σ_δ^2 and could estimate γ, we would still not have enough information to distinguish σ_ε^2 from σ_ζ^2. Multiplying the structural equation through by y, we have

$$\sigma_y^2 - E(y\varepsilon) = \gamma(\sigma_{xy} - E(y\delta)) + E(y\zeta) .$$

Hence

$$(\sigma_\varepsilon^2 + \sigma_\zeta^2) = \sigma_y^2 - \gamma\sigma_{xy} .$$

The sum of these two error components, but not each separately, would be identified.

The structural equations are

$$\eta_1 = \gamma x + \zeta_1, \text{ and}$$

$$\eta_2 = \beta\eta_1 + \zeta_2 .$$

We have

$$\sigma_{xy_1} = \sigma_{x\eta_1} = \gamma\sigma_x^2, \text{ and}$$

$$\gamma = \sigma_{xy_1}/\sigma_x^2 ,$$

the ordinary least squares regression. Substituting the measurement equations into the second structural equation, we get

$$(y_2 - \varepsilon_2) = \beta(y_1 - \varepsilon_1) + \zeta_2 .$$

Multiplying by x and taking expectations

$$\sigma_{xy_2} - E(x\varepsilon_2) = \beta(\sigma_{xy_1} - E(x\varepsilon_1)) + E(x\zeta_2)$$

$$\sigma_{xy_2} = \beta\sigma_{xy_1}$$

$$\beta = \sigma_{xy_2}/\sigma_{xy_1}$$

Multiplying by y_1 and taking expectations

$$\sigma_{y_1y_2} - E(y_1\varepsilon_2) = \beta(\sigma_{y_1}^2 - E(y_1\varepsilon_1))$$

$$\sigma_{y_1y_2} = \beta\sigma_{y_1}^2 - \beta\sigma_{\varepsilon_1}^2$$

$$\sigma_{\varepsilon_1}^2 = \sigma_{y_1}^2 - \sigma_{y_1y_2}/\beta$$

$$\sigma_{\eta_1}^2 = \sigma_{y_1}^2 - \sigma_{\varepsilon_1}^2$$

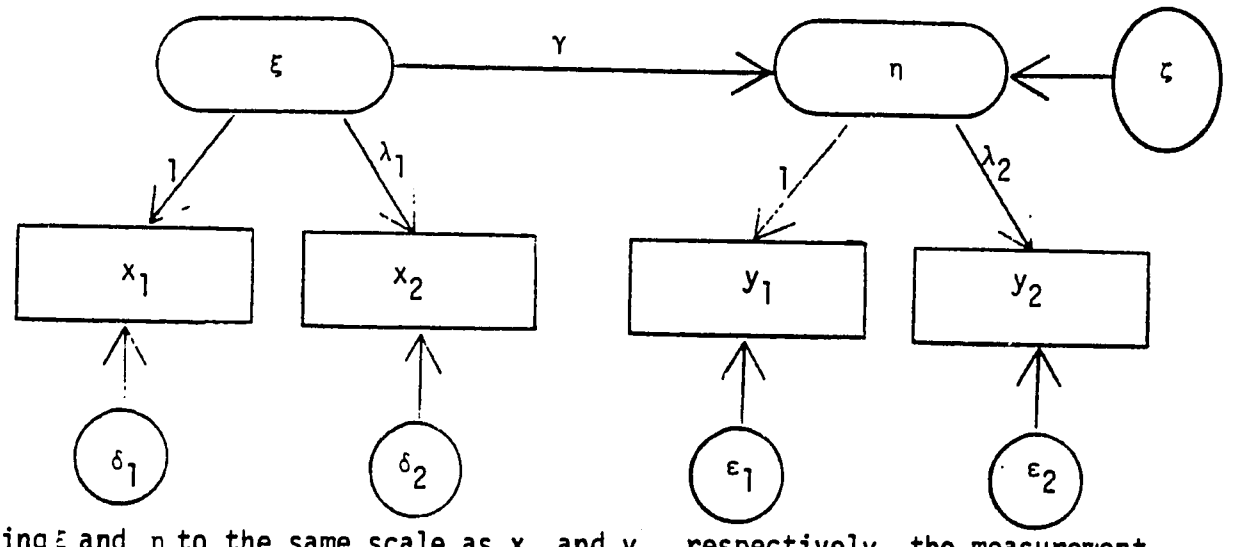
$$\sigma^2_{\zeta_1} = \sigma^2_{\eta_1} - \gamma\sigma^2_x$$

$$(\sigma^2_{\zeta_2} + \sigma^2_{\varepsilon_2}) = \sigma^2_{y_2} - \beta\sigma^2_{\eta_1}$$

Thus, using the instrumental variable, x, we can identify γ, β, ε_1, and ζ_1, but only the sum $(\sigma^2_{\zeta_2} + \sigma^2_{\varepsilon_2})$. Since estimating the structural coefficients, γ and β, is the likely primary object, setting ε'_2 to zero and estimating $\zeta'_2 = \zeta_2 + \varepsilon_2$ would suffice.

With multiple indicators of each latent variable, all of the unknown parameters can be estimated.

Figure A.24



Setting ξ and η to the same scale as x_1 and y_1, respectively, the measurement model equations are:

$$y_1 = \eta + \varepsilon_1$$

$$y_2 = \lambda_2\eta + \varepsilon_2$$

$$x_1 = \xi + \delta_1$$

$$x_2 = \lambda_1\xi + \delta_2$$

The structural equation is

$$\eta = \gamma\xi + \zeta \ .$$

The four observed variables provide 10 variances and covariances to estimate 9 parameters (λ_1, λ_2, γ, σ_ξ^2, σ_ζ^2, δ_1, δ_2, ε_1, and ε_2). We have the covariances

$$\sigma_{x_1y_1} = \sigma_{x_1\eta} = \gamma\sigma_\xi^2$$

$$\sigma_{x_1y_2} = \lambda_2\sigma_{x_1\eta} = \lambda_2\gamma\sigma_\xi^2$$

$$\sigma_{x_2y_1} = \sigma_{x_2\eta} = \lambda_1\gamma\sigma_\xi^2$$

$$\sigma_{x_2y_2} = \lambda_2\sigma_{x_2\eta} = \lambda_1\lambda_2\gamma\sigma_\xi^2$$

Dividing the third and fourth equations by the first two, we get

$$\lambda_1 = \sigma_{x_2y_1}/\sigma_{y_1x_1} = \sigma_{y_2x_2}/\sigma_{y_2x_1}$$

which solves for λ_1 and provides an overidentifying constraint. With λ_1 determined σ_ξ^2 is determined by $\sigma_{x_1x_2} = \lambda_1\sigma_\xi^2$

$$\gamma = \sigma_{x_1y_1}/\sigma_\xi^2$$

$$\lambda_2 = \sigma_{x_1y_2}/\gamma\sigma_\xi^2$$

$$\sigma_\zeta^2 = \sigma_{y_1y_2}/\lambda_2 - \gamma^2\sigma_\xi^2$$

$$\sigma_{\delta_1}^2 = \sigma_{x_1}^2 - \sigma_\xi^2$$

$$\sigma_{\delta_2}^2 = \sigma_{x_2}^2 - \lambda_1\sigma_\xi^2$$

$$\sigma_{\varepsilon_1}^2 = \sigma_{y_1}^2 - \sigma_\eta^2 \qquad (\sigma_\eta^2 = \gamma\sigma_\xi^2 + \sigma_\zeta^2)$$

$$\sigma_{\varepsilon_2}^2 = \sigma_{y_2}^2 - \lambda_2\sigma_\eta^2$$

The simple two or three variable models which are discussed above have been generalized by Jöreskog to the Linear Structural Relations (LISREL)

model. This model includes a structural equation expressing the relations among the latent variables and equations relating the latent variables to observed variables. This general model and its estimation are discussed succinctly in Jöreskog and Sörbom (1979; pp. 106-13).

The computer program, LISREL III, by Jöreskog and Sörbom (1976), which has been slightly modified so as to constrain variances to be non-negative, has been used throughout this section. This provides maximum likelihood estimates of the parameters of the sets of simultaneous equations using an iterative Fletcher-Powell (1963) algorithm for minimization.

While the conditions for identifiability have been worked out for a number of specific models involving latent variables (Goldberger and Duncan (eds.) 1973, and Aigner and Goldberger (eds.), 1977, passim), they had to be worked out on a case by case basis, and sometimes by inductive trial and error, for the complex models developed in this report.

REFERENCES

Aigner, D. J., and Goldberger, A. S., eds., Latent Variables in Socio-Economic Models. Amsterdam: North-Holland, 1977.

Duncan, Otis Dudley. Introduction to Structural Equation Models. New York: Academic, 1975.

Fletcher, R., and Powell, M. J. D., "A Rapidly Convergent Descent Method for Minimization." Computer Journal 6 (1963): 163-168.

Goldberger, Arthur S., and Duncan, Otis Dudley, eds., Structural Equation Models in the Social Sciences. New York: Seminar, 1973.

Jöreskog, Karl G., and Sörbom, Dag. Advances in Factor Analysis and Structural Equation Models. Cambridge: ABT Books, 1979.

Appendix A.11.b

List A.2 (Section A.6, Figure A.15)
VARIABLES IN CALORIES, PROTEIN AND GROWTH MODEL
Age 24-60
All Villages Combined

Symbols

Observed Variables	Constructs	Acronym	Age	Variable Description
y_1	η_1	WT12	36	Weight in kilograms
y_2	η_2	WT14	48	Weight in kilograms
y_3	η_3	WT15	60	Weight in kilograms
y_4	η_4	HT12	36	Height in centimeters
y_5	η_5	HT14	48	Height in centimeters
y_6	η_6	HTC15	60	Height in centimeters
	η_7		30	Calories: $\xi_1 + \xi_2$
	η_8		42	Calories: $\xi_3 + \xi_4$
	η_9		54	Calories: $\xi_5 + \xi_6$
	η_{10}		30	Protein: $\xi_7 + \xi_8$
	η_{11}		42	Protein: $\xi_9 + \xi_{10}$
	η_{12}		54	Protein: $\xi_{11} + \xi_{12}$
x_1	ξ_1	BCHHDC2	30	Home diet calories
x_2	ξ_2	CHSPC18	30	Supplementation calories
x_3	ξ_3	ACHHDC3	42	Home diet calories
x_4	ξ_4	CHSPC19	42	Supplementation calories
x_5	ξ_5	BCHHDC4	54	Home diet calories
x_6	ξ_6	CHSPC13	54	Supplementation calories
x_7	ξ_7	BCHHDP2	30	Home diet protein
x_8	ξ_8	CHSPP18	30	Supplementation protein
x_9	ξ_9	ACHHDP3	42	Home diet protein
x_{10}	ξ_{10}	CHSPP19	42	Supplementation protein

List A.2 continued

Symbols				
Observed Variables	Constructs	Acronym	Age	Variable Description
x_{11}	ξ_{11}	BCHHDP4	54	Home diet protein
x_{12}	ξ_{12}	CHSPP13	54	Supplementation protien
x_{13}	ξ_{13}	SEX		Sex
x_{14}	ξ_{14}	MAHT		Mother's height in centimeters
x_{15}	ξ_{15}	PAHT		Father's height in centimeters
x_{16}	ξ_{16}	CHDIAR20	24-36	Child's incidence of diarrhea: proportion
x_{17}	ξ_{17}	CHDIAR21	36-48	Child's incidence of diarrhea: proportion
x_{18}	ξ_{18}	CHDIAR14	48-60	Child's incidence of diarrhea: proportion
x_{19}	ξ_{19}	MORBOTH2	24-36	Child's incidence of ailments other than diarrhea
x_{20}	ξ_{20}	MORBOTH3	36-48	Child's incidence of ailments other than diarrhea
x_{21}	ξ_{21}	MORBOTH4	48-60	Child's incidence of ailments other than diarrhea
x_{22}	ξ_{22}	WT10	24	Weight in kilograms
x_{23}	ξ_{23}	HT10	24	Height in kilograms

Equation A.2 (Section II.A.6, Figure A.15)

Calories, Protein, and Growth -- Measurement Model Equations

$$\begin{vmatrix} y_1 \\ y_2 \\ y_3 \\ y_4 \\ y_5 \\ y_6 \end{vmatrix} = \begin{vmatrix} 1 & & & & & \\ & 1 & & & & \\ & & 1 & & & \\ & & & 1 & & \\ & & & & 1 & \\ & & & & & 1 \end{vmatrix} * \begin{vmatrix} \eta_1 \\ \eta_2 \\ \eta_3 \\ \eta_4 \\ \eta_5 \\ \eta_6 \\ \eta_7 \\ \eta_8 \\ \eta_9 \\ \eta_{10} \\ \eta_{11} \\ \eta_{12} \end{vmatrix} + \begin{vmatrix} \epsilon_1 \\ \epsilon_2 \\ \epsilon_3 \\ 0 \\ 0 \\ 0 \end{vmatrix}$$

$$\begin{vmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \\ x_5 \\ x_6 \\ x_7 \\ x_8 \\ x_9 \\ x_{10} \\ x_{11} \\ x_{12} \\ \vdots \\ x_{21} \\ x_{22} \\ x_{23} \end{vmatrix} = \begin{vmatrix} 1 & & & & & & & & & & & & & & & & & & & & & & & \\ & 1 & & & & & & & & & & & & & & & & & & & & & & \\ & & 1 & & & & & & & & & & & & & & & & & & & & & \\ & & & 1 & & & & & & & & & & & & & & & & & & & & \\ & & & & 1 & & & & & & & & & & & & & & & & & & & \\ & & & & & 1 & & & & & & & & & & & & & & & & & & \\ & & & & & & 1 & & & & & & & & & & & & & & & & & \\ & & & & & & & 1 & & & & & & & & & & & & & & & & \\ & & & & & & & & 1 & & & & & & & & & & & & & & & \\ & & & & & & & & & 1 & & & & & & & & & & & & & & \\ & & & & & & & & & & 1 & & & & & & & & & & & & & \\ & & & & & & & & & & & 1 & & & & & & & & & & & & \\ & & & & & & & & & & & & \ddots & & & & & & & & & & & \\ & & & & & & & & & & & & & 1 & & & & & & & & & & \\ & & & & & & & & & & & & & & 1 & & & & & & & & & \\ & & & & & & & & & & & & & & & 1 & & & & & & & & \end{vmatrix} * \begin{vmatrix} \xi_1 \\ \xi_2 \\ \xi_3 \\ \xi_4 \\ \xi_5 \\ \xi_6 \\ \xi_7 \\ \xi_8 \\ \xi_9 \\ \xi_{10} \\ \xi_{11} \\ \xi_{12} \\ \vdots \\ \xi_{21} \\ \xi_{22} \\ \xi_{23} \end{vmatrix} + \begin{vmatrix} \delta_1 \\ 0 \\ \delta_3 \\ 0 \\ \delta_5 \\ 0 \\ \delta_7 \\ 0 \\ \delta_9 \\ 0 \\ \delta_{11} \\ 0 \\ \vdots \\ 0 \\ \delta_{21} \\ 0 \end{vmatrix}$$

Equation A.3 (Section A.6, Figure A.15)

Calories, Protein, and Growth -- Structural Equation

$$
\begin{vmatrix}
1 & & & & & & & & & & & \\
\beta & 1 & & & & & & & & & & \\
 & \beta & 1 & & & & & & & & & \\
 & & & 1 & & & & & & & & \\
 & & & \beta & 1 & & & & & & & \\
 & & & & \beta & 1 & & & & & & \\
 & & & & & \beta & 1 & & & & & \\
 & & & & & & & 1 & & & & \\
 & & & & & & & & 1 & & & \\
 & & & & & & & & & 1 & & \\
 & & & & & & & & & & 1 & \\
 & & & & & & & & & & & 1
\end{vmatrix}
+
\begin{vmatrix}
\eta_1 \\ \eta_2 \\ \eta_3 \\ \eta_4 \\ \eta_5 \\ \eta_6 \\ \eta_7 \\ \eta_8 \\ \eta_9 \\ \eta_{10} \\ \eta_{11} \\ \eta_{12}
\end{vmatrix}
=
\begin{vmatrix}
1 & 1 & & & & & & & & & & \\
 & & 1 & 1 & & & & & & & & \\
 & & & & 1 & 1 & & & & & & \\
 & & & & & & 1 & 1 & & & & \\
 & & & & & & & & 1 & 1 & & \\
 & & & & & & & & & & 1 & 1
\end{vmatrix}
\begin{vmatrix}
\gamma & & \gamma & & \gamma & & \gamma \\
1 & & & \gamma & & \gamma & \\
1 & & & & \gamma & & \gamma \\
1 & \gamma & \gamma & \gamma & & \gamma & & \gamma \\
\gamma & \gamma & \gamma & & \gamma & & \gamma & \\
\gamma & \gamma & \gamma & & & \gamma & & \gamma
\end{vmatrix}
*
\begin{vmatrix}
\xi_1 \\ \xi_2 \\ \xi_3 \\ \xi_4 \\ \xi_5 \\ \xi_6 \\ \xi_7 \\ \xi_8 \\ \xi_9 \\ \xi_{10} \\ \xi_{11} \\ \xi_{12} \\ \xi_{13} \\ \xi_{14} \\ \xi_{15} \\ \xi_{16} \\ \xi_{17} \\ \xi_{18} \\ \xi_{19} \\ \xi_{20} \\ \xi_{21} \\ \xi_{22} \\ \xi_{23}
\end{vmatrix}
+
\begin{vmatrix}
\delta_1 \\ \delta_2 \\ \delta_3 \\ \delta_4 \\ \delta_5 \\ \delta_6 \\ 0 \\ 0 \\ 0 \\ 0 \\ 0 \\ 0
\end{vmatrix}
$$

Equation A.4 (Section A.6, Figure A.16)

Growth in Stature -- Cross Sectional Model -- Equations

$$\begin{vmatrix} 1 & \beta & \beta \\ & 1 & \\ & & 1 \end{vmatrix} * \begin{vmatrix} \eta_1 \\ \eta_2 \\ \eta_3 \end{vmatrix} = \begin{vmatrix} \gamma & \gamma & \gamma & & & \gamma & \gamma & & \\ & & & 1 & 1 & & & & \\ & & & & & & & 1 & 1 \end{vmatrix} * \begin{vmatrix} \xi_1 \\ \xi_2 \\ \xi_3 \\ \xi_4 \\ \xi_5 \\ \xi_6 \\ \xi_7 \\ \xi_8 \\ \xi_9 \end{vmatrix} + \begin{vmatrix} \zeta_1 \\ 0 \\ 0 \end{vmatrix}$$

$$\begin{vmatrix} y_1 \end{vmatrix} = \begin{vmatrix} 1 & 0 & 0 \end{vmatrix} * \begin{vmatrix} \eta_1 \\ \eta_2 \\ \eta_3 \end{vmatrix} + \begin{vmatrix} 0 \end{vmatrix}$$

$$\begin{vmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \\ x_5 \\ x_6 \\ x_7 \\ x_8 \\ x_9 \end{vmatrix} = \begin{vmatrix} 1 & & & & & & & & \\ & 1 & & & & & & & \\ & & 1 & & & & & & \\ & & & 1 & & & & & \\ & & & & 1 & & & & \\ & & & & & 1 & & & \\ & & & & & & 1 & & \\ & & & & & & & 1 & \\ & & & & & & & & 1 \end{vmatrix} * \begin{vmatrix} \xi_1 \\ \xi_2 \\ \xi_3 \\ \xi_4 \\ \xi_5 \\ \xi_6 \\ \xi_7 \\ \xi_8 \\ \xi_9 \end{vmatrix} + \begin{vmatrix} 0 \\ 0 \\ 0 \\ \delta_4 \\ 0 \\ 0 \\ 0 \\ \delta_8 \\ 0 \end{vmatrix}$$

Equation A.5 (Section A.7, Figure A.17)

Verbal Development - Structural Equation

$$
\begin{array}{c|ccccc|}
 & 1 & 2 & 3 & 4 & 5 \\
\hline
1 & 1 & & & & \\
2 & \beta & 1 & & & \\
3 & & \beta & 1 & & \\
4 & & & \beta & 1 & \\
5 & & & & \beta & 1
\end{array}
\;*\;
\begin{vmatrix} \eta_1 \\ \eta_2 \\ \eta_3 \\ \eta_4 \\ \eta_5 \end{vmatrix}
=
\begin{array}{c|ccccccccccccc|}
 & 1 & 2 & 3 & 4 & 5 & 6 & 7 & 8 & 9 & 10 & 11 & 12 & 13 \\
\hline
 & \gamma & \gamma & \gamma & \gamma & \gamma & \gamma & \gamma & \gamma & \gamma & \gamma & \gamma & \gamma & \gamma \\
 & \gamma & \gamma & \gamma & \gamma & \gamma & \gamma & \gamma & \gamma & \gamma & \gamma & \gamma & \gamma & \gamma \\
 & \gamma & \gamma & \gamma & \gamma & \gamma & \gamma & \gamma & \gamma & \gamma & \gamma & \gamma & \gamma & \gamma \\
 & \gamma & \gamma & \gamma & \gamma & \gamma & \gamma & \gamma & \gamma & \gamma & \gamma & \gamma & \gamma & \gamma \\
 & \gamma & \gamma & \gamma & \gamma & \gamma & \gamma & \gamma & \gamma & \gamma & \gamma & \gamma & \gamma & \gamma
\end{array}
*
\begin{vmatrix} \xi_1 \\ \xi_2 \\ \xi_3 \\ \xi_4 \\ \xi_5 \\ \xi_6 \\ \xi_7 \\ \xi_8 \\ \xi_9 \\ \xi_{10} \\ \xi_{11} \\ \xi_{12} \\ \xi_{13} \end{vmatrix}
+
\begin{vmatrix} \zeta_1 \\ \zeta_2 \\ \zeta_3 \\ \zeta_4 \\ \zeta_5 \end{vmatrix}
$$

Equation A.6 (Section A.7, Figure A.17)

Verbal Development -- Measurement Model Equations

$$\begin{vmatrix} y_1 \\ y_2 \\ y_3 \\ y_4 \\ y_5 \\ y_6 \\ y_7 \\ y_8 \\ y_9 \\ y_{10} \end{vmatrix} = \begin{vmatrix} 1 & & & & & & & & & \\ \lambda & & & & & & & & & \\ & 1 & & & & & & & & \\ & \lambda & & & & & & & & \\ & & 1 & & & & & & & \\ & & \lambda & & & & & & & \\ & & & 1 & & & & & & \\ & & & \lambda & & & & & & \\ & & & & 1 & & & & & \\ & & & & \lambda & & & & & \end{vmatrix} * \begin{vmatrix} \eta_1 \\ \eta_2 \\ \eta_3 \\ \eta_4 \\ \eta_5 \end{vmatrix} + \begin{vmatrix} \epsilon_1 \\ \epsilon_2 \\ \epsilon_3 \\ \epsilon_4 \\ \epsilon_5 \\ \epsilon_6 \\ \epsilon_7 \\ \epsilon_8 \\ \epsilon_9 \\ \epsilon_{10} \end{vmatrix}$$

$$\Theta_\epsilon = \begin{vmatrix} \sigma^2_{\epsilon_1} & & & & & & & & & \\ \sigma & \sigma^2_{\epsilon_2} & & & & & & & & \\ \sigma & & \sigma^2_{\epsilon_3} & & & & & & & \\ & \sigma & \sigma & \sigma^2_{\epsilon_4} & & & & & & \\ \sigma & & \sigma & & \sigma^2_{\epsilon_5} & & & & & \\ & \sigma & & \sigma & \sigma & \sigma^2_{\epsilon_6} & & & & \\ \sigma & & \sigma & & \sigma & & \sigma^2_{\epsilon_7} & & & \\ & \sigma & & \sigma & & \sigma & \sigma & \sigma^2_{\epsilon_8} & & \\ \sigma & & \sigma & & \sigma & & \sigma & & \sigma^2_{\epsilon_9} & \\ & \sigma & & \sigma & & \sigma & & \sigma & \sigma & \sigma^2_{\epsilon_{10}} \end{vmatrix}$$

$$\sigma_{\epsilon_1,\epsilon_2} = \sigma_{\epsilon_3,\epsilon_4} = \sigma_{\epsilon_5,\epsilon_6} = \sigma_{\epsilon_7,\epsilon_8} = \sigma_{\epsilon_9,\epsilon_{10}}$$

$$\begin{vmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \\ x_5 \\ x_6 \\ x_7 \\ x_8 \\ x_9 \\ x_{10} \\ x_{11} \\ x_{12} \\ x_{13} \end{vmatrix} = \begin{vmatrix} 1 & & & & & & & & & & & & \\ & 1 & & & & & & & & & & & \\ & & 1 & & & & & & & & & & \\ & & & 1 & & & & & & & & & \\ & & & & 1 & & & & & & & & \\ & & & & & 1 & & & & & & & \\ & & & & & & 1 & & & & & & \\ & & & & & & & 1 & & & & & \\ & & & & & & & & 1 & & & & \\ & & & & & & & & & 1 & & & \\ & & & & & & & & & & 1 & & \\ & & & & & & & & & & & 1 & \\ & & & & & & & & & & & & 1 \end{vmatrix} * \begin{vmatrix} \xi_1 \\ \xi_2 \\ \xi_3 \\ \xi_4 \\ \xi_5 \\ \xi_6 \\ \xi_7 \\ \xi_8 \\ \xi_9 \\ \xi_{10} \\ \xi_{11} \\ \xi_{12} \\ \xi_{13} \end{vmatrix} + \begin{vmatrix} 0 \\ 0 \\ 0 \\ 0 \\ 0 \\ 0 \\ 0 \\ 0 \\ 0 \\ 0 \\ 0 \\ 0 \\ 0 \end{vmatrix}$$

I.e., $x_i = \xi_i \quad (i = 1, 13)$

List A.3 (Section A.7, Table A.36)

STATURE, DIET, AND VERBAL DEVELOPMENT VARIABLES

Symbol: Observed Variables	Symbol: Constructs	Acronym	Age	Variable Description
y_1		NAMING1	36	Naming psychometric test
y_2		RECOG1	36	Recognition psychometric test
	η_1		36	Verbal factor
y_3		NAMING2	48	Naming psychometric test
y_4		RECOG2	48	Recognition psychometric test
	η_2		48	Verbal factor
y_5		NAMING3	60	Naming psychometric test
y_6		RECOG3	60	Recognition psychometric test
	η_3		60	Verbal factor
x_1	ξ_1	BCHHDC2	30	Home diet calores
x_2	ξ_2	CHSPC18	30	Supplementation calories
	η_4		30	Calories: $\xi_1 + \xi_2$
x_3	ξ_3	ACHHDC3	42	Home diet calories
x_4	ξ_4	CHSPC19	42	Supplementation calories
	η_5		42	Calories: $\xi_3 + \xi_4$
x_5	ξ_5	BCHHDC4	54	Home diet calories
x_6	ξ_6	CHSPC13	54	Supplementation calories
	η_6		54	Calories: $\xi_5 + \xi_6$
x_7	ξ_7	BCHHDP2	30	Home diet protein
x_8	ξ_8	CHSPP18	30	Supplementation protein
	η_9		30	Protein: $\xi_7 + \xi_8$
x_9	ξ_9	ACHHDP3	42	Home diet protein
x_{10}	ξ_{10}	CHSPP19	42	Supplementation protein
	η_8		42	Protein: $\xi_9 + \xi_{10}$

List A.3 continued

Symbol: Observed Variables	Symbol: Constructs	Acronym	Age	Variable Description
x_{11}	ξ_{11}	BCHHDP7	54	Home diet protein
x_{12}	ξ_{12}	CHSPP13	54	Supplementation protein
	η_9		54	Protein: $\xi_{11} + \xi_{12}$
x_{13}	ξ_{13}	SEX		Sex
x_{14}	ξ_{14}	ZMARD		Mother's literacy and schooling
x_{15}	ξ_{15}	MODVOC		Mother's modernity and vocabulary
x_{16}	ξ_{16}	CONSUMP		Housing, possessions, and salary
x_{17}	ξ_{17}	FMSZO	0	Number of household members at birth
x_{18}	ξ_{18}	CHDIAR20	24-36	Child's incidence of diarrhea: proportion
x_{19}	ξ_{19}	CHDIAR21	36-48	Child's incidence of diarrhea: proportion
x_{20}	ξ_{20}	CHDIAR14	48-60	Child's incidence of diarrhea: proportion
x_{21}	ξ_{21}	MORBOTH2	24-36	Child's incidence of ailments other than diarrhea
x_{22}	ξ_{22}	MORBOTH3	36-48	Child's incidence of ailments other than diarrhea
x_{23}	ξ_{23}	MORBOTH4	48-60	Child's incidence of ailments other than diarrhea
x_{24}	ξ_{24}	DSTRUC3	36	Family structure: nuclear vs. all other
x_{25}	ξ_{25}	HT10	24	Height in centimeters

Equation A.7 (Section A.7, Table A.36)

Stature, Diet, and Verbal Development -- Ages 36, 48, 60: Measurement Equations

$$
\begin{vmatrix} y_1 \\ y_2 \\ y_3 \\ y_4 \\ y_5 \\ y_6 \end{vmatrix}
=
\begin{array}{c}
\begin{array}{ccccccccc} 1 & 2 & 3 & 4 & 5 & 6 & 7 & 8 & 9 \end{array} \\
\begin{vmatrix}
1 & & & & & & & & \\
\lambda & & & & & & & & \\
 & 1 & & & & & & & \\
 & \lambda & & & & & & & \\
 & & 1 & & & & & & \\
 & & \lambda & & & & & & \\
 & & & & & & & & \\
 & & & & & & & & \\
 & & & & & & & & \\
\end{vmatrix}
\end{array}
\begin{vmatrix} \eta_1 \\ \eta_2 \\ \eta_3 \\ \eta_4 \\ \eta_5 \\ \eta_6 \\ \eta_7 \\ \eta_8 \\ \eta_9 \end{vmatrix}
+
\begin{vmatrix} \epsilon_1 \\ \epsilon_2 \\ \epsilon_3 \\ \epsilon_4 \\ \epsilon_5 \\ \epsilon_6 \end{vmatrix}
$$

$$
\begin{vmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \\ x_5 \\ x_6 \\ x_7 \\ x_8 \\ x_9 \\ x_{10} \\ x_{11} \\ x_{12} \\ \vdots \\ x_{20} \end{vmatrix}
=
\begin{vmatrix}
1 & & & & & & & & & & & & & \\
 & 1 & & & & & & & & & & & & \\
 & & 1 & & & & & & & & & & & \\
 & & & 1 & & & & & & & & & & \\
 & & & & 1 & & & & & & & & & \\
 & & & & & 1 & & & & & & & & \\
 & & & & & & 1 & & & & & & & \\
 & & & & & & & 1 & & & & & & \\
 & & & & & & & & 1 & & & & & \\
 & & & & & & & & & 1 & & & & \\
 & & & & & & & & & & 1 & & & \\
 & & & & & & & & & & & 1 & & \\
 & & & & & & & & & & & & \ddots & \\
 & & & & & & & & & & & & & 1
\end{vmatrix}
\cdot
\begin{vmatrix} \xi_1 \\ \xi_2 \\ \xi_3 \\ \xi_4 \\ \xi_5 \\ \xi_6 \\ \xi_7 \\ \xi_8 \\ \xi_9 \\ \xi_{10} \\ \xi_{11} \\ \xi_{12} \\ \vdots \\ \xi_{15} \end{vmatrix}
+
\begin{vmatrix} \delta_1 \\ 0 \\ \delta_3 \\ 0 \\ \delta_5 \\ 0 \\ \delta_7 \\ 0 \\ \delta_9 \\ 0 \\ \delta_{11} \\ 0 \\ \vdots \\ 0 \end{vmatrix}
$$

Equation A.8 (Section A.7, Table A.36)

Stature, Diet, and Verbal Development -- Ages 36, 48, 60: Structural Equation

$$
\begin{vmatrix}
1 & & & \beta & & & \beta & & \\
\beta & 1 & & & \beta & & & \beta & \\
 & \beta & 1 & & & \beta & & & \beta \\
 & & & 1 & & & & & \\
 & & & & 1 & & & & \\
 & & & & & 1 & & & \\
 & & & & & & 1 & & \\
 & & & & & & & 1 & \\
 & & & & & & & & 1
\end{vmatrix}
*
\begin{vmatrix}
\eta_1 \\ \eta_2 \\ \eta_3 \\ \eta_4 \\ \eta_5 \\ \eta_6 \\ \eta_7 \\ \eta_8 \\ \eta_9
\end{vmatrix}
=
\begin{array}{c}
\begin{array}{ccccccccccccccccccccccccc}
1 & 2 & 3 & 4 & 5 & 6 & 7 & 8 & 9 & 10 & 11 & 12 & 13 & 14 & 15 & 16 & 17 & 18 & 19 & 20 & 21 & 22 & 23 & 24 & 25
\end{array} \\
\begin{vmatrix}
 & & & & & & & & & & & & \gamma & \gamma & \gamma & \gamma & \gamma & \gamma & & & \gamma & & & \gamma & \gamma \\
 & & & & & & & & & & & & \gamma & \gamma & \gamma & \gamma & \gamma & & \gamma & & & \gamma & & \gamma & \gamma \\
 & & & & & & & & & & & & \gamma & \gamma & \gamma & \gamma & \gamma & & & \gamma & & & \gamma & \gamma & \gamma \\
1 & 1 & & & & & & & & & & & & & & & & & & & & & & & \\
 & & 1 & 1 & & & & & & & & & & & & & & & & & & & & & \\
 & & & & 1 & 1 & & & & & & & & & & & & & & & & & & & \\
 & & & & & & 1 & 1 & & & & & & & & & & & & & & & & & \\
 & & & & & & & & 1 & 1 & & & & & & & & & & & & & & & \\
 & & & & & & & & & & 1 & 1 & & & & & & & & & & & & &
\end{vmatrix}
\end{array}
*
\begin{vmatrix}
\xi_1 \\ \xi_2 \\ \xi_3 \\ \xi_4 \\ \xi_5 \\ \xi_6 \\ \xi_7 \\ \xi_8 \\ \xi_9 \\ \xi_{10} \\ \xi_{11} \\ \xi_{12} \\ \xi_{13} \\ \xi_{14} \\ \xi_{15} \\ \xi_{16} \\ \xi_{17} \\ \xi_{18} \\ \xi_{19} \\ \xi_{20} \\ \xi_{21} \\ \xi_{22} \\ \xi_{23} \\ \xi_{24} \\ \xi_{25}
\end{vmatrix}
+
\begin{vmatrix}
\zeta_1 \\ \zeta_2 \\ \zeta_3 \\ 0 \\ 0 \\ 0 \\ 0 \\ 0 \\ 0
\end{vmatrix}
$$

List A.4 (Section A.7, Figure A.18)
VARIABLES IN STATURE, DIET, AND VERBAL FACTOR SCORE CROSS SECTION AT, e.g., 48 MONTHS OF AGE

Observed Variable	Construct	Acronym	Age	Variable Description
y_1		NAMING2	48	Naming psychometric test
y_2		RECOG2	48	Recognition psychometric test
	η_1		48	Verbal factor
x_1	ξ_1	ACHHDC3	42	Home diet calories
x_2	ξ_2	CHSPC19	42	Supplementation calories
	η_2		42	Calories
x_3	ξ_3	ACHHDP3	42	Home diet protein
x_4	ξ_4	CHSPP19	42	Supplementation protein
	η_3		42	Protein
x_5	ξ_5	CHDIAR21	36-48	Child's incidence of diarrhea: proportion
x_6	ξ_6	MORBOTH3	36-48	Child's incidence of ailments other than diarrhea
x_7	ξ_7	AT12	36	Height in centimeters
x_8	ξ_8	FMSZ3	36	Number of household members at age
x_9	ξ_9	YOUNG3	36	Number of younger siblings at age
x_{10}	ξ_{10}	DIED3	36	Number of elder siblings deceased at age
x_{11}	ξ_{11}	DSTRUC3	36	Family structure: nuclear vs. all other
x_{12}	ξ_{12}	SEX		Sex
x_{13}	ξ_{13}	ZMARD		Mother's literacy and schooling
x_{14}	ξ_{14}	ZPARD		Father's literacy and schooling
x_{15}	ξ_{15}	MODVOC		Mother's modernity and vocabulary

List A.4 continued

Observed Variable	Construct	Acronym	Age	Variable Description
x_{16}	ξ_{16}	CONSUMP		Housing, possessions, and salary
x_{17}	ξ_{17}	PASTAT		Father's occupation, ownership of land, animals
x_{18}	ξ_{18}	ZMAOCC		Mother's occupation

Equation A.9 (Section A.7, Figure A.18)

Stature, Diet, and Verbal Factor Scores -- Cross Sectional Structural Equation

$$\begin{vmatrix} 1 & \beta & \beta \\ & 1 & \\ & & 1 \end{vmatrix} * \begin{vmatrix} \eta_1 \\ \eta_2 \\ \eta_3 \end{vmatrix} = \begin{vmatrix} & & & \gamma & \gamma & \gamma & \gamma & \gamma & \gamma & \gamma & \gamma & \gamma & \gamma & \gamma & \gamma & \gamma & \gamma \\ 1 & 1 & & & & & & & & & & & & & & & \\ & & 1 & 1 & & & & & & & & & & & & & \end{vmatrix} * \begin{vmatrix} \xi_1 \\ \xi_2 \\ \xi_3 \\ \xi_4 \\ \xi_5 \\ \xi_6 \\ \xi_7 \\ \xi_8 \\ \xi_9 \\ \xi_{10} \\ \xi_{11} \\ \xi_{12} \\ \xi_{13} \\ \xi_{14} \\ \xi_{15} \\ \xi_{16} \\ \xi_{17} \\ \xi_{18} \end{vmatrix} + \begin{vmatrix} \zeta_1 \\ 0 \\ 0 \end{vmatrix}$$

Equation A.10 (Section A.8, Figure A.19)

$$
\begin{vmatrix} 1 & \alpha & \beta \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{vmatrix} = \begin{vmatrix} 0 & 0 & 0 & 0 & 2 & 2 & 2 & 2 & 2 & 2 & 2 & 2 & 2 & 2 & & 2 & 2 & 2 & 2 & 2 & 2 & 2 & 2 & 2 \\ 0 & 0 & 1 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 1 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \end{vmatrix} + \underline{\underline{=}} + \begin{vmatrix} \delta_1 \\ 0 \\ 0 \end{vmatrix}
$$

$$
\eta_1 = \begin{vmatrix} 1 & 0 & 0 \end{vmatrix} * \begin{vmatrix} n_1 \\ n_2 \\ \eta_3 \end{vmatrix} + \begin{vmatrix} 0 \end{vmatrix} \qquad \text{i.e., } \eta_1 = \eta_1
$$

$$
\begin{vmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \\ x_5 \\ x_6 \\ x_7 \\ x_8 \\ x_9 \\ x_{10} \\ x_{11} \\ x_{12} \\ x_{13} \\ x_{14} \\ x_{15} \\ x_{16} \\ x_{17} \\ x_{18} \\ x_{19} \\ x_{20} \\ x_{21} \\ x_{22} \\ x_{23} \\ x_{24} \\ x_{25} \end{vmatrix}
=
\begin{array}{c}
\begin{array}{cccccccccccccccccccccccc} 1 & 2 & 3 & 4 & 5 & 6 & 7 & 8 & 9 & 10 & 11 & 12 & 13 & 14 & 15 & 16 & 17 & 18 & 19 & 20 & 21 & 22 & 23 & 24 \end{array} \\
\begin{vmatrix}
1 & & & & & & & & & & & & & & & & & & & & & & & \\
 & 1 & & & & & & & & & & & & & & & & & & & & & & \\
 & & 1 & & & & & & & & & & & & & & & & & & & & & \\
 & & & 1 & & & & & & & & & & & & & & & & & & & & \\
 & & & & 1 & & & & & & & & & & & & & & & & & & & \\
 & & & & \lambda & & & & & & & & & & & & & & & & & & & \\
 & & & & & 1 & & & & & & & & & & & & & & & & & & \\
 & & & & & & 1 & & & & & & & & & & & & & & & & & \\
 & & & & & & & 1 & & & & & & & & & & & & & & & & \\
 & & & & & & & & 1 & & & & & & & & & & & & & & & \\
 & & & & & & & & & 1 & & & & & & & & & & & & & & \\
 & & & & & & & & & & 1 & & & & & & & & & & & & & \\
 & & & & & & & & & & & 1 & & & & & & & & & & & & \\
 & & & & & & & & & & & & 1 & & & & & & & & & & & \\
 & & & & & & & & & & & & & 1 & & & & & & & & & & \\
 & & & & & & & & & & & & & & 1 & & & & & & & & & \\
 & & & & & & & & & & & & & & & 1 & & & & & & & & \\
 & & & & & & & & & & & & & & & & 1 & & & & & & & \\
 & & & & & & & & & & & & & & & & & 1 & & & & & & \\
 & & & & & & & & & & & & & & & & & & 1 & & & & & \\
 & & & & & & & & & & & & & & & & & & & 1 & & & & \\
 & & & & & & & & & & & & & & & & & & & & 1 & & & \\
 & & & & & & & & & & & & & & & & & & & & & 1 & & \\
 & & & & & & & & & & & & & & & & & & & & & & 1 & \\
 & & & & & & & & & & & & & & & & & & & & & & & 1
\end{vmatrix}
\end{array}
* \begin{vmatrix} \xi_1 \\ \xi_2 \\ \xi_3 \\ \xi_4 \\ \xi_5 \\ \xi_6 \\ \xi_7 \\ \xi_8 \\ \xi_9 \\ \xi_{10} \\ \xi_{11} \\ \xi_{12} \\ \xi_{13} \\ \xi_{14} \\ \xi_{15} \\ \xi_{16} \\ \xi_{17} \\ \xi_{18} \\ \xi_{19} \\ \xi_{20} \\ \xi_{21} \\ \xi_{22} \\ \xi_{23} \\ \xi_{24} \\ \xi_{25} \end{vmatrix}
+
\begin{vmatrix} \delta_1 \\ 0 \\ \delta_3 \\ 0 \\ \delta_5 \\ \delta_6 \\ 0 \\ 0 \\ 0 \\ 0 \\ 0 \\ 0 \\ 0 \\ 0 \\ 0 \\ 0 \\ 0 \\ 0 \\ 0 \\ 0 \\ 0 \\ 0 \\ 0 \\ 0 \end{vmatrix}
$$

List A.5 (Section A.9, Figure A.19)
VARIABLES IN TEACHER ASSESSMENT MODEL

Symbols: Observed Variables	Symbols: Constructs	Acronym	Age	Variable Description
y_1		S22	5-9	Language assessment by teacher
y_2		S21	5-9	Mathematics assessment by teacher
	η_1		5-9	Teacher assessment
x_1	ξ_1	ACHHDC5	66	Home diet calories
x_2	ξ_2	CHSPC14	66	Supplementation calories
x_3	ξ_3	ACHHDP5	66	Home diet protein
x_4	ξ_4	CHSPP14	66	Supplementation protein
	η_2		66	Protein = $\xi_3 + \xi_4$
	η_3		66	Calories = $\xi_1 + \xi_2$
x_5		NAMING4	72	Naming psychometric test
x_6		RECOG4	72	Recognition psychometric test
	ξ_5		72	Verbal factor
x_7	ξ_6	CHDIAR15	60-72	Child's incidence of diarrhea: proportion
x_8	ξ_7	MORBOTH4	60-72	Child's incidence of ailments other than diarrhea: proportion
x_9	ξ_8	HTC15	60	Height in centimeters
x_{10}	ξ_9	FMSZ6	[illegible]	Number of household members at age
x_{11}	ξ_{10}	YOUNG6	72	Number of younger siblings at age
x_{12}	ξ_{11}	DIED6	72	Number of elder siblings deceased at age
x_{13}	ξ_{12}	DSTRUC6	72	Family structure: nuclear vs. all other
x_{14}	ξ_{13}	SEX		Sex

List A.5 continued

Symbols				
Observed Variables	Constructs	Acronym	Age	Variable Description
x_{15}	ξ_{14}	ZMARD		Mother's literacy and schooling
x_{16}	ξ_{15}	ZPARD		Father's literacy and schooling
x_{17}	ξ_{16}	MODVOC		Mother's modernity and vocabulary
x_{18}	ξ_{17}	CONSUMP		Housing, possessions, and salary
x_{19}	ξ_{18}	PASTAT		Father's occupation, ownership of land, animals
x_{20}	ξ_{19}	ZMAOCC		Mother's occupation
x_{21}	ξ_{20}	EDASP		Mother's educational aspirations for child
x_{22}	ξ_{21}	ATTAGE		Child's age in months at beginning of year assessed
x_{23}	ξ_{22}	D6		Atole - Large village
x_{24}	ξ_{23}	D8		Fresco - Small village
x_{25}	ξ_{24}	D14		Atole - Small village

Equation A.10 (Section A.9, Figure A.20)

Equations in Teacher Assessment Model

$$
\begin{vmatrix} 1 & \beta & \beta \\ & 1 & \\ & & 1 \end{vmatrix} * \begin{vmatrix} \eta_1 \\ \eta_2 \\ \eta_3 \end{vmatrix} = \begin{array}{c} \begin{array}{cccccccccccccccccccccccc} 1 & 2 & 3 & 4 & 5 & 6 & 7 & 8 & 9 & 10 & 11 & 12 & 13 & 14 & 15 & 16 & 17 & 18 & 19 & 20 & 21 & 22 & 23 & 24 \end{array} \\ \begin{vmatrix} & & & & \gamma & \gamma & \gamma & \gamma & \gamma & \gamma & \gamma & \gamma & \gamma & \gamma & \gamma & \gamma & \gamma & \gamma & \gamma & \gamma & \gamma & \gamma & \gamma & \gamma \\ & & 1 & 1 & & & & & & & & & & & & & & & & & & & & \\ 1 & 1 & & & & & & & & & & & & & & & & & & & & & & \end{vmatrix} \end{array} * \begin{vmatrix} \xi_1 \\ \xi_2 \\ \xi_3 \\ \xi_4 \\ \xi_5 \\ \xi_6 \\ \xi_7 \\ \xi_8 \\ \xi_9 \\ \xi_{10} \\ \xi_{11} \\ \xi_{12} \\ \xi_{13} \\ \xi_{14} \\ \xi_{15} \\ \xi_{16} \\ \xi_{17} \\ \xi_{18} \\ \xi_{19} \\ \xi_{20} \\ \xi_{21} \\ \xi_{22} \\ \xi_{23} \\ \xi_{24} \\ \xi_{25} \end{vmatrix} \begin{vmatrix} \xi_1 \\ 0 \\ 0 \end{vmatrix}
$$

$$
\begin{vmatrix} y_1 \\ y_2 \end{vmatrix} = \begin{vmatrix} 1 & 0 & 0 \\ \lambda & 0 & 0 \end{vmatrix} \begin{vmatrix} \eta_1 \\ \eta_2 \\ \eta_3 \end{vmatrix} + \begin{vmatrix} \epsilon_1 \\ \epsilon_2 \end{vmatrix}
$$

$$
\begin{vmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \\ x_5 \\ x_6 \\ x_7 \\ \vdots \\ x_{25} \end{vmatrix} = \begin{vmatrix} 1 & & & & & & & & \\ & 1 & & & & & & & \\ & & 1 & & & & & & \\ & & & 1 & & & & & \\ & & & & 1 & & & & \\ & & & & \lambda & & & & \\ & & & & & 1 & & & \\ & & & & & & 1 & & \\ & & & & & & & \ddots & \\ & & & & & & & & 1 \end{vmatrix} * \begin{vmatrix} \xi_1 \\ \xi_2 \\ \xi_3 \\ \xi_4 \\ \xi_5 \\ \xi_6 \\ \xi_7 \\ \vdots \\ \xi_{24} \end{vmatrix} + \begin{vmatrix} \delta_1 \\ 0 \\ \delta_3 \\ 0 \\ \delta_5 \\ \delta_6 \\ 0 \\ \vdots \\ 0 \end{vmatrix}
$$

Appendix A.11.c

Factor Analysis and Scaling

There are a very large number of indicators of various aspects of economic status available in the INCAP nutritional study. Most of these derive from the Rand/INCAP collaborative study conducted in March, 1977. A few additional related variables were generated at earlier dates based upon home interviewers' assessments of the quality of housing and clothing. There is also a smaller set of variables related to parental socialization practices, literacy, and attitudes.

The main purpose for undertaking factor analysis and scaling with these sets of items is to reduce the total number of variables to a more manageable number of dimensions. Two secondary reasons are to generate scales with greater reliability and validity as indicators of constructs such as "affluence" and "literacy," and to reduce the number of missing values by permitting sets of related items to serve as proxies for one another.

Twenty-nine economic indicators and nine "cultural" indicators were selected and analyzed as separate sets. Three factors were extracted from each of these sets. As the attached tables of factor patterns show, there is a high degree of consistency in the clustering of items ("loadings" on factors) when using principal components, orthogonal rotations, and oblique rotations.

The consistent clustering of items indicated by these factor analysis was used to group items for scaling. Three multi-item scales -- labeled MODVOC, PASTAT, and CONSUMP -- were generated by projecting a first principal component, estimating the single common factor among the selected set of items.

Three shorter "scales," represented by only two items each, were calculated as the average of the standardized variables (Z-scores). These scales were labeled ZPARD (father's literacy and schooling), ZMARD (mother's literacy and schooling), and ZMAOCC (mother's current and usual occupation).

The items included in each of the six scales are listed in List A.6, below. Items are ordered within scales by factor loadings.

The four variables representing father's and mother's current and usual occupations which enter into the PASTAT and ZMAOCC scales are based upon a collapsed ranking of the 27 category occupation code used in the Rand/INCAP 1975 study. The ranking is as follows:

4 - Farmer owning his own land

3 - Farmer working on family land, merchant, specialized laborer (i.e., skilled journeyman)

2 - Factory laborer, blue collar non-office employees, renting farmer

1 - Manufacturing laborer (construction, textile, ceramics, hats, etc.)

0 - Day laborers (agricultural and non-agricultural), domestic workers, unemployed, disabled

The scales within each set, calculated independently, are more highly intercorrelated than the oblique factors.

Table A.41

Intercorrelations Among Factor Scales

	MODVOC	ZMARD	ZPARD	CONSUMP	PASTAT	ZMAOCC
MODVOC	1.00					
ZMARD	.36	1.00				
ZPARD	.11	.24	1.00			
CONSUMP	.39	.27	.12	1.00		
PASTAT	.06	-.00	-.02	.28	1.00	
ZMAOCC	.05	.09	.11	.15	.05	1.00

List A.6

Items Included in Six Scales

MODVOC

MODREV	Modernity scale -- modification of Jones-Inkeles O-M scales
PTVOC	Vocabulary of mother
TEACH	Instruction of child by parents
MAINV	Mother's involvement in child rearing
PAINV	Father's involvement in child rearing

ZPARD

PASCH	Years of schooling of father
PAREAD	Literacy of father

ZMARD

MAREAD	Literacy of mother
MASCH	Years of schooling of mother

PASTAT

PAOCC2	Father's usual occupation
PAOCC1	Father's current occupation
LANDAN	Land owned by family
VGRA	Value of five major crops
PLANTY	Value of land planted
ANIMALSS	Value of all animals owned
HSOWN	Ownership of home

CONSUMP

HSTYP	Type of house: formal, semi-formal, rancho, improvised
HOUSER	INCAP 1974 raw house scale
WALLS	Walls material: stucco, adobe, bamboo, reed
HOUSE	INCAP 1974 house scale (SES-14)
ROOMS	Number of rooms in house
COOKPL	Location of cooking place: kitchen, separate, bedroom, none
ROOF	Roof material: metal, tile, straw
COOKTP	Type of cooking place: stone, high formal, low formal, high informal, floor, none
FLOOR	Floor material: blocks, cement, adobe, dirt
SEWING	Sewing machine
SALALL	Salaries earned by all
RADIO	Radio
SANIT	Sanitation: sceptic, latrine, none
RECORD	Record player
WATER	Water source: water system, own well, public well, river/lake
BIKE	Bicycle
DRAIN	Water drainage: hole in ground, floor

ZMAOCC

MAOCC1	Mother's current occupation
MAOCC2	Mother's usual occupation

TABLE A.42

Factor Analyses and Scaling of Parental Socialization Items

Variable	Principal Components			Varimax Rotation			Factor Pattern (δ=-1.5) Oblimin Rotation		
	1	2	3	1	2	3	1	2	3
1 PTVOC	.54	.24	.30	.25	.09	.60	.59*	-.06	-.15
2 MODREV	.46	.30	.33	.19	.01	.61	.61*	.03	-.09
3 MAINV	.24	.12	.36	-.03	.07	.45	.46*	-.06	.11
4 TEACH	.31	.21	.34	.06	.02	.50	.51*	.00	.02
5 PAINV	.33	.22	.13	.21	-.01	.36	.35*	.04	-.15
6 PAREAD	.50	-.66	.10	.10	.82	.06	.03	-.82*	-.03
7 PASCH	.59	-.76	.09	.14	.95	.07	.03	-.95*	-.07
8 MAREAD	.84	.21	-.46	.95	.16	.18	.09	-.10	-.94*
9 MASCH	.75	.22	-.35	.82	.12	.22	.14	-.07	-.80*
% of Var.	32.9%	18.5%	13.7%	54.8%	28.2%	17.0%	54.8%	28.2%	17.0%

	Factor Correlations		
	1	2	3
1	1.00		
2	-.09	1.00	
3	-.26	.14	1.00

Table continued on next page

TABLE A.42 (cont'd.)

PRINCIPAL COMPONENT
FACTOR SCORE (MODVOC)

Variable	Coefficient	N
1 PTVOC	.35	499
2 MODREV	.35	501
3 MAINV	.27	498
4 TEACH	.31	607
5 PAINV	.26	498

MISSING DATA

Items Missing	Frequency	Percent
5	225	26.8%
4	133	13.5
3	3	0.4
2	1	0.1
1	9	1.1
0	489	58.2
Total	840	100.0%
Scores Missing	225	26.8%
Scores Present	615	73.2

CORRELATIONS AMONG SCALES

	MODVOC	ZMARD	ZPARD	N
MODVOC	1.00			(615)
ZMARD	.36	1.00		(650)
ZPARD	.24	.11	1.00	(600)

TABLE A.43

Factor Analyses of Parental Economic Items

Variable	Principal Components			Varimax Rotation			Factor Pattern (δ=-1.0) Oblimin Rotation		
	1	2	3	1	2	3	1	2	3
1 HOUSE	.74	.02	-.04	.70*	.20	.06	.71*	.01	.11
2 CLOTH*	.26	.08	.06	.23	.13	.10	.23	.09	.10
3 HOUSER	.85	-.18	-.10	.84*	.23	-.14	.85*	-.19	.11
4 HSTYP	.87	-.12	-.10	.86*	.22	-.08	.87*	-.14	.10
5 HSOWN	.16	-.12	.20	.08	.26*	-.07	.09	-.09	.25*
6 ROOMS	.63	.12	-.09	.62*	.11	.15	.62*	.11	.02
7 FLOOR	.44	.17	-.05	.43*	.07	.19	.43*	.16	.01
8 ROOF	.59	-.22	-.05	.57*	.19	-.18	.59*	-.22	.11
9 WALLS	.81	-.17	-.12	.81*	.19	-.14	.82*	-.19	.08
10 COOKPL	.62	-.02	-.16	.63*	.06	-.00	.64*	-.04	-.03
11 COOKTP	.58	-.10	-.09	.57*	.13	-.07	.58*	-.11	.05
12 ELEC*	-.01	.27	.04	-.03	-.01	.27	-.03	.27	-.01
13 SANIT	.32	.11	.02	.29*	.10	.13	.29*	.11	.06
14 DRAIN	.10	.10	-.10	.13*	-.08	.09	.13*	.08	-.09
15 WATER	.21	-.07	.15	.15*	.22	-.03	.15*	-.05	.20
16 RADIO	.40	.06	.08	.35*	.19	.10	.35*	.07	.14
17 RECORD	.20	.02	-.15	.24*	-.07	.00	.24*	-.01	-.11
18 TV*	.03	.11	.08	.00	.07	.12	-.00	.12	.07
19 BIKE	.16	.08	.03	.14*	.06	.09	.14*	.08	.04
20 SEWING	.36	.17	-.12	.38*	-.02	.17	.38*	.15	-.07
21 PAOCC1	.27	-.12	.51	.08	.59*	-.02	.08	-.04	.58*
22 PAOCC2	.27	-.14	.56	.07	.63*	-.02	.07	-.05	.62*
23 MAOCC1	.21	.89	-.02	.20	-.12	.88*	.18	.88*	-.15
24 MAOCC2	.21	.83	-.06				.19	.82*	-.17

Table continued on next page

TABLE A.43 (cont'd.)

Variable	Principal Components			Varimax Rotation			Factor Pattern (δ=-1.0) Oblimin Rotation		
	1	2	3	1	2	3	1	2	3
25 PLANTV	.02	.26	.48	-.15	.40*	.33	-.16	.33	.42*
26 VGRA	.12	.24	.47	-.05	.43*	.32	-.06	.31	.44*
27 SALALL	.26	.03	.34	.36*	-.24	-.01	.37*	-.02	-.29
28 LANDAN	.35	-.05	.46	.17	.55*	.05	.17	.02	.53*
29 ANIMALSJ	.38	-.04	.26	.27	.37*	.03	.27	-.00	.34*
% of Variance	20.8%	8.2%	8.0%	60.4%	21.7%	17.9%			

*Omitted from scales.

	Factor Correlations		
	1	2	3
1	1.00		
2	.09	1.00	
3	.14	.05	1.00

TABLE A.44

SCALING OF PARENTAL ECONOMIC ITEMS

Principal Component Factor Scores

PASTAT		
Variable	Coefficient	N
1 HSOWN	.16	660
2 PAOCC1	.29	600
3 PAOCC2	.30	597
4 PLANTV	.21	627
5 VGRA	.22	627
6 LANDAN	.29	627
7 ANIMALSJ	.22	627

CONSUMP		
Variable	Coefficient	N
1 HOUSE	.14	607
2 HOUSER	.16	508
3 HSTYP	.16	660
4 ROOMS	.12	660
5 FLOOR	.09	660
6 ROOF	.12	660
7 WALLS	.15	660
8 COOKPL	.12	660
9 COOKTP	.11	660
10 SANIT	.06	660
11 DRAIN	.02	660
12 WATER	.04	660
13 RADIO	.08	660
14 RECORD	.05	660
15 BIKE	.03	660
16 SEWING	.07	660
17 SALALL	.06	627

Items Missing	Frequency	Percent
7	175	20.8%
6	10	1.2
5	0	0.0
4	28	3.3
3	5	0.6
2	50	6.0
1	3	0.4
0	569	67.7
Scores Missing	185	22.0%
Scores Present	655	78.0

Items Missing	Frequency	Percent
17	122	14.5%
16	42	5.0
15	15	1.8
14	1	0.1
13	0	0.0
12	0	0.0
11	0	0.0
10	0	0.0
9	0	0.0
8	0	0.0
7	0	0.0
6	0	0.0
5	0	0.0
4	0	0.0
3	11	1.3
2	100	11.9
1	80	9.5
0	469	55.8
Scores Missing	180	21.4%
Scores Present	660	78.6

Table continued on next page

TABLE A.44 (cont'd.)

Correlation Among Scales

	CONSUMP	PASTAT	ZMAOCC	N
CONSUMP	1.00			(660)
PASTAT	.28	1.00		(655)
ZMAOCC	.15	.05	1.00	(650)

Appendix A.11.d

Missing Data and Outlier's

The longitudinal study cohorts in the INCAP study consists of children born between January 1969 and February 1973 who lived in one of the four study villages before reaching the age of seven. In 1975, several files of data for 743 children meeting these criteria were provided by INCAP to the Berkeley group for secondary analysis.

The only variables for which all cases had data were sex and birthdate (and, artifactually, aggregated supplementation -- assuming that no report of supplementation implied none consumed). The number of cases having data present for other variables ranged from about 600 to 60.

A large amount of the missing data at the older age levels could readily be accounted for by the fact that the younger cohorts had not yet reached those age levels, e.g., children born in 1973 could not have anthropometry, psychometry, or other measurements for ages 60 and 72 months. A smaller number died at some point before their seventh year, and these dates of death were recorded.

A third factor which would not involve potential self-selection bias is geographic mobility. Both INCAP and Berkeley staff have devised screens based upon the absence of both morbidity and supplementation reports to distinguish cases among the missing data which were "away from the village" from those which were otherwise unaccounted for.

This procedure, however, presumes the completeness of the field work coverage, reporting, and recording which itself might be a source of missing data. Periodic household censuses of the villages were conducted by INCAP, and one special census was conducted in 1975 in

connection with the Rockefeller sub-study. A printout of the last entry for each child in the four villages (dated 7 MAR 1978) was received from INCAP, and the 1975 census data were included in the Rockefeller study materials sent by the Rand Corporation to Berkeley.

Both of these censuses included for each person a date of arrival in the village, or, if departed, a date of departure. (The INCAP 425 census listing, however, overwrote any possible prior information on arrival or departure for those children who reached the age of seven, supplanting this information with the date of their attaining the age of seven -- which is implicit in their birthdate.)

Because these censuses, for those not yet seven years old, contained a positive indication of a reported date of arrival or departure, a collation was made to update our files from these data. Since the INCAP census was a printout, the scan was made visually. The appropriate age groups were sorted out by birthdates from the Rand tapes.

Missing and Duplicate Cases -- All of the 743 cases originally provided to Berkeley were located on the INCAP 425 census. Two of these cases were indicated by the field worker taking the census to be invalid duplicate identification numbers of other persons in the sample. These cases did, indeed, have the same birthdates, sex, and name. A listing of data available here for these two cases showed one to have no data, the other to have only 36 months psychometry (which was missing under the other ID number). These latter data were merged with the data in the proper ID number, and the sample thus reduced to 741.

Since the listing of this census was alphabetical, by name, the search for sample cases was done by scanning birthdates. This led to the discovery of 87 additional cases which had been recorded in these

periodic INCAP censuses, and were born during our sample period, but who were not included in the original sample of 743.

The further subsequent search of the Rand files of the 1975 census located many of these 87 cases and an additional 12 cases also not included in the original 743 (now reduced to 741).

This search for arrival and departure dates in order to account for missing data had the initial effect of exacerbating the volume of missing data by adding 99 cases for which there were no data other than the census data itself and the socioeconomic data describing families.

The discovery of 12 cases in the 1975 census which were not included in the INCAP cumulative census (and many of the 87 found in the cumulative census were not in the 1975 census although reported to have arrived before that date) leaves an uncertainty about the completeness of any of the files, including the present aggregate of 840 cases, as exhausting the target population of interest. This question cannot be answered from data presently at hand in Berkeley. (16 of the additional 99 cases which were located were ultimately eliminated from the sample base -- e.g., children who died before the family immigrated to the study villages.)

After collating arrival and departure dates into the existing files of data, missing data for about 400 variables were disaggregated into these screening categories: deceased, not arrived, departed, not of age, and unaccounted for. A tally of these distributions is attached to this note for the three groups of cases -- the original 741, the 87 from the INCAP listing, and the 12 from the Rand census. A comparison of these distributions shows higher mortality and mobility among the

99 additional cases than among the 741 cases. For example, by age 84 months, only 9 of the 99 cases (9%) should have been expected to have data as compared to 157 of the 741 original cases (21%).

Since the socioeconomic data from the Rockefeller study was available for all 12 cases located in those files and for about a half of the additionally located 87 cases, a comparison was made between the three groups using the five factor scores based on these data. None of these contrasts proved significant. The higher morbidity and mobility rates already mentioned and reflected in the attached tally are the only distinguishing variables among the few available.

TABLE A.45

Frequency Distribution -- Present and Missing Anthropometry (Weight) by Age	1 Number of Cases	2 Not Arrived	3 Departed	4 Not of Age	5 Deceased	6 Data Present	7 Missing	8 7/(6+7)
Birth	840	122	2		23	410	283	.41
12 months	840	86	26		64	463	201	.30
24 months	840	82	53		82	470	153	.26
36 months	840	66	70		84	472	148	.24
42 months	840	48	95		87	472	138	.23
60 months	840	26	88	171	87	375	93	.20
72 months	840	12	69	355	88	241	75	.24
84 months	840	0	47	537	88	122	46	.27
Present and missing morbidity (Diarrhea) when data for any portion of an interval are used								
9 - 12 months	840	90	24		60	482	184	.28
21 - 24 months	840	78	51		79	526	106	.17
30 - 36 months	840	57	64		82	539	98	.15
42 - 48 months	840	44	86		87	537	86	.14
48 - 60 months	840	29	101	19	87	551	53	.09
60 - 72 months	840	17	91	197	87	416	32	.07
72 - 84 months	840	3	70	380	88	282	17	.06
Presence and absence of selected family background variables								
HSSIZ	840	42	88			528	182	.26
ZPARD	840	43	84			600	113	.16
ZMARD	840	40	81			650	69	.10
ZMAOCC	840	40	81			650	69	.10

TABLE A.46

Frequency Distribution
Data Status by Source of Case

Age	Number of Cases in Group	Not Arrived	Departed	Not of Age	Deceased	Data Expected
1/2 month	87	17	1		15	54
	12	1	0		4	7
	741	121	1		19	600
12 months	87	16	3		21	47
	12	1	0		6	5
	741	94	24		39	584
24 months	87	16	8		24	39
	12	1	0		6	5
	741	82	46		52	561
36 months	87	15	13		24	35
	12	1	0		6	5
	741	64	58		54	565
48 months	87	11	16		24	36
	12	1	0		6	5
	741	44	80		57	560
60 months	87	6	11	25	24	21
	12	1	0	1	6	4
	741	25	78	145	57	436
72 months	87	2	6	42	24	13
	12	0	0	2	6	4
	741	12	64	311	58	296
84 months	87	0	6	49	24	8
	12	0	0	5	6	1
	741	0	43	483	58	157

TABLE A.47

Means of Selected Variables for Sub-Samples

	87	12	741	840
Duration	50.0	71.6	61.2*	60.2
Stay	.71	.92	.83	.82
Cohort	70.1	77.0	74.0*	73.6
ZPARD	-.05	.53	.03	
ZMARD	-.17	-.11	.01	-.00
PARSTIM3	-.08	.07	.01	
MODVOC	---	---	.00	.00
SES3	-.02	.10	-.01	-.01
ZMAOCC	.12	-.21	-.00	.00
CONSUMP	-.01	.20	-.01	-.00
PASTAT	-.03	-.34	.01	.01
SEX	.54	.45	.54	.54

TABLE A.48

Distribution of Missing, Unavailable, and Valid Data for 357 Variables Among 840 Cases

	Missing	Not Available	Valid Data
None	0	14	99
0 - 10%	65	219	3
10 - 20	486	281	27
20 - 30	143	80	117
30 - 40	29	31	61
40 - 50	19	45	32
50 - 60	31	53	43
60 - 70	19	46	180
70 - 80	13	30	252
80 - 90	25	41	26
90 - 100	6	0	0
All	4	0	0
Total	840	840	840
Mean	23.7%	26.4%	49.9%
s	19.4%	23.9%	26.7%

TABLE A.49

Numbers of Cases of Valid Data for Selected Cross-Sectional Variables
(Not varying -- or not remeasured -- with age)
Data Items = 9,145

	COHORT	SEX	MAHT	PAHT	VILLAGE	ZPARD	ZMARD	MODVOC	FAMSIZE	PASTAT	ZMAOCC	CONSUMP	NSLDPPR	NSLDTCA
COHORT	840													
SEX	837	837												
MAHT	576	576	576											
PAHT	440	440	415	440										
VILLAGE	840	837	576	440	840									
ZPARD	600	598	485	418	600	600								
ZMARD	650	647	512	420	650	591	650							
MODVOC	615	615	557	430	615	517	551	615						
FAMSIZE	528	528	459	380	528	489	519	487	528					
PASTAT	655	653	516	428	655	600	641	557	522	655				
ZMAOCC	650	647	512	420	650	591	650	551	519	641	650			
CONSUMP	660	657	519	427	660	600	650	559	528	650	650	660		
NSLDPPR	627	625	496	417	627	572	614	535	499	627	622	614	627	
NSLDTCA	627	625	496	417	627	572	614	535	499	627	622	614	627	627

Table continued on next page

TABLE A.49 (cont'd.)

Numbers of Cases Among Those Living in Village at Both Age 36 Months and 48 Months

Data Items = 7,395

	COHORT	SEX	MAHT	PAHT	VILLAGE	ZPARD	ZMARD	MODVOC	FAMSIZE	PASTAT	ZMAOCC	CONSUMP	NSLDPPR	NSLDTCA
COHORT	596													
SEX	596	596												
MAHT	483	483	483											
PAHT	395	395	373	395										
VILLAGE	596	596	483	395	596									
ZPARD	510	510	438	382	510	510								
ZMARD	542	542	463	381	542	501	542							
MODVOC	517	517	480	388	517	466	493	517						
FAMSIZE	466	466	417	347	466	436	458	442	466					
PASTAT	547	547	466	388	547	510	536	497	461	547				
ZMAOCC	542	542	463	381	542	501	542	493	458	536	542			
CONSUMP	551	551	470	388	551	510	542	501	466	545	542	551		
NSLDPPR	527	527	452	379	527	490	517	481	444	527	525	517	527	
NSLDTCA	527	527	452	379	527	490	517	481	444	527	525	517	527	527

TABLE A.50

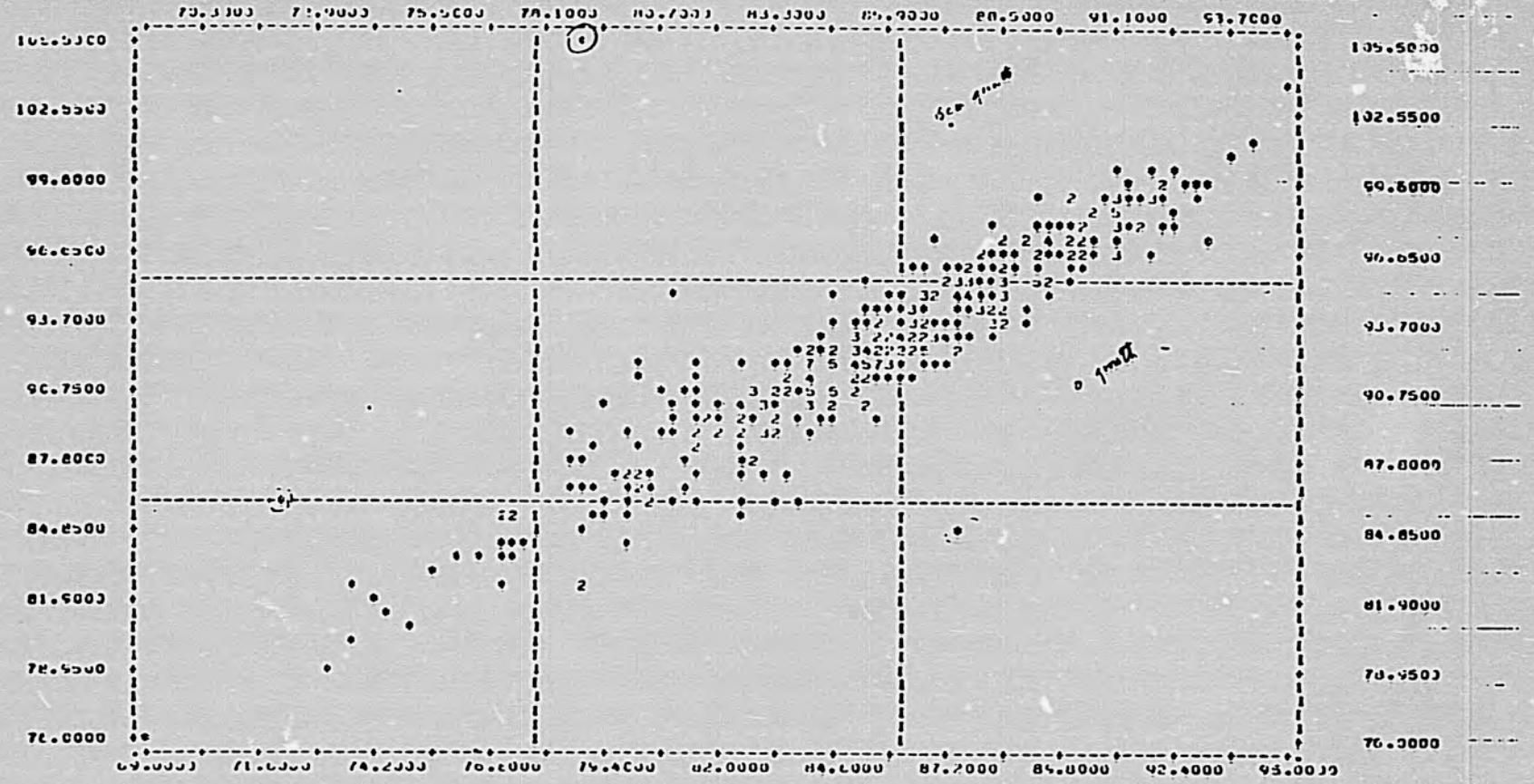
Numbers of Cases of Valid Data for Selected Sequential (Age-level) INCAP Data

Acronym	MONTHS OF AGE							
	12	24	36	48	60	72	84	TOTAL
HT	460	468	430	429	374	241	122	2,524
NAMING			481	507	393	264	139	1,784
RECOG			481	507	393	264	139	1,784
CHSP	705	694	692	689	535	362	192	3,869
ACHHDC		503	514	517	451	173	68	2,226
ACHHDP		503	514	517	451	173	68	2,226
CHDIAR	523	571	552	559	551	416	282	3,454
CHANOR	523	571	552	559	551	416	282	3,454
CHHLTH	523	571	552	559	551	416	282	3,454
MAATT	840	840	840	840	840	840	840	5,880
VILSUP	705	694	692	689	535	362	192	3,869
WTAVP	513	502	516	505	517	392	259	3,204
ACHC		499	513	517	397	172	68	2,166
ACHP		499	513	517	397	172	68	2,166
ACHCP		499	513	517	397	172	68	2,166
ACHPCP		499	513	517	397	172	68	2,166
CHSPP	705	694	692	689	629	362	192	3,963
CHSPC	705	694	692	689	629	362	192	3,963
Total	6,202	9,301	10,252	10,323	8,988	5,731	3,521	54,318

Figure A.25

MODEL VARIABLE DISTRIBUTIONS AND SCATTERGRAMS 01/13/79 17.35.56. PAGE 22

FILE INCAPX (CREATION DATE = 01/12/79)
SORTED BY ATOM ATOM ATOM
SCATTERGRAM OF (ATOM) ATOM 40 40 HEIGHT IN CENTIMETERS (ACROSS) HEIGHT 36 MO. HEIGHT IN CENTIMETERS



MODEL VARIABLE DISTRIBUTIONS AND SCATTERGRAMS 01/13/79 17.35.56. PAGE 23

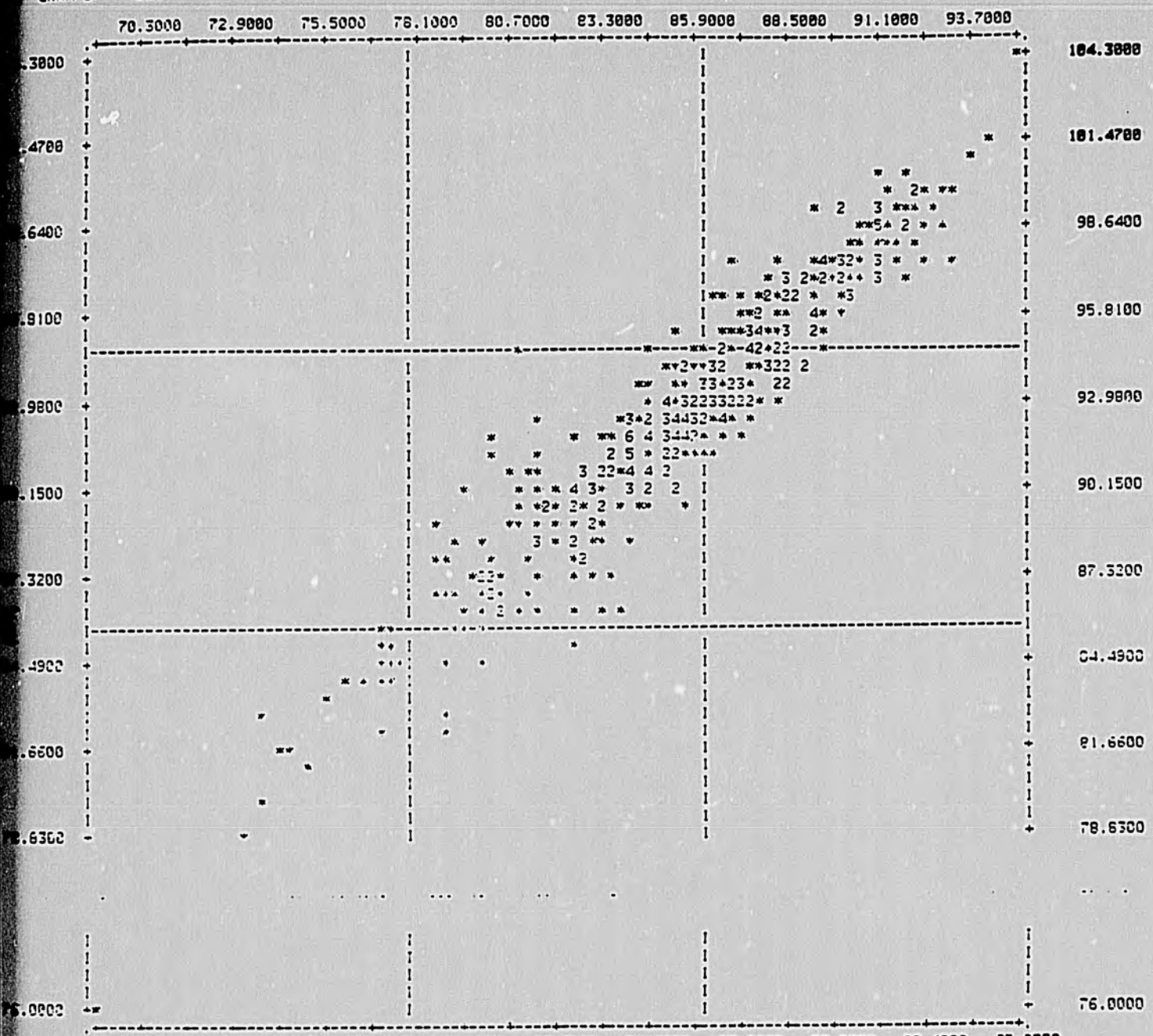
STATISTICS..

CORRELATION (R)-	.91560	R SQUARED -	.83897	SIGNIFICANCE R -	.00001
STD ERR OF EST -	1.72001	INTERCEPT (A) -	11.82128	STD ERROR OF A -	1.72821
SIGNIFICANCE A -	.00001	SLOPE (B) -	.96703	STD ERROR OF B -	.02319
SIGNIFICANCE B -	.00001				
PLOTTED VALUES -	651	LACKING VALUES -	0	MISSING VALUES -	103

******** IS PRINTED IF A VALUE EXCEEDS CAPACITY OF COMPUTER.

Figure A.26

INCAPX (CREATION DATE = 01/14/79)
FR03 AT06 FR08 AT14
GRAM OF (DOWN) HT14 48 MO HEIGHT IN CENTIMETERS (ACROSS) HT12 36 MO HEIGHT IN CENTIMETERS



TEPGRAMS 01/24/79 20.03.03. PAGE 3

ISTICS..

RRELATION (R)-	.94737	R SQUARED -	.89751	SIGNIFICANCE R -	.00001
TD ERR OF EST -	1.34557	INTERCEPT (A) -	8.99538	STD ERROR OF A -	1.37422
IGNIFICANCE A -	.00001	SLOPE (B) -	.98021	STD ERROR OF B -	.01605
IGNIFICANCE B -	.00001				
TED VALUES -	426	EXCLUDED VALUES-	0	MISSING VALUES -	412

******** IS PRINTED IF A COEFFICIENT CANNOT BE COMPUTED

II-B. School Enrollment and Achievement: Effects of Nutrition, Health, and Need for Children's Work

Judith B. Balderston

The objective of this study was to test for the effects of nutrition, health, and parental occupation, education, and affluence upon children's school enrollment and achievement. In the context of the four villages in Eastern Guatemala, such analysis must take into account the fact that, although schooling is compulsory according to law, customarily approximately half of the children in rural areas actually attend school even at the lowest primary levels. Children's work activities generally take place at home or on the family agricultural plots, although it will also be seen that some children also earn money in paid hourly work, mostly agricultural. Many children attend school and undertake work activities during their out-of-school hours.

The central question of the Berkeley study was whether school enrollment and achievement depend on children's nutrition and health. Results of Alan Wilson's analysis reported in Section II-A show that there are nutritional and health effects on children's verbal scores and on school achievement. In the present section, prior nutrition and health will be assumed to have already affected height by age seven. Present nutritional intake and morbidity measures are used as indicators of present health. Family work and the need for children's help will be included in the analysis.

It was hypothesized that several factors would contribute to the family's decisions with respect to the child's schooling. First, the economic need of the family for the child's work would be determined by family occupation and land holdings, family size and composition, including age and sex of all members, as well as family income and accumulated wealth. Second, in addition to these measurable economic factors, attitudinal factors

such as parental attitudes and aspirations for the child would be the result of parental literacy and perceptions of child economic utility. Third, the effects of children's past nutritional intake and health, physical development, and verbal ability would influence the child's enrollment and performance in school.

The direction of these effects is difficult to assess a priori. While Alan Wilson's findings indicate that taller children also achieve higher verbal development and appear to be more successful in school, healthier children may also be more valuable for work on the family farm, making their parents less willing to dispense with their help. (The opposite might happen with weak, small children who, being perceived as unable to profit from schooling, might also be judged to be of little help in the family work activities.) Moreover, healthier, more robust children might be able to carry out both work and school activities. The sign of the effects of nutritional status on school participation is, therefore, not determined unless information is also available on the relation between family economic needs and children's work activities.

Although poor participation, dropout, repetitions, and failure are serious sources of wastage in many educational systems, and although it is suspected that malnutrition and poor health may contribute to these problems, surprisingly few studies have sought to identify the contribution of diet and health to school success. In fact, many school feeding programs have been instituted without evaluation of their results.

Many authors have recognized family background factors inherent in school wastage, and a few studies have attempted to collect information on parental background and income (1 to 5). But, in none of these studies

were students weighed or measured or asked questions about their diet. In the few cross sectional studies relating educational achievement and nutritional and health status, data are too aggregated to test for causation. From these studies, it is apparent that children from low SES families have poorer chances of school success, less adequate diet, and generally less stimulating social environment. Disentangling the nutritional component in these studies is difficult because (a) extensive nutritional data have never been collected; and (b) in purely observational cross-sectional studies, it is difficult to control for SES and isolate nutritional and health effects on school performance.

Recently, important work by Marcelo Selowsky (6) has presented a theoretical framework for examining relationships between early health and nutrition and later school achievement and productivity. Popkin and Lim-Ybanez (7) have sought to explain relationships between nutrition and school achievement. Shortlidges's (8) work on the determinants of school attendance in India employs Becker's theory of household choice to explain parental decisions on the use of children's time. In the present study, we shall approach two of these questions -- the relationships between children's nutrition and schooling and the parental decisions to enroll children based on need for the children's work in family economic activities.

From the longitudinal study carried out by INCAP, it was possible to identify determinants of height, verbal development, and school enrollment; results of such analysis were presented in the previous section of this report by Alan Wilson. Because of the presence of the supplementation program, it was possible to separate out the effects of family economic status from the child's food intake since children at low income levels had the opportunity to obtain improved nutrition than they would otherwise have had. The INCAP experiment and data collection, therefore, permitted us to

analyze how attained size at age seven, as a proxy measure of prior nutritional intake and morbidity, along with current measures of health, affected school enrollment* and performance. The availability of information on family work, income, wealth, parental literacy, and perceptions of the economic utility of children, which had been collected in 1974-75 by Rand, with Rockefeller support, permitted us to analyze the relationships between family economic conditions, children's work activities, and school attendance. Unfortunately, the data collected did not provide a complete set of information on nutrition, health, growth, children's work activities, family economic background, and school attendance for any subjects; for the three sets of data collected, there were not comparable variables. Data collected by INCAP, with NSF support, included information on children's school performance (from the Guatemala Ministry of Education) and on children's work for pay. We therefore had to make use of available data by carrying out the analysis in stages which will be described in a later section.

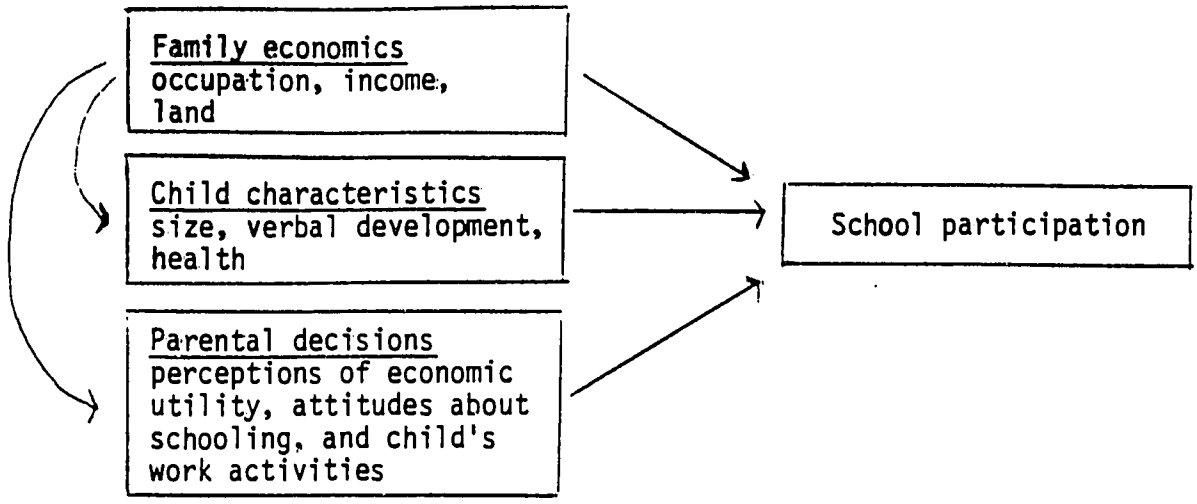
Framework of the Analysis

The participation of a child in school can be seen to be the result of a number of factors which fall into three general categories: the family's economic conditions, the child's individual characteristics, and the parents' perceptions and decisions about the allocating of their children's time. We can portray this in a simple model where we propose that family economics, child characteristics, and parental attitudes contribute directly

---

*Here school enrollment, attendance, and participation will be used interchangeably since the data collected by INCAP defined attendance as completion of the school year as indicated by grades received. Data on enrollment and actual attendance would have been preferable but were not available for all the villages.

to the child's school participation while recognizing that these factors themselves are interrelated in complex ways.



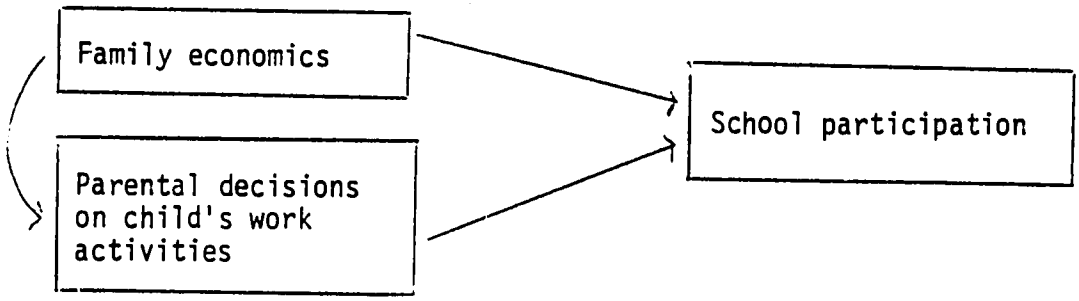
Family economic conditions include family work, income, size of family, and land holdings. Child's characteristics include sex, size, health, and mental development. Family attitudes shape the decisions that parents make, their attitudes about desired family size, and their decisions concerning the allocation of children's time. These attitudes control whether parents are willing to defer the children's present help in favor of potentially higher future earnings due to school attainment. In this paper, we use a scale of parent's perception of children's economic utility, measures of parental educational aspirations for their children, and measures of the work activities actually undertaken by the children to estimate the effect on children's school participation.

The model above involves two kinds of simplification that need to be recognized. First, family economic conditions, of course, affect both the child's development and parental decisions concerning children's work, but the mechanism by which they do this is not completely clear. Second, another arrow is missing that would connect the child's health and development with family decisions about the child's work and schooling; ignoring this

relationship potentially introduces bias. It would be important in a further investigation to look at the relationship of the health and size of the child to the kind of work performed, but in this paper, we had to ignore this relationship. In fact, because of data limitations, we could not estimate the full model as shown because data were not available that would allow us to include school, work activities, health, and development variables for the same group of subjects.

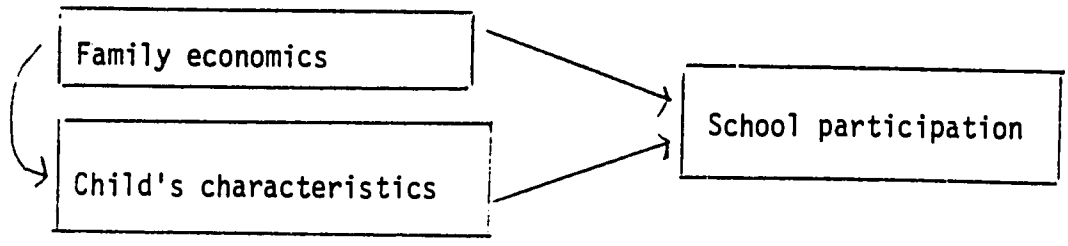
We therefore had to estimate separately the following models:

Model A. The influence of economic need for the child's work was examined by way of a simple model:



Here, we looked only at the economic variables, neglecting health, verbal development, and size of the child because these elements were not available for that sample for which work information was available. We concentrated on how family economics appeared to affect children's work and school participation. By specifying the model in this form, we introduce bias because of the neglected variables of child's health and size which are correlated with economic group and affluence. Therefore, we tend to overestimate the effect of the need for the children's work as a cause of non-enrollment in school. We expect that in poorer families, school enrollments would be lower because of the need for children's economic contribution and also because poorer families suffer from poorer nutrition and health; by omitting health/nutrition variables, which are highly correlated with the included variables, we tend to overstate the effect of the included variables.

Model B. Similarly, we looked at the influence of family economics and size and health and verbal development on school enrollment. Because of the presence of the supplementation experiment and the introduction of health care into the four communities in the study, it was possible, partially, to separate the influence of health and physical development from family economic background. Thus, we have another simple model:



In this case, the model tends to overstate the importance of health and size by ignoring the fact that less affluent families may need to have their children work more and attend school less, while more affluent families whose children may be better nourished and developed are also more likely to be encouraged to attend school. The presence of the INCAP experiment over the long period of time may have served to break the connection between family affluence and child's physical growth.

Model C. To avoid the problems of either A or B above, we sought a method that would take advantage of the two sets of data and that would improve the power of each of the two oversimplified models. We therefore used the results of A, where family economic group appeared to relate to school enrollment, and partitioned the sample into four economic groups -- three groups of farmers and one non-farming group. Then, returning to the three cohorts for which we had size and health data, we used model B and examined school enrollment patterns within economic groups to test for the independent effects of health and size on school enrollment and achievement. We also split the sample by villages and noted that there

were different patterns of school participation among villages due to differing economic characteristics of the villages.

Information Sources

There were three sources of information available to us on children's school and work activities:

1.) The "R09 Files" consisting of information collected by Rand with the support of the Rockefeller Foundation, which provided detailed measures of children's daily activities collected over four rounds. The information was reported by mothers of 552 children over seven years of age in 1974-75.

2.) The INCAP schooling study consisting of information on 714 children born between 1962 and 1968 for whom there were data on school enrollment and achievement for 1972 to 1978. Included in these data was information on children's paid work activities.

3.) Information about school enrollment and achievement for the 543 children born between 1969 and 1971 who had been followed since birth in the INCAP longitudinal study and who were eligible to enter school by 1978.

These files were merged by individaul child with other data available from Rand and INCAP. In children born before 1969, the other data came chiefly from Rand files and did not include measures of children's growth or health. In children born between 1969 and 1971, a large set of data on anthropometry, morbidity, and diet were available in addition to information on the family from the Rand files.

1.) R09 Files from Rand for Children Born Before 1967

The earlier cohorts studied in 1974-75 by Rand under a project financed by the Rockefeller Foundation collected data on family work, family size, decisions, parents' expectations, and an ambitious study of children's activities. The last mentioned study was based on information collected by Rand and

INCAP from mothers of children who were over seven years of age. There was also a corresponding set of information collected on the activities of children under seven, but only when the child accompanied the mother in undertaking certain activities. Thus, for children under seven, the study is more accurately one of mothers' activities rather than children's, and we decided to use only the material on children's independent activities. In children age seven and over, the list of activities included chores to help the family, work for pay, farm work, etc.[1] The method of the inquiry was to ask mothers if their children performed these activities during the previous day. Interviews were conducted over four rounds.[2] We selected information from the third round which was considered to be more accurate than the first two rounds and which gave activites for a day when school was in session. The information thus collected from the Rand study of children's activities was used in conjunction with other information collected by Rand under other parts of its study. By merging files from the various parts of the Rand studies and matching by family and child, we were able to examine how children differed in the activities performed according to sex, age, village, and family size and economic need.

From the list of all activities, a variable, "work," was computed to indicate whether the child undertook any work activities. Then, forming four variables, "work only," "school only," "work and school," and "neither work nor school," we performed regression analysis on these variables to test hypotheses that family work, size, village, and relative wealth

---

[1]The complete list of activities included prepared food for family, took care of children, went shopping, carried water, carried wood, did laundry, carried meals, did farm work, sold products, worked away from house for pay, worked within household, other activities inside house, visited relatives and other activities outside house, attended school, visited clinic, and spend nights away from house.

[2]For detailed analysis of these achievements, see Carol Clark, "Relation of Economic and Demographic Factors to Household Decisions Regarding Education of Children in Guatemala," Ph.D. dissertation, University of Michigan, 1979.

affected the disposition of the children's time. We also disaggregated by farm groups (non-farmers, subsistence farmers, semi-subsistence, and commercial farmers) to examine how children's activities appear to be related to parental economic group.

2.) The Schooling Study From INCAP for Children Born From 1962 to 1968

In a separate set of information collected by INCAP in 1978 under a study sponsored by the National Science Foundation, data were available to us to examine children's patterns of paid work, for children of age 10 to 15 in 1978, as well as their school enrollment and achievement from 1972 to 1978. Missing in this part of the study was information on children's help to their families, the kind of chores done without compensation which were studied in the Rand activities study. Using the information so collected in conjunction with the same set of economic variables that we employed with the Rand data, we were able to extend the amount of information available to us on which children worked for pay and which children attended school.

This set of data enabled us to look at the most recent evidence on whether children were attending school and/or had worked for pay in the past. In this case, we were looking at "school," "work for pay," "school and work for pay," and "neither school nor work for pay." We carried out regression analysis, using Ordinarly Least Squares, to test hypotheses that enrollment and paid work were affected by family economic activities, village, sex, and age. We aggregated, as above, all occupational groups together and then repeated the statistical tests by the four categories (non-farmers, subsistence farmers, semi-subsistence, and commercial).

Here again, as in 1.) above, we did not have diet, growth, or health measurements. Thus, this part of the study was useful in showing relationships only between economic variables and schooling/work patterns.

3.) Schooling Data for Children Born Between 1969 and 1971

Schooling data collected by INCAP from the Guatemelan Ministry of Education records (with NSF support) were available for all children who participated in the longitudinal study who were born before December 31, 1972.

For these children, a complete set of psychological anthropometric, food intake, morbidity variables was potentially available and was used by Alan Wilson in his work reported in Section II-A. For the purposes of the present section, variables selected included only cross sectional data at age seven, the school entering age.

Method

We examined school enrollment, as the dependent variable, using three approaches which correspond to the models desribed above.

Model A - Explanation of school enrollment by work activities for children born between 1962 and 1968 as the result of village location, age, sex, father's occupational group*, family income, mother's perception of the economic utility**, family size and order. This was done using two different sets of data on work activities:

1.) Those describing all work activities during one day, as collected in the Rand/Rockefeller file R09.

2.) Those describing only paid work activities over a several year period as collected by INCAP in their schooling study.

Model B - Explanation of school enrollment by child's developmental, height and health characteristics for children born between 1969 and 1971:

---

*For a full discussion of occupational groupings, see Part II-C of this report by Maria Freire.

**Measures of mother's perception of economic utility of children were obtained with slight modification from the scale constructed by Mari Simonen and discussed in Part II-D in this report.

1.) School enrollment was hypothesized to be determined by prior health and nutrition as measured by attained physical height and verbal development at age seven, as well as concurrent home diet, supplementation, health, and a variety of other variables describing family affluence, literacy, and family size.

2.) School enrollment was hypothesized to be determined by height, verbal development, health, family economics, and literacy of parents.

Model C - Explanation of school enrollment by child's developmental characteristics partitioned by village and economic groups for children born between 1969 and 1971:

1.) Using the same variables as in (B-1), data were partitioned by village.

2.) Using the same variables as in (B-1), data were partitioned by economic group.

Findings

We begin with the study of children's activities from Rand's R09 files and children's paid work from INCAP's files and summarize the results so obtained. With those results as background, we then discuss the findings of the analysis of school participation for the children followed in the longitudinal study.

Model A - Results of the Study of School Enrollment And Work Activites:

1.) Starting with the Rand data, which provided detailed information on children's activities during one typical day when school was in session, we see from Table B-1 that for all economic groups and villages combined that there are some statistically significant variables which explain the mix of children's activities. We note that R^2 results are low, indicating that there are other explanatory factors besides the ones that are included

in the model that explain the greatest part of the variance in behavior.

The method used here was to consider school only, work only, school and work, and neither school nor work*, as four dichotomous dependent variables; in each case, whether the child did the specific activity was hypothesized to be determined by a variety of factors. Since we recognized that performance of any of these activities was related to non-performance of the others, the results must be interpreted carefully. Moreover, regression analysis for dichotomous dependent variables is suitable when the probability of an event occuring is near .5 but is inappropriate where probabilities are near 0 or 1. Under the latter conditions, a preferable method is logit analysis, but in our work, due to limitations on time and computer budget, we had to confine ourselves to regression analysis.

We see in Table B-1 that the mean for the occurrence of work only is close to .5, but that school only and no school, no work occur rarely. We therefore concentrate on the results of the regression of work only. Determinants of whether the child does or does not work only are: sex (girls work without going to school more than boys), birth year (older children are more likely to work without going to school than younger ones), and children from village 6 are very likely to work only, while those from village 14 are likely not to work only. Although not as highly significant, it also appears that affluence, as shown by house type and commercial farming group, contributes negatively to inclusion in the work only group. Mothers' perception of the economic utility of children makes a positive contribution to the child's participation in the work only group.

---

*In the appendix to this chapter, the reader will find other information about school and work activities by village, sex, and age. This material may be useful in interpreting the results summarized here.

TABLE B-1

REGRESSION OF WORK/SCHOOL CHOICES FROM RØ9.
N=515
All Villages and Occupations

VARIABLE	Mean	S.D.	School Only		Work Only		School & Work		No School, No Work	
			Coeff.	T-Stat.	Coeff.	T-Stat.	Coeff.	T-Stat.	Coeff.	T-Stat.
School	.0893	.2855								
Work	.4951	.5005								
School & Work	.2777	.4483								
No School	.1379	.3451								
PLB1	.5308	.3247	.0211	.5201	.1144	1.6442	-.0733	-1.1120	-.0623	-1.2263
PLB2	.0000	.0001	-149.3545	-.9694	-64.6865	-.2461	322.0873	1.2905	-107.8644	-.5606
COM2	.2350	.4244	.0155	.4825	.0117	.2128	-.0345	-.6652	.0674	.1853
COM3	.3476	.4767	.0287	.9692	-.0912	-1.8019	.1205*	2.5448	-.0580	-1.5700
PERUTIL	2.5010	1.7537	-.0012	-.1711	.0217	1.7428	-.0298**	-2.5349	.0094	1.0379
HSTYPE	2.7320	.6096	.0099	.4367	-.0596	-1.5320	.0168	.4563	.0328	1.1570
SEX (m=1, f=2)	1.4913	.5004	-.0617*	-2.5660	.1366***	3.3237	.0026	.0068	-.0776**	-2.5848
Y63	.1049	.3067	.0186	.4469	-.1721*	-2.4194	.2671***	3.9638	-.1135*	-2.1856
Y64	.1223	.3280	.1407***	3.5987	-.2337***	-3.4937	.2079***	3.2819	-.1149*	-2.3533
Y65	.1243	.3302	.0568	1.4627	-.1977**	-2.9771	.1969**	3.1315	-.0560	-1.1553
Y66	.1437	.3511	.1179**	3.1898	-.0796	-1.2584	.0666	1.1128	-.1050*	-2.2741
Y67	.1359	.3450	.0224	.5934	-.1611*	-2.4976	.1455*	2.3811	-.0067	-.1426
V2	.2796	.4492	-.0289	-.8994	.1266*	2.2991	-.0938	-1.7997	-.0038	-.0945
V3	.1961	.3974	.0089	.2306	.0442	.6720	-.0124	-.1990	-.0407	-.8471
V4	.2175	.4129	.2045***	5.9718	-.3391***	-5.7887	.0081	.1462	1.765**	2.9577
FAMSZ	7.4893	1.6984	.0154*	2.1536	.0024	.1958	-.0058	-.5067	-.0119	-1.3356
			R2 = .155		R^2 = .195		R^2 = .101		R^2 = .100	

*$p \leq .05$

**$p \leq .01$

***$p \leq .001$

Source: Rand File R09

We also can look at the school and work results for indications of which children perform both kinds of activities. Here, perceived economic utility has a significant negative coefficient, belonging to the commercial farmers group has a positive coefficient, and birth year and village appear to exert influences that are oppositve in sign for those noted in the work only results above.

Table B-2 shows differences in distributions for children of the four occupational groups. We see that children's activities differ according to whether thier fathers are non-farmers, or subsistence, semi-subsistence, or commercial farmers. Children of non-farmers are least likely to attend school only, and they are most likely to do neither work nor attend school. Of the three groups of farmers, children of commercial farmers are most likely to work, whether in conjunction with attending school or not. We see that semi-subsistence farmers' children tend to work only, and commercial farmers' children are also the most likely ones to attend school and are least likely to do neither work nor school. Earlier, we suggested that this model tends to overstate economic factors because of the absence of health and height variables which are likely to be correlated with affluence. Here, the greater affluence of children of commercial farmers may contribute to their greater ability to both work and go to school, while their less affluent peers may not be robust enough to undertake both simultaneously.

We see here that the more affluent groups, commercial farmers and non-farmers, are likely to use their children's help while the children attend school. Few children of any occupational category go to school without participating in any work activity. Children of non-farmers are more likely than other occupational groups to engage in little work activity, showing that the need for children's help appears to be concentrated in tasks that help farming families.

TABLE B-2

COMPARISON OF MEANS AND STANDARD DEVIATIONS OF VARIABLES USED IN REGRESSIONS OF SCHOOL AND WORK ON FAMILY BACKGROUND AND ECONOMIC VARIABLES, OLDER COHORTS, FOUR VILLAGES

Variable	Children of Non-Farmers N=60		Children of Subsistence Farmers N=178		Children of Semi-Subsistence Farmers N=121		Children of Commercial Farmers N=179	
	Mean	Stan. Dev.	Mean	Stan. Dev.	Mean	Stan. Dev.	Mean	Stan. Dev.
School only	.0167	.1291	.1067	.3907	.1074	.3110	.0782	.2693
Work only	.3667	.4860	.4944	.5014	.5455	.5000	.4804	.5010
School and work	.3833	.4903	.2528	.4358	.1818	.3873	.3631	.4823
No school, no work	.2333	.4265	.1461	.3542	.1653	.3730	.0782	.2693
PLB1	.1438	.3033	.6325	.3277	.6206	.2920	.4614	.2497
PLB2	.0000	.0000	.0000	.0001	.0000	.0000	.0000	.0001
ACTGY	136.1567	294.3516	59.4579	162.0781	68.0934	220.0047	123.2872	227.1168
SALESA	35.0000	66.7802	0	0	63.5207	54.9921	335.6704	417.9128
PERUTIL	2.1833	1.7611	2.1461	1.6945	3.1157	1.5011	2.5642	1.8360
HSTYPE	2.9667	.6630	2.6348	.5276	2.5950	.5565	2.8547	.6370
SEX (m=1, f=2)	1.5667	.4997	1.4888	.5013	1.4545	.5000	1.4972	.5014
Y63	.0667	.2515	.0843	.2786	.1240	.3309	.1173	.3227
Y64	.1667	.3758	.1236	.3300	.1322	.3402	.1117	.3159
Y65	.1000	.3025	.1404	.3484	.0992	.3001	.1341	.3417
Y66	.1167	.3237	.1292	.3364	.1570	.3653	.1620	.3695
Y67	.1333	.3428	.1517	.3597	.1240	.3309	.1229	.3292
V2	.1333	.3428	.2528	.4358	.2810	.4514	.3464	.4771
V3	.2833	.4544	.1292	.3364	.0744	.2635	.3240	.4693
V4	.1833	.3902	.3483	.4778	.2562	.4383	.0782	.2693
FAMSZ	6.0833	1.6395	7.3483	1.3909	7.8264	1.8105	7.6983	1.7477

* $p \leq .05$

** $p \leq .01$

*** $p \leq .001$

Source: R09 Files

Regressions were also done by occupational groups separately. Results of that examination showed that when we control for family economic activities, it is possible to improve the R^2, but there is the same problem with the use of regression analysis as was noted before. Meaningful results appeared for the work only variable and showed that different patterns emerged by family occupational group. Because of the small sample size, there were few significant coefficients, and they were chiefly sex, age, and village.

2.) Results From the Schooling Study

Looking at the schooling study carried out by INCAP in 1978, we have information on school enrollment, achievement, and paid work over a several year period. We break down activities this time into school only, paid work only, school and paid work, and neither school nor paid work. Recall that this information differs from the Rand data since it was collected four years later, reveals several years of activity rather than a single day, but, most important, does not include non-paid work done for the family.

Table B-3 presents regression coefficients for the four types of activities as dichotomous dependent variables. Here, as in (a) above, we approach with caution the regression results, examining those for school only and school and paid work because these activities occur nearly 50% of the time. We see that significant negative coefficients appear in the school only regression for semi-subsistence and commercial farm groups for older children and significant positive coefficients for the mothers' expectations of children's grade attainment. Other less highly significant determinants are sex (girls are more likely than boys to go to school without working for pay) and village (location in village 6 produces a negative coefficient). For school and paid work, we see that positive, highly significant coefficients occur for semi-subsistence and

TABLE B-3

OLS COEFFICIENTS FOR REGRESSION OF YEARS OF SCHOOL ENROLLMENT AND PAID WORK ACTIVITIES OF CHILDREN BORN 1962-1968
From Schooling Study

Variable Name	Mean	Stan. Dev.	School Only N=277		Paid Work Only		Paid Work and School		No School and No Paid Work	
			Coefficient	T-Statistic	Coefficient	T-Statistic	Coefficient	T-Statistic	Coefficient	T-Statistic
School only	.5379	.4995								
Paid work only	.0469	.2119								
School and paid work	.3646	.4822								
No school, no paid work	.0505	.2195								
PLB1	.5011	.3332	-.0739	-.7103	.0908*	2.0003	-.0190	-.1841	.0021	.0427
PLB2	.0000	.0001	-50.2894	-.1602	1.0458	.0076	115.7854	.3725	-66.5418	-.4529
COM2	.2202	.4151	-.1937*	-2.4040	-.0235	-.6677	.1887*	2.3656	.0284	.7522
COM3	.3430	.4756	-.1515*	-2.0921	-.0201	-.6374	.1483*	2.0673	.0234	.6888
PERUTIL	2.4585	1.7721	-.0122	-.6933	-.0033	-.4345	.0158	.9018	-.0002	-.0221
SALALL	289.0903	451.3604	.0001	.7655	.0000	.0112	-.0001	-.7561	-.000001	-.0457
ACTGY	108.4152	261.0222	-.0001	-.4965	-.00001	-.2918	-.00002	-.1838	.0001	1.7213
HSTYPE	2.7329	.6145	-.0528	-.8858	-.0133	-.5123	.0514	.8693	.0148	.5301
SEX (m=1, f=2)	1.4910	.5008	.1090	1.8186	-.0051	-.1939	-.1335*	-2.2492	.0296	1.0549
Y63	.1264	.3328	-.3176**	-3.0435	.0635	1.3944	.2732**	2.6433	.0191	-.3909
Y64	.1444	.3521	-.2069*	-2.1027	.0750	1.7484	.1557	1.5981	-.0239	-.5188
Y65	.1444	.3521	-.1889	-1.9457	.0901*	2.1279	.1213	1.2606	-.0225	-.4940
Y66	.1841	.3883	-.0719	-.7840	.0310	.7738	.0632	.6959	-.0227	-.5185
Y67	.2202	.4151	ne		ne		ne		ne	
Y68	.1805	.3853	.0321	.3435	.0047	.1162	-.0882	-.9523	.0514	1.1726
V2	.1552	.3628	-.1539	-1.6973	.1154**	2.9163	.0767	.8532	-.0381	-.8972
V3	.2852	.4523	-.0834	-1.0011	.0084	.2318	.0826	1.0001	-.0076	-.1949
V4	.1697	.3760	.1250	1.3619	-.0632	-1.5772	-.0055	-.0605	-.0564	-1.3112
Family Size	7.3502	1.7867	.0203	1.2004	-.0061	-.8287	-.0198	-1.1861	.0057	.7177
RWMA	.2527	.4354	.0911	1.2565	-.0392	-1.2387	-.0247	-.3445	-.0272	-.8007
Expected Grades	5.1480	1.5095	.0647**	2.9801	-.0041	-.4287	-.0448*	-2.0852	-.0158	-1.5542
RWPA	.2852	.5597	.0201	.3457	-.0343	-1.3499	.0133	.1966	.0028	.1035
			R^2 = .172		R^2 = .124		R^2 = .128		R^2 = .060	

* $p \leq .05$

** $p \leq .01$

*** $p \leq .001$

ne = not entered in equation

Source: ESCOL and ESCLIS Files

commercial farmers, sex (boys are more likely to have done paid work and also gone to school), and older ages. Grade expectation produces a negative coefficient here. These results show consistent determinants, opposite in sign, for the school only ones, since the occurrence of the other two kinds of activity combinations is of extremely low frequency.

In Table B-4, we have comparisons of means and standard deviations for the same set of variables as in B-3, but disaggregated by economic groups. We see, from the breakdowns of activities, that children of all groups are equally likely to go to school at some time and that children of semi-subsistence and commercial farmers are more likely to work than children of other families. We also see a difference in mothers' perceptions of the economic utility of children; higher perceived utility appears for children of semi-subsistence farmers, while educational aspirations are lowest for subsistence farmers. Measures of economic activity and affluence show that non-farmers and commercial farmers earn more than the other two groups. Family size of non-farmers is lower and literacy of mothers higher.

Regressions were also done by occupational groups separately which tended to raise the R^2 results within each group. Especially noticeable among the findings was the fact that children of subsistence farmers in village 6 are most likely to do paid work (shown by the large positive coefficient and high t-statistic). It is also noticeable that mothers' educational aspirations are clearly related to attendance in school for children of subsistence and semi-subsistence farmers. Other results from these disaggregated regressions are more difficult to interpret due to the small sample size and the infrequency of occurrence of paid work activities.

From the schooling study, we also examined years of schooling in relation to economic and family background that appeared as variables in the other models. Table B-5 shows that significant determinants of number of years

TABLE B-4

COMPARISONS OF MEANS AND STANDARD DEVIATIONS OF VARIABLES USED IN REGRESSIONS OF SCHOOL AND PAID WORK
From Schooling Study by Occupation of Father

Variable Name	Non-Farmers N=43		Subsistence Farmers N=99		Semi-Subsistence Farmers N=61		Commercial Farmer N=95	
	Mean	Stan. Dev.	Mean	Stan. Dev.	Mean	Stan. Dev.	Mean	Stan. Dev.
School only	.5814	.4992	.6364	.4835	.4590	.5025	.4737	.5020
Paid work only	.0698	.2578	.0606	.2398	.0328	.1796	.0316	.1758
School and paid work	.3023	.4647	.2727	.4476	.4426	.5008	.4316	.4979
No school, no paid work	.0465	.2131	.0303	.1723	.0656	.2496	.0632	.2445
PLB1	.2001	.3060	.5956	.3478	.5980	.3128	.4238	.2509
PLB2	.0000	.0000	.0000	.0001	.0000	.0001	.0000	.0001
PERUTIL	2.1628	1.7173	2.3131	1.8386	3.0984	1.5460	2.2947	1.7617
SALALL	453.8605	567.6657	334.0404	494.7766	159.8033	267.1866	330.7579	504.0498
ACTGY	189.4302	418.4296	62.6273	168.8538	81.3459	241.4953	123.8368	229.8465
HSTYPE	2.7442	.6580	2.6869	.5277	2.6230	.5821	2.8526	.6518
SEX (m=1, f=2)	1.4419	.5025	1.5051	.5025	1.4918	.5041	1.4842	.5024
Y63	.0698	.2578	.1111	.3159	.1967	.4008	.0947	.2944
Y64	.2791	.4539	.1313	.3359	.1311	.3404	.1474	.3564
Y65	.0930	.2939	.1717	.3791	.0820	.2766	.1579	.3666
Y66	.1628	.3735	.1818	.3877	.1475	.3576	.2105	.4098
Y67	.2093	.4116	.2121	.4109	.1967	.4008	.2421	.4306
Y68	.1860	.3937	.1919	.3958	.2459	.4342	.1474	.3564
V2	.0930	.2939	.1717	.3791	.1475	.3576	.1579	.3666
V3	.4186	.4992	.1717	.3791	.1148	.3214	.4632	.5013
V4	.1163	.3244	.2828	.4527	.2131	.4219	.0421	.2019
Family Size	5.5581	1.9062	7.3838	1.3455	7.7705	2.0363	7.4632	1.7370
RWMA	.3721	.4891	.1919	.3958	.2623	.4435	.2947	.4583
Educational Asperations	5.4884	1.4699	4.8182	1.6682	5.2623	1.3404	5.2842	1.3889
RWPA	.1163	.9053	.1717	.3791	.2787	.5811	.4632	.5614

Source: ESCOL and ESCLIS Files

TABLE B-5

OLS COEFFICIENTS FOR REGRESSION OF SCHOOL AND PAID WORK ACTIVITIES FOR CHILDREN BORN 1962-1968 FROM SCHOOLING STUDY, 1978

N=277

Variable Name	Mean	Stan. Dev.	Schooling -- Years Enrolled		Paid Work	
			Coefficient	T-Statistic	Coefficient	T-Statistic
ECS22 (Number of Years Enrolled)	3.5704	2.0269				
FOR16 (Work for Pay? 1, 2)	1.4116	.4930				
PLB1	.5011	.3332	-.0883	-.2592	.1008	1.0606
PLB2	.0000	.0001	-205.5630	-.1876	86.5527	.2832
COM2	.2202	.4151	.1022	.3651	.1754*	2.2458
COM3	.3430	.4756	-.0321	-.1266	.1328	1.8755
PERUTIL	2.4585	1.7721	.0221	.3563	.0117	.6760
HSTYPE	2.7329	.6145	.3835	1.9439	.0273	.4962
SEX	1.4910	.5008	-.2359	-1.1209	-.1406*	-2.3952
Y63	.1264	.3328	1.8702***	5.0920	.3396***	3.3151
Y64	.1444	.3521	1.5472***	4.4955	.2243*	2.3372
Y65	.1444	.3521	1.2950***	3.7842	.2109*	2.2094
Y66	.1841	.3883	.7338*	2.2718	.0945	1.0487
Y67	.2202	.4151	0	0	0	0
Y68	.1805	.3853	-.6199	-1.9033	-.0944	-1.0394
V2	.1552	.3628	-.1733	-.5449	.1854*	2.0894
V3	.2852	.4523	.3596	1.2455	.1009	1.2530
V4	.1697	.3760	.8983**	2.8290	-.0553	-.6239
FAMSZ	7.3502	1.7867	-.0241	-.4058	-.0264	-1.5954
RWMA	.2527	.4354	.6592**	2.5812	-.0624	-.8761
FOR13 (mother's grade expectations)	5.1480	1.5095	.3794***	5.0166	-.0510*	-2.4180
RWPA	.2852	.5597	.3373	1.6429	-.0229	-.3991
			R^2 = .371		R^2 = .17	

* $p \leq .05$

** $p \leq .01$

*** $p \leq .001$

Source: ESCOL and ESCLIS Files

enrolled appear to be age, mother's literacy, and educational aspirations. Since the children included in this set of data were ten to fifteen years of age, we treated the whole group as equivalent in terms of the potential length of time available to them to attend school. This is not too damaging an assumption since most school participation appears to be between seven and ten years of age.

Models B and C - School Enrollment Models for the Cohorts Born 1969 to 1971:

In this section, we describe results of two kinds of models that relate the child's characteristics with school participation. In each case, we first looked at all villages and economic groups combined (Model B above) and then disaggregated by village and economic group to test for the effects of child's health and development within each group (Model C). Based on information collected by INCAP in the longitudinal study, we examined school enrollment and achievement of children born between 1969 and 1971. These children, having attained school starting age at seven, permitted us to hypothesize relationships between family work, parental literacy, economic factors, child's health, size, verbal fluency, sex, birth order, age, and village, in relation to the child's school enrollment and achievement. In this analysis, we did not have data on the out-of-school activities of the children.

The results of this examination are shown in several tables. Table B-6 compares the means and standard deviations, and B-7 shows results of the regression of school enrollment on family health, food intake (home diet and supplementation were entered separately), and some family background variables. This was done, first, for all villages combined and then for each village separately. Height at age seven is entered here, but the results show that concurrent nutritional variables provide more powerful

TABLE B-6

MEANS AND STANDARD DEVIATIONS OF VARIABLES FOR REGRESSIONS OF SCHOOL ENROLLMENT ON CHILD NUTRITION, HEALTH, SIZE, AND FAMILY BACKGROUND VARIABLES
Longitudinal Cohorts Born 1969-1971
By Village and All Villages Combined

Village	All Villages		Village 3		Village 6		Village 8		Village 14	
	Mean	S.D.	Mean	S.D.	Mean	S.D.	Mean	S.D.	Mean	S.D.
(N)	(138)		(33)		(41)		(30)		(34)	
ENROLL	.623	.486	.818	.392	.317	.471	.733	.450	.706	.463
SEX (f=0, m=1)	.486	.502	.515	.508	.512	.506	.467	.507	.441	.504
YEAR	.094	.293	.12	.33	.098	.304	.033	.183	.118	.327
RWMA	.261	.441	.394	.496	.220	.419	.333	.479	.118	.327
RWPA	.399	.491	.546	.506	.317	.471	.533	.507	.235	.431
VGRA$	296.9	420.4	211.17	288.6	364.4	464.6	459.0	597.1	155.9	145.5
SALALL	180.9	280.3	250.0	385.0	229.7	303.5	106.4	126.7	120.8	200.5
ACTGY	75.3	185.2	75.4	217.4	67.9	198.4	79.4	134.3	80.6	181.3
CONSUMP	.073	.938	.730	.655	-.158	.981	-.502	.768	.223	.857
ORDR6	.608	.160	.609	.164	.636	.171	.605	.167	.576	.136
FMSZ6	6.94	1.79	6.97	1.94	7.19	1.89	6.533	1.93	6.97	1.34
HTC16	104.9	4.36	104.0	3.99	105.7	4.72	104.1	3.92	105.4	4.55
CHSPP14	20.82	30.64	0	0	42.7	35.2	0	0	33.0	26.8
ACHHDP5	134.88	33.91	142.2	29.0	120.5	30.9	146.5	34.5	134.8	35.8
CHSPC14	92.33	68.6	80.06	49.9	111.9	92.3	85.57	36.4	86.54	70.54
ACHHDC5	1047.0	288.98	1042.2	247.5	1007.7	288.8	1234.78	305.09	933.46	238.9
CHHLTH6	.695	.336	.645	.344	.630	.353	.722	.280	.717	.318
CHDIAR16	.009	.028	.0074	.014	.0079	.016	.011	.022	.011	.048
RECOG4	33.67	3.82	34.03	3.84	34.34	3.38	32.2	3.46	33.82	4.37
NAMING4	21.09	4.71	21.73	5.17	21.93	4.17	18.73	3.66	21.53	5.16

TABLE B-7

COEFFICIENTS FOR REGRESSIONS OF SCHOOL ENROLLMENT ON CHILD NUTRITION, HEALTH, SIZE, FAMILY BACKGROUND VARIABLES
Longitudinal Cohorts Born 1969-1971
By Village and All Villages Combined

Variable	All Village		Village 3		Village 6		Village 8		Village 14	
	Coeff.	T-Stat.	Coeff.	T-Stat.	Coeff.	T-Stat.	Coeff.	T-Stat.	Coeff.	T-Stat.
SEX (f=0,m=1)	.147	1.90	.088	.456	.181	.934	.095	.354	.629*	2.59
YEAR	-.491**	-3.98	-.229	-.779	-.655	-1.89	-.770	-1.27	-.001	-.003
RWMA	.0114	.124	-.044	-.250	.341	1.50	.045	.118	.018	.050
RWPA	.077	.999	.243	1.44	.087	.465	.201	.848	.148	.539
VGRA$	-.00001	-.790	.0002	.054	-.00002	-.094	-.0003	-1.16	-.001	-1.62
SALALL	-.00001	-.3128	-.0002	-.906	-.0002	-.480	-.0006	-.591	.00004	.099
ACTGY	-.0001	-.626	.0003	.391	-.001	-2.02	.0009	.974	.0005	.782
CONSUMP	.157*	2.74	.080	.443	.336*	2.79	-.022	-.101	-.0148	-.099
ORDR6	-.273	-.981	-1.29	-2.06	-.204	-.350	-.278	-.394	.371	.357
FMSZ6	.029	1.135	.062	.904	.020	.310	.016	.211	.021	.251
HTC16	.009	1.07	.0055	.239	-.002	-.091	.022	.580	.001	.061
CHSPP14	--	--	--	--	.0001	.047	--	--	not entered	
ACHHDP5	.0022	1.916	.007	1.62	.007	1.19	.0011	.147	.004	.702
CHSPC14	.0044	1.229	-.002	-1.22	not entered		-.003	-.942	.004	1.69
ACHHDC5	-.0002	-.973	-.001	-1.08	-.001	-1.70	.00002	.024	-.0002	-.308
CHHLTH6	.199	1.65	-.165	-.666	-.295	-.842	.343	.684	.641	1.60
CHDIAR16	--	--	-6.45	-.542	-3.38	-.620	.156	.039	-1.743	-.748
RECOG4	.015	1.07	.048	1.63	.005	.131	.045	-1.169	.019	.525
NAMING4	.004	.318	.034	1.21	.024	.633	-.002	-.060	-.026	-.902

*p ≤ .05 R^2 = .367 R^2 = .489 R^2 = .635 R^2 = .613

**p ≤ .01

explanations than height when both sets of variables are included. For all villages combined, we see that, of the 138 eligible children, approximately 60% enrolled in school. Great variation between villages, as noticed before, appears again, with village 6 having a very low school participation rate. Regression analysis here shows that age is closely tied to enrollment (children of ages 8 and 9 are more likely to participate in school than seven year olds), that affluence affects enrollment positively, as does home diet protein, and that child's proportion of the days healthy affects enrollment significantly and positively. Boys are more likely to enroll in school than girls.

Disaggregating by village, we see that determinants of enrollment differ by village. In village 6, affluence of family appears to have a significant positive effect on enrollment, while amount of economic activity of the family has a significant negative effect. In the other villages, these coefficients are not statistically significant. Village 3 shows a positive effect of home diet protein as well as a negative effect for the proportion of older siblings in the subject's family.

Table B-8 shows results of regressing school enrollment on height, health, and the same family background factors as above in a stepwise regression. Here, we see from the series of related models which involve an increasing number of explanatory variables how economic factors, then parental literacy, and, finally, child development variables appear to affect enrollment. These results provide strong evidence, controlling for economic factors and other characteristics of the family, that the child's height, verbal development at age seven, and health appear to affect school participation significantly.

Table B-9 shows the results of regressions done by village. Since village 6 has consistently shown low school participation and high involvement

TABLE B-8

OLS Coefficients for Stepwise Regressions for Enroll

N=184

Levels of Significance in ( )

Variable	Mean	S.D.	Step 1		Step 2		Step 3		Step 4	
Enroll	.6467	.4793								
Sex	.5217	.5009	.1508**	(.014)	.1497***	(.013)	.1562**	(.009)	.1513**	(.011)
Y69	.4837	.5011	.4478**	(.000)	.4527***	(.000)	.4494***	(.000)	.4460***	(.000)
Y70	.4347	.4971	.3890***	(.001)	.3696***	(.001)	.3738***	(.001)	.3718***	(.001)
V2	.2880	.4541	-.2893***	(.001)	-.2780**	(.002)	-.3047***	(.001)	-.3665***	(.000)
V3	.2337	.4243	.2038*	(.036)	.1811	(.059)	.1764	(.062)	.1112	(.248)
V4	.2120	.4098	-.0163	(.857)	.0174	(.847)	-.0012	(.989)	-.0334	(.706)
VGRA$	347.1641	623.4997	-.00002	(.966)	-.00002	(.977)	-.0001	(.817)	-.00001	(.880)
ORDR6	.5973	.1576	-.5311*	(.026)	-.5537*	(.019)	-.5260*	(.024)	-.4907*	(.032)
FMSZ6	6.9402	1.7561	.0155	(.454)	.0230	(.268)	.0212	(.299)	.0282	(.168)
SALALL	196.6141	299.8552	.0002	(.147)	.0002	(.103)	.0002	(.104)	-.0002	(.106)
ACTGY	67.1647	174.7513	.00001	(.958)	.00002	(.882)	-.00001	(.974)	-.0001	(.755)
PLB1	.6179	.3380	.0442	(.668)	.0717	(.487)	.0566	(.577)	.0356	(.724)
PERUTIL	2.4022	1.7806	-.0324	(.069)	-.0342	(.052)*	-.0344*	(.047)	-.0317	(.064)
CONSUMP	.0343	.9991	.1663***	(.000)	.1510***	(.000)	.1298***	(.001)	.0963*	(.018)
RWMA	.2391	.4277			.0567	(.448)	.0399	(.588)	.0388	(.594)
RWPA	.4022	.4917			.1438*	(.029)	.1450*	(.025)	.1325*	(.038)
RECOG4	33.7609	4.0065					.0195**	(.011)	.0154*	(.048)
HTC16	104.8109	4.3869							.0150*	(.043)
CHHLTH6	.6685	.3234							.1728	(.077)
			R^2=.359		R^2=.384		R^2=.407		R^2=.431	

* $p \leq .05$

** $p \leq .01$

*** $p \leq .001$

TABLE B-9

OLS COEFFICIENTS FOR REGRESSION OF ENROLL WITH VILLAGE 6 AND VILLAGES 3, 8, 14

Variable	Village 6 (N=53)				Villages 3, 8, 14 (N=131)			
	Mean	S.D.	Coeff.	Level of Signif.	Mean	S.D.	Coeff.	Level of Signif.
ENROLL	.3396	.4781	---	---	.7710	.4218	----	---
SEX	.5660	.5004	.1107	(.412)	.5038	.5019	.1835**	(.005)
Y69	.4906	.5047	.4681	(.068)	.4809	.5016	.4648***	(.000)
Y70	.4151	.4975	.2515	(.330)	.4427	.4986	.4292**	(.001)
RWMA	.2075	.4094	.2248	(.195)	.2519	.4358	.0435	(.610)
RWPA	.3396	.4781	.2164	(.151)	.4275	.4966	.1590*	(.035)
VGRA$	541.3140	990.9916	-.00001	(.933)	268.6149	364.4617	.0001	(.585)
ORDR6	.6136	.1627	-.6454	(.161)	.5907	.1556	-.3280	(.232)
FMSZ6	7.0377	1.7316	.0209	(.650)	6.9008	1.7709	.0351	(.136)
SALALL	246.7358	326.7757	-.00004	(.848)	176.3359	287.0863	.0001	(.264)
ACTGY	52.4981	176.3489	-.0007	(.082)	73.0985	174.4287	.0002	(.238)
PLB1	.6409	.3246	-.2990	(.264)	.6086	.3441	.1523	(.180)
PERUTIL	3.0566	1.4730	.0007	(.990)	2.1374	1.8304	-.0315	(.086)
HTC16	106.3189	4.0927	-.0078	(.634)	104.2008	4.3691	.0255**	(.002)
CHHLTH6	.6052	.3567	.2025	(.321)	.6941	.3066	.0435	(.706)
RECOG4	34.2830	3.5701	.0004	(.985)	33.5496	4.1643	.0167*	(.049)
CONSUMP	-.2025	1.0398	.1730*	(.041)	.1301	.9698	-.0062	(.880)

R^2 = .493 R^2 = .360

* $p \leq .05$

** $p \leq .01$

*** $p \leq .001$

of children in the work force, our breakdown was done by village 6 compared to villages 3, 8, and 14 combined. Different patterns emerge as explainers of school enrollment. In village 6, we see that affluence has a significant positive effect on school attendance, while the proportion of older siblings and amount of economic activity are negatively and significantly related to enrollment. This is consistent with the results found when nutritional data were included above. In villages 3, 8, and 14, boys, older children, more verbal children, and taller children are more likely to be in school.

Table B-10 gives means and standard deviations for variables used in the regressions of school enrollment divided by fathers' occupational groups. Here, we see differences in the proportion of children participating in school by family occupation. Non-farming families in the longitudinal sample had to be excluded because of low sample size, but, for the others, in Table B-11, giving the results of regressions, we see that patterns differ by level of economic activity and land holings. Again, we see that village 6 is a significant negative influence on school enrollment, and village 14 has a positive influence. Among semi-subsistence farmers, the child with fewer older siblings is more likely to go to school, and the taller, more verbal child is also more likely to go to school. More affluent families also tend to have increased positive effect on school enrollment.

A final table, B-12, shows the results of regressing school achievement measures on the same set of variables as before. School achievement was measured in terms of school promotion rates and scores in language and mathematics which were assigned by teachers at the completion of the school year. Few significant explainers appear here. The factors that explained participation in school did not appear to determine school

TABLE B-10

COMPARISONS OF MEANS AND STANDARD DEVIATIONS OF VARIABLES USED IN REGRESSIONS OF CHILDREN'S SCHOOL ENROLLMENT ON SIZE, HEALTH, FAMILY ECONOMIC VARIABLES
Children Born 1969-1971

Variable	Non-farmers		Subsistence Farmer		Semi-Subsistence		Commercial	
	Mean	S.D.	Mean	S.D.	Mean	S.D.	Mean	S.D.
	N=12		N=65		N=47		N=65	
ENROLL	.8333	.3892	.5385	.5024	.6596	.4790	.7231	.4510
SEX	.3333	.4924	.5385	.5024	.5745	.4998	.5231	.5034
Y69	.5000	.5224	.4308	.4990	.5745	.4998	.4615	.5224
Y70	.2500	.4523	.4923	.5038	.3404	.4790	.4615	.5024
V2	.1667	.3892	.2462	.4341	.3191	.4712	.3385	.4769
V3	.2500	.4523	.1846	.3910	.0851	.2821	.3846	.4903
V4	.1667	.3892	.3538	.4819	.2128	.4137	.0769	.2685
RWMA	.6667	.4924	.1692	.3779	.2766	.4522	.2462	.4341
RWPA	.5833	.5149	.2308	.4246	.4468	.5025	.5231	.5034
VGRA$	208.9806	229.5051	91.4631	75.3192	196.5499	113.1571	746.2240	916.8300
ORDR6	.7017	.1563	.5800	.1476	.5808	.1503	.6120	.1706
FMSZ6	6.5833	1.3790	6.8308	1.6160	6.8723	1.9849	7.1231	1.7545
SALALL	603.8333	507.0945	196.0769	299.1494	142.8085	176.5176	213.3692	348.3545
ACTGY	55.2000	172.0033	53.7846	157.4257	62.6723	183.2678	80.8354	182.3494
PLB1	.6063	.4482	.9608	.1519	1.0000	0	.5894	.2949
PERUTIL	1.7500	1.7645	2.1077	1.8466	2.9149	1.5719	2.4154	1.7931
HTC16	106.4417	3.4516	104.1092	4.8542	104.9362	4.4806	105.2877	3.8158
CHHLTH6	.5903	.3292	.6652	.3321	.6622	.3095	.6792	.3334
NAMING4	23.7500	3.8168	19.5538	5.0468	22.7447	4.4841	21.0308	4.0349
RECOG4	35.7500	3.0488	32.6154	4.2489	34.7872	3.9379	34.0308	3.6656
CONSUMP	.9338	.9088	-.2221	.9392	-.0150	.9772	.2394	.9913

OLS COEFFICIENTS FOR REGRESSION OF CHILDREN'S SCHOOL ENROLLMENT ON FAMILY ECONOMIC AND BACKGROUND VARIABLE AND SIZE AND HEALTH BY FAMILY OCCUPATION
Cohorts Born 1969-1971

Variable	Subsistence Farmers		Semi-Subsistence Farmers		Commercial Farmers	
	Coeff.	Level of Signif.	Coeff.	Level of Signif.	Coeff.	Level of Signif.
ENROLL	---	---	---	---	---	---
SEX(1)	.1861	(.180)	.3482**	(.010)	.1879	(.129)
Y69	.4447	(.063)	.1541	(.455)	.7118***	(.001)
Y70	.3659	(.120)	.1273	(.582)	.6152**	(.005)
V2	-.4043	(.077)	-.2714	(.118)	-.3719*	(.033)
V3	.2271	(.313)	1.1003**	(.003)	-.1084	(.544)
V4	.1233	(.487)	-.1977	(.237)	.0672	(.794)
RWMA	-.1189	(.503)	.1380	(.318)	.0057	(.964)
RWPA	.1328	(.416)	.1499	(.242)	.1498	(.196)
VGRA$	.0001	(.559)	.0001	(.831)	-.00004	(.417)
ORDR6	-.1215	(.821)	-1.3199**	(.009)	-.5306	(.279)
FMSZ6	.0347	(.463)	.0537	(.151)	.0441	(.299)
SALALL	.0003	(.198)	-.0001	(.849)	.0001	(.475)
ACTGY	.0001	(.881)	-.0003	(.429)	-.0004	(.191)
PLB1	-.5468	(.200)	not entered		-.2426	(.291)
PERUTIL	-.0560	(.099)	-.0045	(.922)	.0102	(.766)
HTC16	.0813	(.561)	.0287*	(.050)	.0065	(.717)
CHHLTH6	.2031	(.306)	-.1774	(.373)	.2100	(.236)
RECOG4	.0109	(.501)	.0401	(.026)	.0243	(.149)
CONSUMP	.0943	(.355)	.2391**	(.013)	.0323	(.638)
	R^2 = .497		R^2 = .741		R^2 = .482	

(1) Male = 1
Female = 0

* $p \leq .05$

** $p \leq .01$

*** $p \leq .001$

TABLE B-12

REGRESSION RESULTS FOR ACHIEVEMENT ASSUMING ENROLL=1
N=119

Variable	PROMOTION RATE				AV. MATH SCORE		AV. LANG. SCORE	
	Mean	S.D.	Coeff.	T-Stat.	Coeff.	T-stat.	Coeff.	T-stat.
PROMOTION RATE	.5421	.4249	---	---	---	---	---	---
AVERAGE MATH	60.1244	17.6321	---	---	---	---	---	---
AVERAGE LANG.	60.7008	18.5801	---	---	---	---	---	---
SEX	.5630	.4981	-.1251	-1.6182	-4.0895	-1.2956	-4.8213	-1.4949
Y69	.5042	.5021	.0584	.2670	2.6058	.2925	-1.0851	-.1192
Y70	.4622	.5007	.0692	.3194	.2157	.0244	-1.3326	-.1474
V2	.1513	.3598	-.3336*	-2.5595	-12.3837*	-2.3267	-18.0132***	-3.3122
V3	.2857	.4537	-.1693	-1.3893	-5.0128	-1.0072	-5.7285	-1.1265
V4	.2353	.4260	-.2660*	-2.4660	-14.4254***	-3.2748	-8.9604*	-1.9915
RWMA	.2857	.4537	-.0603	-.6668	4.9890	1.3513	4.0111	1.0633
RWPA	.4790	.5017	-.0378	-.4507	-2.4349	-.7101	-2.6753	-.7636
VGRA$	330.3182	439.9829	.0002	1.8330	.0075	1.7974	.0091*	2.1344
ORDR6	.5838	.1630	.0539	.1666	4.4937	.3399	2.5477	.1886
FMSZ6	6.9076	1.8226	-.0461	-1.5540	-1.6038	-1.3233	-2.0784	-1.6783
SALALL	211.7563	336.6810	-.0001	-.7978	-.0020	-.3734	-.0028	-.5072
ACTGY	77.4840	181.3413	.0000	.1538	.0076	.8309	.0068	.7266
PLB1	.6063	.3499	-.0973	-.7660	3.9463	.7610	1.5253	.2879
PERUTIL	2.1597	1.8411	-.0154	-.7099	-.1231	.1390	-.0642	-.0710
HTC16	105.3008	3.9990	.0112	1.0704	.5745	1.3455	.2084	.4776
OLDSIB	.5966	.8667	-.0030	-.0466	-.6231	-.2337	1.1628	.4268
CHHCTH6	.7198	.2972	.0138	.1010	-4.1278	-.7375	-5.4443	-.9520
RECOG4	34.2941	3.8628	.0268**	2.6413	1.1018**	2.6546	1.4947***	3.5246
CONSUMP	.2412	.9452	.0588	1.0148	3.1250	1.3217	3.8264	1.5839
			R^2 = .280		R^2 = .303		R^2 = .344	

performance once the child attended. Only verbal development appears to be of positive significance in determining achievement. These results are consistent with those obtained by Alan Wilson in Part II-A of this report.

Summary of Results

We see that school enrollment appears to be affected strongly by the economic conditions of the family. When we control for economic factors by separating the subjects by family occupation or by village, thereby controlling for economic differences, we find that children's health, size, and verbal development influence significantly and positively their school participation.

We also find that school participation differs by village which apparently reflects work opportunities for children and attitudes about the value of schooling by parents. In the village where work is readily available, school participation is far lower than in the other villages.

We find that there are differences in school participation by sex, schooling for girls is less frequent than for boys, and their work in the household apparently more valued. The perceived value of literacy for girls is apparently lower than for boys.

We see that size and health of children, when all other economic and family background factors are kept constant, are significant determinants of children's school attendance and performance. Thus, prior nutrition appears to make a difference in height, and height, along with present health and verbal development, makes a difference in school attendance and achievement. The sign of effects of nutritional status on school participation appears to be positive, even with the competing needs for the children's work.

BIBLIOGRAPHY

(1) Alexander and Simmons. The Determinants of School Achievement in Developing Countries. World Bank Working Paper, #201, 1975.

(2) Christiansen, N.; Vuori, L.; Mora, J. O.; and Wagner, M. "Social Environments As It Relates to Malnutrition and Mental Development," in Early Malnutrition and Mental Development, Proceedings of a Symposium of the Swedish Nutrition Foundation. Cravioto, J; Hambraeus, L.; and Vahlquist, B., eds. Uppsala, 1974.

(3) McKay, H. E.; McKay, A. C.; and Sinisterra, L. "Behavioral Effects of Nutritional Recuperation and Programmed Stimulation of Moderately Malnourished Pre-school Age Children." Paper presented at the Meeting of the American Association for the Advancement of Science Symposium. Boston, Mass., 1969.

(4) Mora, J. O.; Amezquita, A.; Castro, L.; Christiansen, J.; Clement-Murphy, J.; Cobos, L. F.; Cremer, H. D.; Dragastin, S.; Elias, M. F.; Franklin, D.; Herrera, M. G.; Ortiz, N.; Pardo, F.; Paredes, B. de; Ramos, C.; Riley, R.; Rodriguez, H.; Vuori-Christiansen, L.; Wagner, M.; and Stare, F. J. "Nutrition and Social Factors Related to Intellectual Performance." World Review of Nutrition and Dietetics 19 (1974): 205-236.

(5) Sharma, R. C., and Sapra, C. L. Wastage and Stagnation in Primary and Middle Schools in India. New Delhi: National Council of Educational Research and Training, 1969.

(6a) Selowsky, Marcelo. "Nutrition, Health and Education: The Economic Significance of Complementarities at Early Ages." Unpublished paper for forthcoming conference.

(6b) Selowsky, Marcelo. The Economic Dimensions of Malnutrition in Young Children. The World Bank, 1978.

(7) Popkin, Barry, and Lim-Ybanes. "Nutrition and School Achievement." To be published in Social Science and Medicine.

(8) Shortlidge, Richard. "A Socioeconomic Model of School Attendance in Rural India." Cornell University.

Appendix to Section II-B

The tables included in this appendix may be useful for the reader who is interested in more information on school and work activities. Although numbers of cases differ between tables, they were collected on related samples.

Appendix B

TABLE 1

SCHOOL PARTICIPATION RATES BY VILLAGE AND SEX FOR CHILDREN BORN BETWEEN 1962 AND 1971

Number of children completing at least 1 year of school (and percentage of children completing a year from those eligible)

Village	Boys		Girls		All Children	
		%		%		%
3	129	61%	113	55%	242	58%
6	84	35%	106	50%	190	42%
8	89	66%	80	55%	169	60%
14	75	52%	71	46%	146	49%
All	377	52%	370	51%	747	52%

TABLE 2

FOR16

DOES CHILD EARN MONEY YET?

(For Children Born 1962 - 1968)
(1=no, 2=yes, 9=no information)

	Village							
	3		6		8		14	
	m	f	m	f	m	f	m	f
NO	34 54%	44 66%	24 32%	36 47%	25 50%	29 54%	23 56%	35 81%
YES	29 46%	23 34%	51 68%	41 53%	25 50%	25 46%	18 44%	8 19%

TABLE 3

Average Earnings by Village in Quetzales X 100 Per Day

For Children, Ages 10 to 15

Type of work	Village									
	3		6		8		14		All	
0									0	
1	50								50	
2					20				20	
3	145		152		98		122		134	
4										
5	105		152		75		50		110	
6	150		275		0		0		212	
7	106		60		79		121		95	
All	m	f	m	f	m	f	m	f	m	f
	145	96	147	142	93	94	129	96	133	115

TABLE 4

Daily Earnings of Children Born 1962 - 1968
For Whole Population. All 4 Villages Combined.

0 = No work
1 = rope, twine making
2 = raffia
3 = agricultural work
4 = peels roots
5 = sells
6 = works in factory
7 = other
9 = no information

Average Earnings per Day
All Ages

Activity	M	F	All	# of Cases
0	0	0	0	242
1	50	50	50	1
2	0	20	20	2
3	129	142	133	156
4	0	0	0	0
5			110	18
6			212	4
7	162	75	95	40
9	--	--	--	251

TABLE 5

SUMMARY OF ACTIVITIES from RØ9

Variable	Description	Boys	Girls	Villages
ACRONYM				
BFMFOOD	Prepared food for family	on very few	✓	All
BCARE	Took care of children	very few	✓✓	All
BSHOP	Went shopping	✓	✓✓	All but few boys except in Village 6
BWATER	Carried water	few	✓✓	Few boys in Village 6
BWOOD	Carried wood	✓✓	Fewer girls	
BLAUND	Did laundry	0	✓	Not in #14
BFOOD	Carried meals	few	some	Mostly Village 6
BFMWK	Did farm work	✓✓	many in Village 6, otherwise few	In #6 and #8 girls, otherwise boys in all
BPROD	Sold products	0	few	Very few
BPDWK	Worked away from house for pay	few	few	Few in #3, 8
BWKHM	Worked within household	few	few	Few

Table continued on next page

SUMMARY OF ACTIVITIES from RØ9 (cont'd.)

Variable	Description	Boys	Girls	Villages
BOTWK	Other activities inside house	few	✓	Few altogether
BRELAT	Visited relatives and other activities outside house	✓	✓	All
BSCH	Schools			
BCLIN	Clinic			
BNITE	Nights away from house			

TABLE 6

Village Comparisons

Years of School Passed

Years	Village	3		6		8		14		All	
		#	%	#	%	#	%	#	%	#	%
	0	144	72.7%	187	90.8%	107	72.8%	140	84.8%	578	80.7%
	1	40	20.2%	9	9.4%	30	20.4%	20	12.1%	99	13.8%
	2	12	6.1%	7	3.4%	8	5.4%	5	3.0%	32	4.6%
	3	2	1.0%	3	1.5%	2	1.4%	0	0%	7	1.0%
N		198		206		147		165		716	

$$\text{Promotion rate} = \frac{\text{years passed}}{\text{years exposed}}$$

Village	3		6		8		14		All	
	m	f	m	f	m	f	m	f	m	f
Rate	.84	.82	.84	.83	.76	.79	.73	.76	.80	.80

YEARS OF SCHOOL EXPOSED

Years	Village	3		6		8		14		All	
		#	%	#	%	#	%	#	%	#	%
	0	130	65.7%	170	82.5%	90	61.2%	121	73.3%	511	71.4%
	1	42	21.2%	18	8.7%	30	20.4%	29	17.6%	119	16.6%
	2	21	10.6%	13	6.3%	18	12.2%	13	7.9%	65	9.1%
	3	5	2.5%	5	2.4%	8	5.4%	2	1.2%	20	2.8%
	5	0	0%	0	0%	1	.7%	0	0%	1	.1%
N		198		206		147		165		716	

TABLE 7

VILLAGE COMPARISONS

Mother's Educational Aspirations for Children's Grade Attainment
EDASP

Village	3		6		8		14		All	
		%		%		%		%		%
EDASP 1	5	7.8%	3	4.7%	0	0%	1	1.9%	9	3.9%
2	7	10.9%	13	20.3%	12	24.0%	5	9.6%	37	16.1%
3	2	3.1%	13	20.3%	2	4.0%	6	11.5%	23	10.0%
4	4	6.3%	8	12.5%	1	2.0%	8	15.4%	21	9.1%
5	39	60.9%	25	39.1%	29	58.0%	28	53.8%	121	52.6%
6	7	10.9%	2	3.1%	6	12.0%	4	7.7%	19	8.3%
#	64		64		50		52		230	

Missing = 594

Section II-C: EDUCATION AND AGRICULTURE EFFICIENCY

by Maria E. Freire[1]

A recurrent theme in this report has been that the relationships between agricultural production, education, and nutrition can be examined from two points of view. On the one hand, economic factors affect nutritional intake and educational outcomes; agricultural production, income, and wealth determine the amount of food available to the household and, in combination with health and socio-environmental factors, affect the child's capacity to learn. On the other hand, educational and nutritional conditions affect economic performance since (a) more educated farmers are likely to be better managers, more receptive to innovation, and more productive, and (b) better nourished people, being more energetic, are also likely to be more productive.

The determinants of child's nutrition and health status as well as the influence of nutrition and health status on learning and educational performance was the subject of Sections A and B. It is our task now to investigate whether enhanced educational attainments of farmers will be reflected in increased agricultural production. Because appropriate information is lacking, links between nutritional status and productivity of farming labor can only be tested indirectly.

---

[1]This report is a selective summary of findings presented elsewhere at greater length: See: M. Freire, "Assessing the Role of Education in Rural Guatemala: The Case of Farm Efficiency," (Unpublished doctoral dissertation, Department of Economics, University of California--Berkeley, 1979). I would like to thank Drs. Lawrence Lau and Dean Jamison for making available to me some of their work which had not yet been published. See: D. T. Jamison and L. J. Lau, Farmer Education and Farm Efficiency, (John Hopkins University Press, Baltimore), forthcoming, and M. Lockeed, D. Jamison, and L. J. Lau, "Farmer Education and Farmer Efficiency -- A Survey," Economic Development and Cultural Change, forthcoming. The first study was especially useful in preparing the first chapter of the above mentioned dissertation.

In this section, we seek to close the intergenerational loop as follows: if it can be shown (as it was in sections A and B) that improvements in children's nutritional intake improves their educational performance, then, as they become adults, they may be more productive and, in turn, better able to provide good nutritional conditions for their own children. To do this we therefore need to investigate the link between adult educational attainment and agricultural productivity.

Framework

The present analysis focuses on 510 family farmers in the four villages under study for which farm management data were available. Data were collected by the RAND Rockefeller Guatemala Project in 1975 and contain cross-sectional information for farming activities carried out during one whole year -- 1974. Data collection was made at a single point in time, relying on the recollection of the farmers answering the questionnaire and checked by a second interview some time later.

Basic to the analysis is the assumption that education, or other factors associated with it[2], influences farm production and farm management by means of enhancing farmers' decision making process. Education (formal education and/or literacy) is supposed to convey information or to improve the ease of gathering and using information which, in turn, will help farmers in making decisions both in every day life and at specific points in time. According to this rationale, educated farmers are supposed

---

[2]The identification of which factors are responsible for the educational effects is quite an unsolved problem in the current literature (Griliches 1977). Besides the economic conditions which, in general, are positively correlated with education but for which one is able to control, other personal characteristics also correlated with education but difficult to measure (drive, personal motivation, energy, and so on) may be the ones which explain the consistently found positive relationship between education and economic performance in rural management.

to be more able to organize production activities in order to maximize their objective functions (usually production and/or profit level) subject to constraints in fixed factors. Testing for the effects of education amounts then to testing for i) whether educated farmers have the same objective functions as uneducated farmers, and ii) whether educated farmers are more able to maximize that same objective function.

The typical pattern of traditional tasks in rural sectors does not help our analysis of the effect of education on performance. Current literature in economic development has emphasized the optimality of current practices in traditional agricultural sectors (Reynolds 1975) reflecting the long process of peasants learning how to produce as much as possible within the constraints of traditional technologies and limited physical resources. In such cases, education would have little to do with farm efficiency since both educated and less educated farmers would be using the same traditional practices and techniques. However, whenever change occurs (due, for instance, to the introduction of new practices, new crops, fertilizers, etc.), breaking the secular and traditional ways of cultivation, people are likely to react in different ways, educated farmers being likely to accept the new practices with less resistance than their uneducated neighbors. Consequently, it is in environments where change in farming do occur that the influence of education on farmer's technical and managerial skills can be more easily detected (Schultz 1975).

In fact, the effect of education on agriculture production in developing countries seems today to be well accepted (M. Lockeed, J. Jamison, and L. Lau 1978), especially under conditions of the introduction of new fertilizers, new crops, and new seeds. Although only time-series studies can show in detail the process of adoption of innovation and the

consequent diffusion process, cross-sectional analyses performed for many countries show the definitive advantage that educated people have in using chemical inputs and introducing mechanization.[3]

In the present case, simple production indices computed for the Guatemalan farmers show that educated (literate) farmers (1) obtain, on the average, higher land and labor productivity levels, (2) operate at a larger scale of operation, and (3) tend to use different factor input combinations than illiterate farmers. Whether this reflects higher managerial and technical skills of educated farmers rather than differences in constraints faced by these farmers entering the market, and/or in goals, is difficult to assess a priori. Differences in scale of enterprise can reflect (1) differences in constraints -- if illiterate, farmers face higher costs for some inputs, the optimum size of the operation will be smaller than for literate farmers who have access to cheaper inputs -- or (2) differences in technical and managerial skills -- with constant returns to scale, better managers, having more ability (in decisions about factor inputs), will operate at a larger scale, and firm size will be determined by the owner's knowledge. It is, however, impossible to assess whether differences in size of enterprise reflect differences in farmer's goals; if illiterate farmers are not profit maximizers, they will not have an incentive to operate at the profit maximization level, and their firm size will be below or above the profit maximizing size. Thus, from the analysis of the size of the firm, one cannot make inferences about the differences in

---

[3]The effect of education on adoption of innovation stresses the aspect of education as a vehicle for collecting and using information. Given the importance that risk and uncertainty play in acceptance of new practices, especially among poor farmers, and the likely correlation between education and inherited wealth, the detected influence of education on adoption of change can reflect the situation of better off farmers more likely to innovate unless one is able to control for the interactions between basic wealth and education.

profit maximizing behavior among farmers. As for factor input combinations, the elements of Table C-2 indicate that illiterate farmers face higher prices both for land and input. Higher land cost, for example, and profit maximization imply that the land-labor ratio will be lower for illiterate than for literate farmers. However, neither lack of managerial ability nor the absence of profit maximizing behavior suggests anything with respect to such input ratios. A management error could cause input ratios to be wrong in either direction. If one considers that there is an average ratio corresponding to the profit maximizing behavior, that group of farmers with the smaller variance in input combinations will be the one which commits fewer errors in farming management. In our case, significant differences in variances did not appear among groups. However, if farmers commit similar management errors, this does not imply that their goals and objective functions in terms of profit maximization have to be the same; differences in scale and input combinations may be the result of managerial and technical skill differences as well as differences in constraints faced by farmers on entering the market alone. We shall develop this idea below.

The Data and Some Preliminary Analysis

The collected information for the Guatemalan farmers is fairly complete. Table C-1 and C-2 provide some descriptive statistics for two groups of farmers: illiterate and literate farmers. The reason for this disaggregation is that in a society where the average educational level is very low (a 35 percent literacy rate and less than one grade of schooling), the most important educational indicator that might affect the decision on farming procedures is the capacity to read and write rather than the number of years of school attended or grades achieved. Splitting the sample into these two groups made the analysis much easier for comparisons of the structural differences among farmers who follow different production practices.

TABLE C-1

Descriptive Statistics

	FULL SAMPLE		ILLITERATE FARMERS		LITERATE FARMERS	
	Mean	Stan. Dev.	Mean	Stan. Dev.	Mean	Stan. Dev.
Number of Cases	510	---	323	---	187	---
Land (Cuerdas)[a]	79.9	118.4	78.3	107.5	82.8	136.2
Land (Quetzales Value)	77.2	77.7	69.1	66.9	91.9	93.3
Rented Land/ Planted Land	.368	.398	.381	.397	.344	.400
Labor (man-days)	243.3		238.9	253.8	253.3	288.1
Daily Wage Received	1.89	3.92	1.52	1.24	2.57	2.32
Production (Quetzales)	322.6	488.2	261.5	407.3	407.3	623.0

[a]Cuerda = .0046 hectares

TABLE C-2

Comparison of Ratios

	Illiterate Farmers	Literate Farmers	Ratio of Means
Production/Land Value	3.77	4.50	.838
Production/Labor	3.93	5.20	.750
Labor/Land Value	4.14	3.60	1.15
Inputs/Land Value	.229	.332	.692
Inputs/Labor	.070	.110	.66

Note: Labor measured in man-days.

The value of agricultural production (both sold and consumed by the household), cost of cash inputs and hired labor, and size and value of planted land are available for each farm and disaggregated at the crop level. Labor data were also collected in terms of man-days and contain elements on both family and hired labor, as well as on family labor used in the farm and work for wages. Wage data exist both for wages paid to hired labor (in terms of average by villages) and wages perceived in the labor market by family labor.

In order to illustrate the methodology used, let us assume that both literate and illiterate farmers are profit maximizers, and that there exists a production function y(l,t,i) -- denoting how output y relates to the factor inputs labor, l, land, t, and cash inputs, i -- common to both groups. Differences in productivity levels among groups can then be explained in terms of differences in optimal input combinations chosen according to the constraints operating in each group.

From the indices in Table C-2, one observes that educated (literate) farmers attain consistently higher levels of yield both per unit of land (in standardized quality units) and per unit of labor (man-days). Having assumed that the production function is the same for both groups as well as that farms in both groups are profit maximizers, the observed differences in labor-land and cash inputs-land ratios suggest that labor is the factor which is relatively more expensive for the literate group, while land and cash inputs are relatively more expensive for the illiterate group. Consequently, literate farmers will tend to substitute cash inputs for labor, while the opposite happens for the illiterate group: labor substitutes for land and cash inputs.

Possible explanations for these differences in factor prices and factor ratios are multiple: (1) in terms of cash inputs, illiterate farmers can

face higher input costs due to their lower scale of operations which, in turn, may be related to the relatively lower size of landholdings (in quality terms); (2) having access to less land, illiterate farmers can also face constraints in credit availability necessary for the purchase of cash-inputs; (3) lastly, lower use of chemical inputs may be related to prudence in adopting chemical fertilizers or introducing cash crops, more intensive in chemical inputs than the traditional crops; (4) as for labor inputs, the relative higher use of labor inputs by illiterate farmers may reflect differences in opportunity costs of family labor among the two groups -- the market wage rate of family labor in illiterate farmers is 60 percent lower than the wage rate of illiterate farmers.

If labor markets were competitive and farmers wanted to maximize total income accruing to the household, it would be better for both groups if the literate group used more of illiterate labor. The difference between the values of labor marginal product of the two groups (Table C-13) suggests that room exists for literate farmers to employ more labor and illiterate farmers to sell their services at wage rates higher than the marginal productivity obtained in farming their own plots.

In fact, for a same production function and without constraints in purchasing cash inputs, both groups would be operating at the same labor-land ratio.[4] Moreover, literate farmers having access to a greater amount of cash flow to finance the purchase of cash inputs would be better off using greater labor-land ratios, for increased inputs would lead to higher levels of labor productivity for the same amount of land.[5]

---

[4]See Appendix C-3 for technical discussion of this point.

[5]Since g/ (l/t) increases with i/t, literate farmers able to use higher cash inputs-land ratios would operate at larger labor-land ratios (l/t)*. That is, for the same wage w, marginal labor productivity of literate farmers greater than that obtained by the illiterate would lead to higher labor-land ratios among educated farmers.

In reality, this does not happen; literate farmers use labor-land inputs that are much lower than those used by the illiterate farmers, substituting chemical inputs for labor. The levels of marginal productivity among these two groups are also different, literate farmers operating at a level of marginal productivity substantially higher than the illiterate ones.

If this is the case, one of our initial assumptions is incorrect; either farmers are not profit or income maximizers, or other constraints exist to prevent labor from moving away from farming their own plots to working on other people's farms. This lack of labor mobility inside the villages may indicate that not only social considerations may be important in choosing to work for others, but also that small farmers attach greater value to agriculture production for household consumption than to similar income (in money terms) obtained from working for wages. That is, in a setting of small and quasi-subsistence farmers, maximization of farm production for consumption may operate as the objective function instead of maximization of profits or monetary income. This concern with securing food consumption for the household is exemplified in our sample where 38 percent of the farmers produce home consumption crops only (corn and beans), and even the 20 percent of the farmers who grow cash crops to sell in the market keep a level of crop production for home consumption twice as big as those who do not sell. This concern with home consumption production may reflect risk attitudes preventing small farmers from entering the labor market as well as precluding them from shifting from producing food staples towards more profitable cash crops before basic food requirements are secured. In such circumstances, labor to be hired by large landowners is limited to landless household members, forcing the wage rate of hired labor to be greater than it would otherwise be if small landowners (with high labor-land

combinations and low marginal labor productivities which are considerably less than the wage rates of landless workers) were to allocate part of their working time to work outside their own plots. Moreover, social and environmental circumstances may also be present to explain the reluctance of small farmers to sell their labor services in local labor markets, wage labor being performed by landless household members.

Thus, if it is true that farmers behave according to different objective functions, that is, small farmers possibly trying to maximize household consumption while larger farmers are being oriented towards profit maximizing, our tests for the effect of education on managerial and technical skills of rural farmers must take these differences into consideration. As stated before, differences in managerial skills can only be tested when people have common goals and objective functions. Accordingly, it would be theoretically correct to test whether, for small farmers producing for household consumption, education was associated with cost minimization, while for larger farmers oriented towards the market, education could explain differences in profit maximization behavior. Distinguishing between these two cases is necessary since the usual maximizing profit conditions derived in terms of factor prices and value of factor marginal products do not make sense for farmers producing only food crops for household consumption; these conditions should be assumed only for those who produce for the market. For the former group, testing for differences in management ability could be done by testing whether educated farmers operate at lower costs than uneducated ones, subject to constraints in minimum production levels. Determining these minimum levels of food production-consumption is difficult, however, and even more difficult is differentiating the objectives of minimizing costs from those of maximizing household food consumption.

In the absence of information on minimum production requirements, we chose to split the sample into three farm groups according to their market participation. We assume that the members of each group have similar goals in production. The analyses of Fisk (1964, 1977) and Nakajima (1969) based in A. K. Sen (1966) provided the rationale for the grouping task as described below.

Market Integration: What Differences Does It Make?

Both Fisk and Nakajima explain economic development in rural subsistence economies as a longitudinal process along which progressive contact with the market favors specialization and increases the value of monetary income relative to the value of production for consumption purposes. Their framework is applied to the case where there is no restriction on land, and cultural values play an important role in explaining participation in labor and commodity markets.

For our purposes, three phases of the Nakajima-Fisk model were chosen: (1) the first phase, the subsistence unit phase, corresponds to the situation when farmers do not have contact with money markets. They produce only staple goods for home consumption, and no market transactions are carried out. Their allocation of time and production is determined solely by their fixed resources (land and labor) and their utility function (in terms of income-leisure variables).

The second phases of the Nakajima-Fisk model, called subsistence with complementary cash income, arises with the first contacts with the market. Farmers begin to attach value to money income, to grow more than their immediate consumption needs, and to sell excess production from which they obtain certain money income. This farmer obtains from his fixed resources more than what he needs to feed his family and sells the remaining produce to earn money income. His participation in labor markets is often a way of

earning additional income, but he does not substitue food production by any other type of income.

The last phase in Nakajima-Fisk -- market integration with subsidiary food production -- denotes market specialization. Farmers allocate resources to products which are mainly sold in the market, although they may continue to produce food staples if this is more efficient. Their participation in the labor markets (in both buying and selling labor services) depends on efficiency criteria and not on preference towards production income versus money income. It reflects the situation when farmers value income in monetary values and are indifferent about the sources of such income.

In the present case based on a cross-sectional sample, constraints on land and associated aspects of risk attitudes and social values substitute for the temporal factor in the Nakajima-Fisk model and determine the existence of differences in market integration among the farmers of this cross-sectional study. We define three groups: the first group -- subsistence farmers -- who grow only food crops for household consumption. No sales occur, but farmers may sell labor services to complement agricultural income itself, constrained by their landholdings size which is, on the average, between 39 and 48 units of land value. The second group -- semi-subsistence farmers -- enters the market to sell production not used for household consumption. The nature of the sample does not allow us to identify whether such sales were the consequence of a good agricultural harvest or if such sales occur normally in every year. The average size of landholdings of this group varies between 66 and 86 units of land value. The last group -- commercial farmers -- corresponding to the third group of the Nakajima-Fisk model, sell and grow cash crops together with food crops. Their landholdings vary, in terms of quality units, between 126 and 165 units.

These groupings help us to analyze in further detail the differences observed among literate and illiterate farmers shown in Tables C-1 and C-2. The context of these villages is certainly different from the free-land scenario of Fisk and Nakajima. It is possible, however, for us to assume that land limitations operate in explaining the existence of these groups and analyze the influence of education within each of these groups. Table C-3 summarizes some indices of Table C-4.

TABLE C-3

PRODUCTION INDICES FOR THREE FARMING GROUPS BY LITERACY

	Subsistence		Semi-Subsistence		Commercial	
	Illit.	Liter.	Illit.	Liter.	Illit.	Liter.
Number of cases	140	56	129	83	54	48
Production/Land Value*	2.94	2.88	4.27	5.64	5.66	7.33
Production/ Labor	.73	1.10	1.34	2.36	1.48	1.79
Inputs/Land*	.181	.189	.189	.226	.457	.682
Labor/Land*	4.48	3.23	3.95	3.68	4.05	4.02
Average Area*	42.8	45.6	66.9	88.2	142.7	149.3
% Hired Labor	8.0	8.0	11.3	26.2	25.8	41.8
Income/Capita	63.7	108.3	89.8	136.0	147.3	224.3

*Value of Area Planted.

The elements presented above show that literate and illiterate farmers behave differently within each group. For example, within the commercial group, while both the average value for land planted and the labor-land ratios are quite similar for the two literacy groups, literate farmers use input-land ratios 49 percent higher and obtain levels of labor productivity 21 percent higher than those of the illiterate farmers. Although literate farmers cultivate 3 percent more of cash crops, this does not explain the difference among the cash input-land ratios. Illiterate farmers should be facing higher prices for cash inputs than illiterate farmers. As for

subsistence and semi-subsistence groups, the analysis made before applies: illiterate farmers operate with higher labor-land and lower input-land ratios. Again, inputs and land seem relatively more expensive for subsistence and semi-subsistence illiterate farmers. Substituting cash inputs for labor permits literate farmers to attain higher levels of both land and labor productivity. In all cases, the heavier use of labor by illiterate farmers comes from the more intensive use of family labor. Even among commercial farmers, the use of more family labor by illiterate farmers is apparent in Column 8.1 of Table C-4 or in "% of Hired Labor" of Table C-3.

What can we deduce from these empirical observations? Literate farmers tend consistently to use more cash inputs, to substitute labor by cash inputs, and to attain productivity levels consistently higher than the illiterate farmers. Such substitution is greater in the subsistence and semi-subsistence groups. In the commercial group which is oriented towards the market (80 percent of output sold), very similar labor-land ratios are depicted showing a likely influence of the market on farm management practices. Other questions can be raised. For example: "Is it the influence of literacy that makes farmers more willing to use cash inputs?; Can this be a proxy for greater availability of credit facilities or previous wealth?; Are there any other factors upon which education may act besides increasing input-land ratios?; Can we detect different attitudes in profit maximizing among these groups?; Are literate farmers more knowledgeable about their possibilities?; How does specialization interfere with these processes?; How can one derive from these indices sensible recommendations in terms of rural development policies?"

It is to answer these questions that the following analysis is addressed. First, we shall describe briefly the villages in the sample. After that,

TABLE C-4

ELEMENTS FOR THREE FARM GROUPS BY FARMER'S LITERACY

	TOTAL	SUBSISTENCE			SEMI-SUBSISTENCE			COMMERCIAL		
		Total	Illit.	Literate	Total	Illit.	Literate	Total	Illit.	Literate
Number of cases	510	196	140	56	212	129	83	102	54	48
INCOME SOURCES		387.4	27.5	10.1	41.6	25.3	16.3	20.0	10.6	9.4
			43.3	29.9		39.9	44.4		16.8	25.7
1. Total Gross Income (Quetz.)	603.7	406.5	373.6	488.8	570.2	472.1	722.6	1052.5	921.4	1199.9
2. Farm Gross Income (Quetz.)	348.9	117.9	114.1	127.2	311.3	269.9	375.7	871.1	796.9	954.5
3. Total Net Income (Quetz.)	537.2	396.5	363.0	480.4	534.1	446.9	669.6	813.9	737.0	900.3
4. Farm Net Income (Quetz.)	256.1	92.6	88.5	103.1	245.5	220.9	283.7	592.2	581.1	604.6
5. Wage Income (Quetz.)	205.4	231.5	214.0	275.1	211.1	177.4	263.6	143.5	116.3	174.0
6. Activity Income (Quetz.)	75.7	72.4	60.5	102.2	77.5	48.6	122.3	78.2	39.6	121.7
% Farm Income (4/3)	.477	.234	.244	.215	.460	.494	.423	.727	.788	.672
% Wage Income (5/3)	.382	.584	.590	.573	.395	.397	.394	..176	.158	.193
% Activ. Income (6/3)	.141	.183	.167	.213	.145	.109	.183	.096	0.54	.135
7. Net Income/P. Capita (Quetz.)	110.9	74.4	63.7	108.3	107.9	89.8	136.0	183.5	147.3	224.3
LABOR ALLOCATION										
8. Agriculture Labor	243.3	145.5	156.6	117.6	210.4	212.7	206.8	499.9	504.1	495.1
8.1 Family Agriculture Labor	189.4	133.4	143.8	107.5	171.6	187.7	146.6	333.9	374.2	288.5
% (8 · 1/8)	.883	.92	.92	.92	.83	.887	.738	.668	.742	.582
8.2 Hired Labor	52.9	12.1	12.9	10.3	35.8	24.0	54.2	166.2	130.0	206.9
8.2.1 % (8 · 2/8)	.217	.08	.08	.08	.170	.113	.262	.332	.258	.418
Daily Wage Paid (Quetz.)	.828	.817	.806	.832	.819	.817	.822	.848	.846	.849
Daily Wage Paid (Cases)	(205)	(43)	(25)	(18)	(94)	(43)	(51)	(68)	(35)	(33)
9. Family Wage Labor (Man-days)	109.9	137.9	142.1	127.6	105.9	105.6	106.2	64.6	56.2	74.0
9.1 Daily Wage Rec. (Quetz.)	1.90	1.63	1.44	2.12	1.62	1.49	1.85	3.40	1.99	4.67
Daily Wage Rec. (Cases)	(360)	(149)	(106)	(43)	(156)	(99)	(57)	(55)	(26)	(29)
10. Family Activ. Labor (Man-days)										
Net Income/Day (Quetz.)										
11. Total Labor (8+9+10) (Days)	354.2	284.1	300.1	245.2	318.6	319.9	316.6	563.5	558.3	569.3
11.1 % Family Labor (11/12)	.846	.899	.904	.886	.834	.899	.733	.768	.803	.729
11.2 % Hired Labor (8·2/12)	.152	.097	.090	.114	.163	.101	.260	.233	.198	.272
12. Family Size	5.35	5.40	5.66	4.77	5.21	5.10	5.39	5.53	5.63	5.42
13. Labor Family Agr./Capita (8·1/13)	41.5	27.5	28.5	25.1	39.9	44.6	32.6	71.7	80.3	62.0
14. Labor Family/Capita (11/13)	64.0	54.6	54.9	53.8	63.0	68.1	55.0	84.2	90.6	76.9
15. Labor Total/Capita (12/13)	76.6	57.3	57.7	56.1	73.4	74.9	71.9	120.4	120.0	120.8
FARM PRODUCTION										
Land										
16. Land (Cuerdas)	79.9	64.8	60.4	75.6	92.2	90.3	95.2	83.6	95.3	69.9
17. Land (Hectare)	3.75	3.04	2.83	3.55	4.32	4.23	4.46	3.92	4.49	3.28
18. Value Area Planted (Rental Value)	77.21	43.57	42.76	45.6	75.30	66.99	88.20	145.8	142.7	149.3
19. % Good Land	51.5	42.2	44.6	36.2	47.8	44.8	52.4	76.9	72.5	81.9
20. Intensity of Use*	1.30	1.32	1.33	1.29	1.21	1.24	1.15	1.46	1.44	1.49
21. Land Rented	36.2	29.4	30.3	25.8	38.8	37.2	40.1	41.7	32.8	49.5
21.1% Land Rented (20/17)	.36	.37	.41	.26	.37	.38	.35	.36	.36	.47

Table continued on next page

TABLE C-4 (cont'd.)

	TOTAL	SUBSISTENCE			SEMI-SUBSISTENCE			COMMERCIAL		
		Total	Illit.	Literate	Total	Illit.	Literate	Total	Illit.	Literate
21.2 Rent Rate (var)	.161	.175	.170	.194	.176	.135	.245	.106	.101	.111
Rent Rate (cases)	(251)	(86)	(68)	(18)	(111)	(70)	(41)	(54)	(25)	(29)
Rent Rate (land r)	1.55	.702	2.64	.963	1.41	.964	2.30	1.17	1.31	2.14
PRODUCTION										
22. Value 5 Major Crops	300.2	102.5	98.9	111.4	234.7	223.3	252.4	816.3	755.5	884.4
23. Value Agric. Product	322.6	102.5	98.9	111.4	281.6	246.1	336.7	830.8	765.5	904.2
24. Farm Production										
24.1 Food Crop	186.4	102.6	99.0	111.5	234.9	223.3	252.6	247.1	239.7	255.4
a) Consumed	137.4	89.9	85.8	100.2	157.3	150.5	167.9	187.3	206.4	165.8
b) Sold	49.1	12.7	13.3	11.3	77.5	72.9	84.7	59.8	33.2	89.7
24.2 Cash Crops	136.2	---	---	---	---	----	----	583.7	525.8	648.8
a) Consumed	12.3	---	---	---	---	----	----	61.6	54.9	69.0
b) Sold	123.9	---	---	---	---	----	----	522.1	480.9	579.8
VARIABLE INPUTS & CAPITAL										
25. Agricultural Labor	243.3	145.5	156.6	117.6	210.4	212.7	206.8	499.9	504.1	495.1
25.1 Hired Labor (8.2)	52.9	12.1	12.9	10.3	35.8	24.0	54.2	166.2	130.0	206.9
25.2 Wages Paid	.828	.817	.806	.832	.819	.817	.822	.848	.846	.849
26. Cash Inputs	22.9	6.29	5.76	7.62	13.84	11.53	17.43	73.5	63.0	85.4
27. Animal Food	.802	1.20	1.61	0.18	.476	.628	.241	.706	.556	.875
28. Animal Value	187.9	176.4	210.1	92.2	186.6	196.5	171.1	212.5	228.9	194.0
29. Physical Capital	2.36	1.89	1.91	1.86	2.79	2.43	3.34	2.36	1.87	2.92
PRODUCTIVITY INDICES										
30. In Terms of Area Owned										
30.1 Yield/Land (Quetz./Cuerdas)	5.87	2.71	2.59	2.99	4.51	3.89	5.48	14.8	12.5	17.3
30.2 Profit/Land (Quetz./Cuerdas)	4.68	2.30	2.19	2.56	3.96	3.54	4.62	10.7	9.20	12.5
30.3 Labor/Land (Days/Cuerdas)	4.60	3.63	3.80	3.19	3.42	3.56	3.21	8.81	8.27	9.42
30.4 Inputs/Land (Quetz./Cuerdas)	.394	.158	151	.173	.188	.174	.210	1.27	.966	1.62
31. In Terms of Value Land Cultivated										
31.1 Yield/Land Value	4.40	2.92	2.94	2.88	4.80	4.27	5.64	6.44	5.66	7.33
31.2 Profit/Land Value	3.67	2.45	2.47	2.40	4.20	3.83	4.76	10.7	9.19	12.5
31.3 Labor/Land Value	3.99	4.12	4.48	3.23	3.84	3.95	3.68	4.04	4.05	4.02
31.4 Input/Land Value	2.67	.183	.181	.189	.203	.189	.226	.563	.457	.682
32. In Terms of Labor										
32.1 Yields/Labor (Quetz./Days)	1.36	.84	.73	1.10	1.74	1.34	2.36	1.62	1.48	1.79
32.2 Profit/Labor	1.16	.687	.597	.911	1.55	1.16	2.17	1.26	1.17	1.37
32.3 Input/Labor (Quetz./Days)	.09	.050	.044	.067	.095	.083	.114	.150	.129	.174
33. Sale Prices										
33.1 Price Corn (Cases)	5.79 (166)	---	---	---	5.85 (126)	5.88 (78)	5.96 (48)	5.63 (40)	5.30 (21)	5.99 (19)
33.2 Price Beans (Cases)	16.41 (71)	---	---	---	15.98 (41)	15.6 (25)	16.6 (16)	17.0 (30)	16.7 (19)	17.5 (11)
33.3 Price Tomatoes (Cases)	2.39 (83)	---	---	---	---	---	---	2.39 (83)	2.28 (46)	2.52 (37)
33.4 Price Chili (Cases)	3.81 (43)	---	---	---	---	---	---	3.78 (43)	3.98 (21)	3.58 (22)
33.5 Price Maicillo (Cases)	5.48 (43)	---	---	---	5.51 (39)	5.34 (24)	5.73 (15)	5.13 (4)	5.50 (2)	4.75 (2)

a Cobb-Douglas production function is estimated to detect differences among farmers in organizing inputs to maximize production. Lastly, profit maximization and technical efficiency are tested together using the L. Lau and Yotopoulos (1971, 1972) method of normalized restricted profit functions.

Farm Efficiency Among Literacy Groups

Tests on scale and factor proportions are consistent with the view that illiterate farmers have higher land and cash input costs, but they cannot rule out the possibilities that illiterate farmers have less managerial knowledge or different goals. Even though literate and illiterate farmers act as if they face different factor prices, if one assumes that they operate under the same production functions, it should be possible to test whether they are equally efficient. As emphasized by Lau and Yotopoulos (1971, 1973), efficiency has two components. "Technical efficiency refers to a scale parameter in the production function." Price efficiency describes maximizing behavior.

To illustrate these two concepts, let us assume a fixed form for the production function

$$Y = F(L,T,I) \tag{1}$$

with arguments labor, L, land, T, and cash inputs, I. Define any K production functions which differ by a scale parameter, A_i, with i=1, . . ., K,

$$Y_i = A_i F(L,T,I) \tag{2}$$

Technical efficiency, or differences in technical efficiency, can be tested by means of testing for equality among the scale parameters A_i's. Thus, one says that farmer i is technically more efficient than farmer j, if their production function differs by only scale parameters and A_i is greater than

A_j; that is, with the same factors of production, farmer i attains greater output than farmer j by a factor A_i/A_j.

As seen before, literate farmers attained higher levels of land and labor productivity using higher input-land ratios but lower land ratios than the illiterate farmers. To know whether illiterate farmers are less efficient than literate farmers requires, then, that we test whether illiterate farmers produce less than would be expected given that they use less inputs per land unit and more labor than literate farmers. An analysis of technical efficiency in terms of land ratios can be carried out by modifying the explicit form of the production function in Equation (2).

Assuming that the production function F(.) is homogeneous of the first degree, Equation (2) can be written in terms of land ratios as

$$Y_i/T_i = A_i G(L/T,\ I/T) \qquad i = 1, \ldots, k \qquad (3)$$

where G(L/T, I/T) = F(L/T, 1, I/T). Equation (3) enables a preliminary examination of differences in technical efficiency based on the input combination ratios depicted in Tables C-2 and C-3. In fact, if Y_i/T, L/T, and I/T are approximately the same for the two literacy groups, one concludes that their scale parameter A_i is the same and, therefore, their level of technical efficiency will be the same as well. If the three ratios are not the same, which is the case, one must assume something more specific about the form of F(.) and take into account factor prices in order to make the test correctly. Examining the ratios as we did before had the virtue of simplicity but did not allow us to draw any definitive conclusions due to the differences in labor-land ratios between literate and illiterate farmers withing subsistence and semi-subsistence farms (see Table C-3). For commercial farmers, there were no differences in labor-land ratios. Therefore, in order to explain productivity differences between educated and illiterate

farmers, one has to use more complete analysis. We can adopt a specific form for the production function which allows correction for the variation in factor inputs across groups. However, price inefficiency creates an identification problem for tests of technical efficiency. As Yotopoulos and Lau showed, one should simultaneously test for both types of efficiency if price inefficiency is suspected.

Let us then define price efficiency (or allocative efficiency) before presenting the results of the empirical analysis which was carried out to test for differences in technical efficiency among literacy groups.

For a production function as in Equation (1), one says that price efficiency is achieved when the value of the marginal product of each factor equals its market price. Thus, for a generic factor of production x, [$x \in (L,I,T)$] with price p, price efficiency implies that the value of its marginal productivity equals its price

$$\partial Y/\partial x = p_x \tag{4}$$

If farmers fail to achieve price efficiency, their choice of factor combinations fail to comply with (4), and the degree of price inefficiency can be expressed by

$$\partial Y/\partial x = k_x p_x \tag{5}$$

where k stands for the inefficiency factor with respect to input x. Allowing for differences in inefficiency among farmers, one writes

$$\partial Y_i/\partial x_i = k_{xi} p_x$$

where k_{xi} is the inefficiency parameter of farmer i. Consequently, one can test whether farmers i and j differ in price efficiency by means of testing

whether the parameter k_{xi} and k_{xj} differ among farmers. This test can be carried out either indirectly through the computation of the factors' marginal productivities derived from estimated production functions or directly by estimating a normalized profit function as in Yotopoulos and Lau (1971, 1973).

By now, one concentrates in analyzing differences in technical parameters across farmers by fitting a Cobb-Douglas function both to the overall farm production as well as to specific crops in such a way as to control for the effects of literacy in production as well as the effects of market integration alone and together with literacy. The results are described below.

Technical Efficiency

A production function as

$$\ln Y = a + b_1 \ln L + b_2 \ln T + b_3 \ln I + \Sigma c_i V_i + dR + \Sigma e_j Com_j \qquad (6)$$

where (i=2,3,4) and (j=2,3), was fitted to the overall sample as well and to five different groups: literate and illiterate farmers, subsistence, semi-subsistence, and commercial farmers. Y, L, T, and I are agricultural production, labor, land value, and cash inputs respectively. V_i (i=2,3,4), R, and Com_j (j=2,3) are (0,1) variables denoting whether or not the farmer is located in village i, is literate, and whether the farm belongs to one of the groups defined with respect to market integration. The coefficients estimated for v_i will denote structural differences among villages; those estimated for r and e_j will represent differences in the technological parameters described above.

Definitions of the variables as well as means and standard deviations are presented in the following tables. Regression coefficients are presented

for the whole sample, the individual crops, and by literacy and commercial groups. The following conclusions seem evident.

1) All of the regression models seem to explain a fairly large proportion of the variation in total farm production. For most of the models, physical input coefficients are significant at the 1 percent level of confidence; land, labor, and input coefficients have the highest t-statistic values.

2) Education coefficients are significant at the 1 percent confidence level in the regression for the whole sample, decreasing in magnitude with index variables to denote the degree of commercialization of each farm entering the regression. The same variables play an important role in the regressions fitted for the literate and illiterate groups. One observes that land, labor, and input coefficients decline when Com_2 and Com_3 enter the regressions, denoting the correlations between these inputs and the degree of market integration of each farm.

3) Although the ltieracy coefficient in the overall regression is significantly different from zero, the test for structural differences among literate and illiterate groups leads to the conclusion that we cannot reject the null hypothesis that the two groups have the same production function. Literacy thus appears to affect the technical scale factor but not the physical input elasticities.

4) Disaggregation of the sample into three different groups according to the <u>degree of commercialization</u> leads to different conclusions, however. Each group seems to have a different production function. It is particularly interesting to note that the coefficient of cash inputs increases in magnitude and t-statistic value from the subsistence group to the commercial, that land elasticity is much higher than labor elasticity among the semi-subsistence group, and that labor elasticity is particularly high among

TABLE C-5

VARIABLE DEFINITIONS

VARIABLE	DEFINITION
LNY	Value of total farm production (ln)
LNLAB	Labor (man-days)
LNLD	Area of land used (cuerdas; 1 cuerda = .0469 hectares)
PC34	Percentage of land irrigated and good for corn
LNLV	Value (rental value) of land cultivated (quetzales/manzana; 1 manzana = 16 cuerdas)
LNINP	Value of purchased inputs other than hired labor
LNKA	Value of capital animal (ln)
LNKB	Value of capital (ln)
V2	Dummy = 1 if village = 6
V3	Dummy = 1 if village = 8
V4	Dummy = 1 if village = 14
COM2	Dummy = 1 if food crops sold
COM3	Dummy = 1 if cash crops sold
RWPA	Dummy = 1 if farmer literate
GRADPA	Farmer's educational grades
GRADMA	Farmer's wife's educational grades
AGR7	Average education grades of household members aged 7 or more
AGR16	Average education grades of household members aged 16 or more
AGEPA	Farmer's age
PCOWN	Percentage of land owned by farmer

TABLE C- 6

MEANS AND STANDARD DEVIATIONS; ALL SAMPLE, ILLITERATE AND LITERATE FARMERS, SUBSISTENCE, SEMI-SUBSISTENCE, AND COMMERCIAL FARMS

VARIABLES	ALL SAMPLE		BY EDUCATION				BY COMMERCIALIZATION					
			Illiterate		Literate		Subsist.		Semi-Subsist.		Commercial	
	Mean	S.D.	Mean	S.D.	Mean	S.D.	Mean	S.D.	Mean	S.D.	Mean	S.D.
N. Cases	510		323		187		196		210		102	
LNY	5.21	1.09	5.09	1.03	5.44	1.16	4.48	0.82	5.38	0.85	6.30	0.93
LNLAB	5.11	0.96	5.11	0.96	5.11	0.97	4.89	0.88	5.07	0.87	6.01	0.65
LNLD	3.911	0.91	3.91	0.90	3.91	0.91	3.63	0.92	4.09	0.89	4.09	0.76
PC34	0.3	0.46	0.36	0.46	0.44	0.47	0.42	0.73	0.48	0.71	0.77	0.79
LNLV	3.98	0.86	3.91	0.83	4.10	0.88	3.55	0.69	3.99	0.80	4.78	0.67
LNINP	2.08	1.43	1.92	1.34	2.27	1.56	1.36	1.12	1.90	1.16	3.80	1.14
LNKA	3.47	2.50	3.57	2.53	2.30	1.55	3.03	2.62	3.62	2.42	4.01	2.30
LNKB	0.62	0.79	0.57	0.75	0.70	0.86	.505	0.73	0.71	0.87	0.64	0.72
V2	0.34	0.47	0.35	0.48	0.31	0.46	0.23	0.42	0.23	0.42	0.76	0.43
V3	0.20	0.40	0.15	0.36	0.28	0.45	0.21	0.41	0.21	0.41	0.15	0.37
V4	0.19	0.39	0.22	0.42	0.13	0.34	0.25	0.43	0.21	0.41	0.04	0.20
COM2	0.41	0.49	0.40	0.49	0.44	0.50	--		--		--	
COM3	0.20	0.40	0.17	0.37	0.26	0.44	--		--		--	
RWPA	0.37	0.48	--	--	--	--	0.29	0.45	0.39	0.49	0.47	0.50
GRADPA	1.09	1.64	0.30	0.88	2.46	1.75	1.02	1.68	1.25	1.69	1.40	1.64
AGEPA	44.3	14.3	46.5		40.4		44.2	13.8	44.7	15.4	43.5	13.2
GRADMA	0.74	1.53	0.56	1.41	1.06	1.67	0.73	1.57	0.69	1.58	0.92	1.57
AGR7	1.22	1.22	0.87	0.98	1.85	1.34	1.25	1.16	1.36	1.28	1.49	1.18
AGR16	1.16	1.30	0.73	0.97	1.92	1.45	1.19	1.83	1.28	1.38	1.44	1.31
PCOWN	43.6	42.0	45.6	41.7	40.0	42.3	36.8	40.3	46.8	41.8	49.8	44.2

TABLE C- 7

OLS ESTIMATES OF PRODUCTION FUNCTION COEFFICIENTS.
ALL SAMPLE, LITERATE AND ILLITERATE FARMERS, SUBSISTENCE, SEMI-SUBSISTENCE, AND COMMERCIAL FARMS.

			LITERACY				DEGREE OF COMMERCIAL		
	ALL SAMPLE		ILLITERATE		LITERATE		Subs.	Semi-Sub.	Commerc.
Cases	510		323		187		196	212	102
LNLAB	.291* (6.93)	.268* (6.09)	.290* (5.80)	.256* (4.49)	.374* (4.29)	.310* (3.64)	.244* (3.49)	.064 (1.02)	.715* (5.67)
LNLD	.284* (6.76)	.244* (4.21)	.354* (4.59)	.304* (4.16)	.218* (2.15)	.157* (1.55)	.298* (3.10)	.375* (4.08)	.019 (.154)
LNINP	.285* (10.18)	.224* (7.47)	.276* (2.23)	.205* (5.69)	.249* (4.88)	.249* (4.52)	.139* (2.84)	.205* (4.46)	.236* (3.63)
LNKA		.039* (3.25)	.023 (1.53)	.030* (2.14)	.072* (3.00)	.057* (2.59)	.044* (2.44)	.040* (2.35)	.007 (.269)
LNKB	.207* (5.31)	.178* (4.45)	.092 (1.80)	.131* (2.67)	.250* (3.52)	.234* (3.39)	.120 (1.71)	.279* (5.17)	.157 (1.74)
V2		-.119 (-1.34)	-.151 (-1.43)	-.202 (-1.94)	-.026 (.171)	-.035 (-2.09)	-.118 (-.849)	-.127 (-1.06)	-.203 (-.659)
V3		.110 (1.06)	-.035 (-.265)	-.055 (-.434)	.372 (1.99)	.344 (1.91)	-.343 (-1.97)	.011 (.072)	.481 (1.43)
V4		.103 (1.30)	.026 (.591)	.044 (.500)	.271 (1.53)	.211 (1.26)	.140 (1.228)	.056 (.509)	-.205 (-.505)
RWPA	.176* (2.98)	.131* (2.34)	--- ---	--- ---	--- ---	--- ---	.127 (1.34)	.200* (2.53)	.016 (.135)
COM2		.499* (8.32)		.430* (6.32)		.602* (4.97)	--- ---	--- ---	--- ---
COM3		.626* (6.14)		.602* (5.39)		.577* (3.22)	--- ---	--- ---	--- ---
CONST		.160* (7.86)	1.596*	1.80* (8.22)	1.52* (3.87)	1.76* (4.76)	1.92* (6.58)	2.76 (9.75)	.960 (1.54)
R^2		.723	.695	.736	.657	.702	.541	.616	.677
R. ADJ.		.717	.687	.728	.642	.685	.519	.599	.646
ESS	167.3		89.4		74.4		60.56	58.9	28.3
D.F.	498		312		174.0		186	202	92

T-statistics between parentheses

*Statistically significant at the 1 percent level of confidence.

TABLE C-9

VARIABLE DEFINITIONS, MEANS AND STANDARD DEVIATIONS, FIVE MAJOR CROPS.

	Corn	Beans	Tomato	Chili	Maici.	Corn[1]	Beans[1]	Corn[2]	Beans[2]
NUMBER OF CASES	500	305	91	42	114	267	267	44	44
LNYi, Value of production	4.60 (.866)	3.74 (1.18)	5.08 (1.10)	6.15 (1.23)	3.53 (.909)	4.70 (.802)	3.72 (1.24)	5.11 (0.68)	4.68 (0.99)
LNLBi, Labor (man days) (ln)	4.75 (.717)	3.56 (1.15)	4.69 (.610)	5.22 (1.07)	3.41 (1.10)	4.92 (.643)	3.57 (1.14)	5.21 (0.57)	4.12 (1.13)
LNLDi, Area planted (cuerdas)	3.87 (.664)	2.75 (.900)	2.189 (.829)	2.09 (1.01)	2.65 (.912)	4.00 (0.63)	2.76 (0.94)	4.32 (0.54)	3.24 (0.93)
LNINPi, Cash inputs (ln)	1.33 (.985)	1.90 (1.21)	3.56 (1.28)	3.60 (1.27)	-1.49 (.902)	1.45 (0.93)	1.91 (1.22)	1.87 (0.90)	2.49 (1.11)
D3, dummy = 1, if land is good to grow corn	.382 (.488)	.381 (.487)	.363 (.483)	.238 (.431)	.272 (.447)	.416 (0.49)	.375 (0.48)	.614 (0.49)	.59 (0.50)
D4, dummy = 1, if land irrigated	.056 (.230)	.022 (.147)	.264 (.443)	.524 (.506)	0 --	.023 (0.15)	.023 (0.15)	.023 (0.15)	.023 (0.15)
RWPA, dummy = 1, if farmer is literate	.364 (.482)	.363 (.482)	.407 (.494)	.500 (.506)	.325 (.470)	.363 (0.48)	.363 (0.48)	.318 (0.47)	.318 (0.47)

[1]Group farmers who grow both crops.

[2]Group farmers who sell both crops.

TABLE C-10

OLS ESTIMATES FOR PRODUCTION COEFFICIENTS, FIVE CROPS.
Dependent Variable: Value produced.
Sample: All who grow each crop.

VARIABLES	Corn	Beans	Tomato	Chili	Maicillo
LNLB, labor	.196* (2.96)	.169* (2.54)	.674* (2.89)	.929* (4.55)	.324* (4.15)
LNLD, land	.507* (6.60)	.250* (3.02)	.337* (2.06)	.206 (1.14)	.089 (.946)
LNINB, inputs	.249* (7.21)	.513* (11.8)	.216* (2.06)	-.03 (-.25)	.34* (4.01)
D3, dummy good land	.163* (3.27)	.217* (2.39)	-.34 (-1.43)	.231 (.66)	-.061 (-.36)
D3, dummy irrigated land	.565* (4.96)	-.028 (-.093)	-.327 (-1.61)	-.25 (.83)	--- ---
RWPA, dummy literacy	.062 [1] (1.21)	.245* [2] (2.67)	.152 (.80)	.246* (.70)	.150 (1.00)
Constant	1.25* (6.47)	1.30 (8.46)	.429 (.474)	1.15 (1.63)	2.67* (7.84)
R^2	.617	.631	.569	.720	.412
R^{*2}	.612	.623	.539	.669	.380
..ES	142.1	139.0	45.21	17.6	45.85
D.F.	493	266.0	77	34	96

T statistics between parentheses

(*) significantly different from zero at the 1 percent of confidence level

1) RWPA = .008 (t=.065) when COM2 and COM3 enter the regression

2) RWPA = .215 (t=2.44) when COM2 and COM3 enter the regression

TABLE C-11

OLS ESTIMATES PRODUCTION COEFFICIENTS, FOOD CROPS, AND TOTAL
Dependent Variable: Production value (PV) and Value added (VA)
Groups: (1) Farmers who grow both, (2) Farmers who sell both

FARMER GROUPS	Sell and Grow Both (N = 267)				Sell Both Corn and Beans (N = 44)				Total Sample (N=510)	
CROP	CORN		BEANS		CORN		BEANS			
Dependent Variable	PV	VA	PV	VA	PV	VA	PV	VA	PV	VA
LNLB, labor	.285* (2.74)	.199 (1.53)	.177* (2.46)	.269* (2.59)	.598* (2.12)	.551 (1.89)	.174 (1.29)	.279 (1.64)	.483* (8.51)	.677 (9.67)
LNLD, land	.662 (4.10)	.727* (5.43)	.285* (3.17)	.583* (4.85)	.153 (.049)	.276 (.899)	.132 (.44)	.480 (2.36)	.206* (3.25)	.296* (32.9)
LNINP, inputs	.223 (4.72)	---	.508* (10.80)	---	.106 (.955)	---	.505* (3.74)	---	.239* (8.53)	---
D3, dummy land good	.048 (.77)	.108 (1.35)	-.022 (.068)	.035 (.075)	.350 (.073)	.328* (2.18)	.170 (.28)	.239 (.973)	.210* (3.70)	.271* (3.35)
D4, dummy land irrigated	.208 (.985)	.124 (.466)	.287* (2.90)	.382* (2.63)	.327* (2.19)	.368 (.74)	.194 (1.03)	.357 (.453)	.590* (6.28)	.498* (3.92)
RWPA, dummy literacy	.096 (1.46)	.186* (2.38)	.327* (3.27)	.525* (3.59)	.342* (2.09)	.328* (2.17)	.252 (1.24)	.477 (1.86)	.205* (3.54)	.309* (3.77)
CONSTANT	1.07 (3.73)	.589 (1.91)	1.11* (6.61)	.521 (2.11)	.813 (1.07)	.636 (6.81)	2.08* (5.33)	1.51 (3.10)	.07 (5.24)	-.02 (.080)
R^2	.610	.453	.615	.376	.618	.563	.708	.523	.693	.489

(t-statistics between parenthesis)

*Significant at the 1% level of confidence.

value and significance. The fact that specific prices are used may also explain the increased value in the coefficients. As shown in Table C-3, educated farmers seem to be able to obtain higher selling prices than illiterate farmers.

In short, the analysis of production functions seem to indicate that education (literacy) is related to higher levels of agriculture productivity and that education works as a scale parameter, shifting the production function without changing the output elasticities of each input. Differences in market integration seem, however, to be most important; production functions were detected to be structurally different between groups, with education appearing to have the greatest influence on agricultural productivity for the semi-subsistence group. This is consistent with the findings obtained for the effect of education on the production function for beans and for that group of farmers which sells both beans and corn.

Relative Technical and Price Efficiency

The fact that price inefficiency creates an identification problem for tests of technical efficiency requires that we test whether the findings obtained for the individual crops persist in the simultaneous test for technical efficiency and price, or allocative, efficiency. The method of Yotopoulos and Lau (1971) will be used. The important tools are a definition of price efficiency and a variable profit function.

Price efficiency occurs when the farmer equates the marginal product of a given input to its price. For a function $y=f(x)$, price efficiency will be obtained whenever $f_x=p_x/p_y$. Conversely, one can define an index of price inefficiency as k^i, and write $f_x=k^i(p_x/p_y)$ as the maximization profit rule of a farmer i who perceives the current prices affected by an error k^i. If any two farmers have the same error k, then their

relative efficiency will be equal although both might be inefficient in operating their farms. If $k^i=k^j=1$, both farmers will be absolutely efficient in terms of price or allocative efficiency. Since farmers can differ in technical constants a_i, as seen before, a simultaneous test for $a_i=a_j$ and $k_i=k_j=1$ has to be carried out.

For a Cobb-Douglas function, and using duality theory, Lau and Yotopoulos show that there is a variable profit function corresponding to the production function defined by (6) as follows:

$$P = P(w,T,I) = A^{(1-b_1)^{-1}} (1 - b_1)^{-\alpha} w^{\alpha} T^{\beta} I^{\gamma} \qquad (7)$$

where P is the normalized variable profit, w is the normalized wage rate, $\alpha = -b_1(1-b_1)^{-1}$, $\beta = b_2(1-b_1)^{-1}$, and $\gamma = b_3(1-b_1)^{-1}$.

Variable profit is a decreasing function of the normalized wage rate and an increasing function of b and inputs. Variable profit depends upon the wage rate and the fixed factors (cash inputs being associated with credit availability and considered among the fixed factors).

If farmers vary in technical efficiency, one can substitute A^i for A in (7); if the two groups of farmers vary in their price efficiency, however, this straightforward substitution is not possible. To incorporate the inefficiency scale factors as defined in (5) above, Lau and Yotopoulos show that the observed variable profit function is

$$P_i = A^*_i w^{\alpha} T^{\beta} I' \qquad (8)$$

where for i=1,2,

$$A^*_i = A_i^{(1-b_1)^{-1}} (1 - b_1/k_i)(k_i)^{-\alpha} b_1^{-\alpha} \qquad (9)$$

The goal is to test whether or not $A_1=A_2$, where 1 denotes illiterates and 2 denotes literate farmers by using a dummy variable to estimate A_2^*/A_1^*. The test will not work if $k_2 \neq k_1$ because the ratio A_2^*/A_1^* involes the constants k_1 and k_2, as follows:

$$\frac{A_2^*}{A_1^*} = \left(\frac{A_2}{A_1}\right)^{(1-b_1)^{-1}} \left(\frac{1 - b_1/k_2}{1 - b_1/k_1}\right)\left(\frac{k_2}{k_1}\right)^{\alpha} \tag{10}$$

If $k^2=k^1$, however, the ratio simplifies to the following:

$$A_2^*/A_1^* = (A_2/A_1)\ (1-b_1)^{-1} \tag{11}$$

First, we must test $k_2=k_1$. To do so, Yotopoulos and Lau exploit a convenient property of the Cobb-Douglas Function. When factor prices vary, quantities adjust to keep factor shares constant. The same is true for the ratio of variable factor costs to variable profit. Suppose a firm responds efficiently to the wage rate, i.e., $k_i=1$. Then

$$-wL/P = -b_1(1-b_1)^{-1} = \alpha \tag{12}$$

But, if the farmers are acting as if the wage rate is $k_i w$, then their ratio $k_1 wL/P$ will be constant as w varies, not wL/P. We do not observe $k_1 w$, but the authors show that when $k_i \neq 1$,

$$-wL/P = (1-b_1)(k_i(1-b_1/k_i))^{-1} = \alpha_i \tag{13}$$

This equation defines α_i. One can, therefore, test for $k^i=k^j$ by testing whether or not the ratio of the observed wage bill to variable profit is significantly different for illiterate and literate farmers. This can be done by means of

$$-wL/P = a^i/D^i + a^1 D^1$$

where D^i and D^1 are (0,1) variables denoting illiterate and literate farmers.

If the test shows equal relative price efficiency across groups, we can proceed to test absolute price efficiency and relative technical efficiency with the Cobb-Douglas profit function.

The two profit functions are

$$P_2 = A_2^* \, w^{\alpha} T^{\beta} I^{\gamma} \tag{14}$$

$$P_1 = A_1^* \, w^{\alpha} T^{\beta} I^{\gamma} \tag{15}$$

If one replaces A_1^* in (15) by $A_1^*(A_2^*/A_1^*)$, one can estimate (15) by means of

$$\ln P = \ln A^* + d\, D_2 + \alpha \ln w + \beta \ln T + \gamma \ln I \tag{16}$$

where $d = \ln (A_2^*/A_1^*) = (1/1-b_1) \ln (A_2/A_1)$. If d is significantly different from zero, then A_2 and A_1 are not equal. We can also test absolute price efficiency because if $k_2=k_1=1$, then $\alpha_1=\alpha_2=\alpha$, as can be seen by inspecting the equation defining α in (12).

Coefficients of (14) and (15) estimated for each crop are shown in Tables C-13a and C-13b.

The results are shown by both the Zellner method and OLS. The test of equal technical efficiency and price efficiency is performed by imposing the equality on the appropriate coefficients: $\alpha_1 = \alpha_2$.

As we can see, only for beans and tomato can we reject the hypothesis of equal price efficiency. For the remaining cases, price efficiency differences do not seem to be important, and proceeding to test for differences in technical efficiency does not yield rejection of equal technical efficiency. The simultaneous test for technical and price efficiency, thus, shows that the differences in technical efficiency detected in the bean production function analysis might be due to differences in price efficiency; the remaining crops do not benefit much from this method of analysis.

TABLE C-12

VARIABLE DEFINITIONS, MEANS AND STANDARD DEVIATIONS, NORMALIZED RESTRICTED PROFIT FUNCTION FOR HIRING LABOR FARMS AND FIVE CROPS

VARIABLES		CORN	BEANS	TOMATO	CHILI	MAICILLO
Ln π^*	Normalized profit (ln)	3.13 (.879)	1.76 (.923)	4.10 (1.21)	4.37 (1.19)	1.87 (1.22)
Ln W*	Normalized wage (ln)	-1.96 (.231)	-2.98 (.206)	-1.09 (.41)	-1.44 (.433)	-1.97 (.190)
D	Dummy Literacy	.470 (.502)	.478 (.505)	.500 (.500)	.529 (.507)	.412 (.507)
Ln LD	Area Planted (ln)	3.98 (.760)	3.38 (.944)	2.41 (.914)	2.17 (.995)	2.96 (1.08)
VIL2	Dummy, vil=6	.434 (.499)	.652 (.482)	.932 (.255)	.529 (.507)	---
VIL3	Dummy, vil=8	.157 (.366)	.044 (.206)	---	---	---
VIL4	Dummy, vil=14	.133 (341)	.174 (383)	.046 (211)	.029 (172)	.235 (437)
SHW	Labor Share ($-w^*L/\pi^*$)	-.455 (10.2)	-.056 (7.50)	-.485 (1.31)	-.523 (.978)	-.981 (1.99)
Number of cases		87	47	50	37	19

TABLE C-13a

ESTIMATION OF A NORMALIZED RESTRICTED PROFIT FUNCTION AND LABOR DEMAND FUNCTION; HIRING LABOR FARMS; CORN AND BEANS

(t-statistics between parentheses)

Parameters	BEANS				CORN			
	OLS	ZELLNER'S METHOD			OLS	ZELLNER'S METHOD		
		Unrest.	1 Restr.[a]	2 Restr.[b]		Unrest.	1 Restr.[a]	2 Restr.[b]
PROFIT FUNCTION (1)		(2)	(3)	(4)	(5)	(6)	(7)	(8)
ln A*	-2.70 (1.91)	-2.89 (-1.86)	-2.90 (-1.85)	-3.13 (-3.81)	.175 (.161)	.346 (.332)	.380 (.364)	.150 (1.54)
d (literacy)	.445 (2.02)	.362 (1.69)	.209 (1.03)	.197 (.979)	-.205 (.910)	-.204 (.907)	-.279 (-1.29)	-.277 (-1.28)
a (ln w*)	-.407 (-1.81)	-.766 (-1.64)	-.801 (-2.70)	-.877 (-4.64)	-1.34* (-2.90)	-1.26* (-2.85)	-1.26* (-2.84)	-1.38* (-3.48)
b (ln LD)	.618* (5.43)	[illegible] (6[illegible]	.771* (6.15)	.774* (6.22)	.077 (.545)	.074 (.548)	.074 (.548)	.076 (.555)
V_2(viL2)	-.240 (-.781)	-.419 (-1.34)	-.407 (-1.29)	-.402 (-1.29)	-.310 (-1.97)	-.24 (-1.6)	-.310 (-1.97)	-.222 (-1.30)
V_3(viL3)	-.597 (-.95)	-.183 (-.310)	-.208 (-.350)	-.176 (-.304)	-.357 (-1.59)	-.458 (-2.11)	-.357 (.225)	-.500 (-.206)
V_4(viL4)	-.399 (-1.20)	-.366 (-1.00)	-.370 (-1.09)	-.356 (-.989)	-.579* (-2.35)	-.577 (-2.43)	-.579 (-2.35)	-.584 (-2.19)
LABOR DEMAND FUNCTION								
$r_1(D_1)$	-1.33* (-4.89)	-1.33* (-4.89)	-.894* (-4.32)	-.877* (-4.63)	-2.83* (-2.38)	-1.26 (-2.16)	-1.82 (-2.13)	-1.38 (-3.48)
$r_2(D_2)$	-.431 (-1.54)	-.431 (-1.54)	-.894 (-4.32)	-.877 (-4.63)	-.783 (-.650)	-.510 (-7.13)	-1.82 (-2.13)	-1.38 (-1.38)
R^2(OLS)	.546				130			
F computed			(1,84) 5.55	(2,84) 2.78			(1,168) 1.48	(2,168) 907
Critical values for 1% confidence level			6.90	4.85			6.80	4.74
Critical values for 5% confidence level			3.95	3.10			3.90	3.05

TABLE C- 13b

COEFFICIENTS OF NORMALIZED RESTRICTED PROFIT AND LABOR DEMAND FUNCTIONS: TOMATO, CHILI, MAICILLO

(t-statistics between parenthesis)

	TOMATO			CHILI			MAICILLO		
	Unrestr.	1. Restr.	2. Restr.	Unrestr.	1. Restr.	2. Restr.	Unrestr.	1. Restr.	2. Restr.
VARIABLE PROFIT FUNCTION									
ln A^*	2.99* (2.82)	3.01* (2.87)	3.09* (2.95)	-.250 (-.188)	-.209 (-.157)	-.062 (-.043)	.989 (.324)	1.09 (.355)	-.359 (-1.24)
d (RWPA)	-.121 (-.37)	-.187 (-.58)	-.185 (-.58)	.179 (.380)	.118 (.250)	.031 (.060)	-.087 (-.099)	-.396 (-.465)	-.428 (-.467)
a(ln w*)	-.058* (-2.21)	-.059* (-2.23)	-.504 (-3.50)	-1.79* (-3.32)	-1.79* (-3.33)	-1.58* (-2.79)	.014 (.013)	.004 (.004)	-.653 (-7.34)
b(ln LD)	.676* (3.55)	.677* (3.54)	.691* (3.73)	.575* (2.22)	.574* (2.22)	.605 (2.12)	.366 (1.07)	.352 (1.02)	-.211 (.625)
V_2(viL2)	-1.24 (-1.42)	-1.24 (-1.41)	-1.27 (-1.44)	.394 (.749)	.392 (.745)	.356 (.617)	-3.48* (-2.16)	-3.96* (-2.07)	-2.93 (-1.73)
V_4(viL4)	-1.25 (-1.13)	-1.24 (-1.13)	-1.29 (-1.17)	.275 (.187)	.271 (.185)	.387 (.239)	-3.19 (-.377)	-3.52 (.414)	-11.76 (-.198)
LABOR DEMAND FUNCTION									
r_1 (D_1)	-.139 (-.626)	-.471 (-2.76)	-.504 (-3.50)	2.42 (.038)	-.127 (-.071)	-1.58 (-2.79)	-1.60 (1.15)	-.653 (-.734)	-.653 (-.734)
r_2 (D_2)	-.884* (-3.55)	-.471 (-2.76)	-.504 (-3.50)	-2.29 (.964)	-.127 (-.071)	-1.58 (2.79)	-1.60 (1.15)	-.653 (-.734)	-.653 (-.734)
R^2	.346			.390			.41		
F computed		4.96	2.55		1.81	1.32		1.62	1.28
Critical value for 1%		6.80	4.70		7.05	4.95		7.62	7.50
Critical value for 5%		3.89	3.05		3.90	3.14		4.19	3.32

TABLE C-14
FORMAL EDUCATION AND AGRICULTURE PRODUCTIVITY
Education Variable: Farmer's Literacy

FARM GROUP	Dependent Variable	Coefficient Education's Influence on Farm Output	t statistics	R^2	Increase % in Production Related With One Year of Schooling*	Source in Text	Comments
All, N-510	Agricult. Production	.131	2.12	.689	5.69%	2.2.2	
	Farm Production	.153	2.55	.689	6.72%	2.2.2	
	Value Added	.309	3.77	.489	14.7%	2.2.9	Possible bias due to inputs
Subsistence Farms N=196	Farm Production	.127	1.34	.541	5.5%	2.2.4	excluded
Semi-subsist. N=212	Farm Production	.200	2.53	.616	9.0%	2.2.5	
Commercial Farms, N=102	Farm Production	.016	0.135	.677	0.6%	2.2.5	
Semi-subsist. and Commercial N=314	Agricult. Production	.156	1.93	.599	6.9%	---	

table continued on next page

TABLE C-14 (cont'd.)

FARM GROUP	Dependent Variable	Coefficient Education's Influence on Farm Output	t statistics	R^2	Increase % in Production Related With One Year of Schooling*	Source in Text	Comments
Corn Producers N=500	Value of Harvest Av. Prices	.062	1.21	.617	2.6%		
Bean Produc. N=273	Value Bean Harv. Av. Prices	.245	2.67	.623	11.3%		
Tomato Producers N=91	Value of Tomato Production	.152	.80	.539	6.7%		coefficient not signif. diff. of zero
Chili Producers N=42	Value Chili Production	.246	2.70	.669	11.3%		
Maicillo Producers N=114	Value of Maicillo Produced	.150	1.00	.380	6.6%		coefficient not signif. diff. of zero
Sell Corn or Beans N=267	Corn Harvest[1]	.095	1.46	.607	4.1%		coefficient not significant
	Bean Harvest[1]	.327	3.27	.615	15.7%		

table continued on next page

In summary, the conclusions of this analysis were as follows:

i) Education seemed to be relevant to explain differences in agricultural productivity. Literate farmers attained higher levels of both land and labor productivity, partly because of the higher use of chemical inputs and partly because a relatively higher proportion of their total output was in cash crops. Even after controlling for those factors, education seemed to be significant in explaining variations in agricultural production among farmers. This was especially important for semi-subsistence farmers.

ii) Analysis of allocative or market efficiency did not lead to definitive results on the influence of education. Both subsistence and semi-subsistence farmers operate at very low levels of labor productivity which may reflect the fact that either they value wage income less than farm production or that they use labor whose opportunity cost is well below the market wage rate. In the first case, our basic assumption that farmers attempt to maximize does not hold, and the results of the analysis are useless as a test for market efficiency. In the second case, since information is not available on the actual work of each family member, we are also unable to reach conclusions on efficiency.

iii) The fact that educated farmers were slightly better off than their illiterate peers in terms of underemployment may indicate that they valued alternative non-farm occupations for their children or other family members or merely that, for the same input combinations, a slight edge in land quality increased the average labor productivity considerably. Only further research on this subject can shed some light on the matter.

Appendix C-1

A Word on Poverty in the Villages

On the average predominantly poor (the income per capita in 1974 was around 132 dollars), these villages also show a substantial degree of inequality within themselves. In terms of land distribution, the upper quintile of the farmers control more than 40 percent of the land. As for total income, wages and other income sources compensated for a small part of the skewed distribution in land; but, even so, 5 percent of the wealthiest families received 25 percent of the income. Gini coefficients computed for these two variable distributions were .555 and .495 respectively. With small alterations, the picture for each village did not change. Total agriculture production was equally highly concentrated -- more than 32 percent accruing to the top quintile of the distribution.

The direct consequence of such unequal distribution of both means of production and income inequality is that a considerable number of families lie below the poverty level, poverty being defined by local standards. The average wage for all four villages is taken as the lower limit of acceptable income for a family of 4.8 persons (approximately 220 quetzales a year). A second poverty measure was computed by estimating the cost of buying the basic food products (corn and beans) at the average village prices. This amounts to 45.1 quetzales per person/year for 220 kg. corn and 39 kg. beans year/person. For an average of a 4.8 person family, the computed poverty level equals 207.5 quetzales/year in 1974 prices.

Disregarding other living expenses and using these values to define a poverty level, we can conclude that 35 percent of the families in the villages were below the poverty level. The consequences for them in terms of child malnutrition, poor school attendance and performance are certainly to be expected. Table C-15 depicts some indicators for the two groups of families divided according to their relative position with

TABLE C-15

The Profile of Poverty in the Four Rural Communities, 1974
(In Percent)

	Poor Families	All Others
Age of Head of Household		
14-29	15	19
30-60	70	64
61+	15	17
Education of Head of Household		
Literate	26	43
Illiterate	74	57
Educational Grades Household Head		
0	72	54
1-2	12	19
3-4	11	21
5+	5	8
Position in Occupation of Head of Household		
Employer (hiring wage labor)	3	20
Self Employed	31	25
Share Cropper	51	34
Wage-laborer in community	9	11
Employed outside community	6	10
Total Land Exploited	(44.4)	(77.8)
Nuclear Land	61	71
Extended Family Land	9	7
Rented Land	21	19
Municipal Land	9	3
Sources of Income	(211)	(713)
Agriculture Income	48	45
Wages	36	40
Activities, arts and crafts	16	15
Number of Children 0-10		
0	30	19
1-2	34	29
3-4	27	38
5+	8	15
Number of cases	164	430

respect to the poverty level, now measured in terms of income per capita (45 quetzales/year) to control for the effect of family size.

The differentiating characteristics of poverty emerge clearly in Table C-15: low levels of education (or at least relatively lower), small land-holdings and greater proportion of family-farms as opposed to commercial farms, both cultivating land owned by the family or by share-cropping. Meanwhile, poor families have fewer children than "all the others" (1.25 and 2.15 respectively), which may be a consequence of their own poverty or a slightly higher average age of household head.

Appendix C-2

The Villages

Available information allows characterization of the four rural villages for a single year, 1974. Because of that, some features cannot be generalized. For example, one cannot be sure whether food staple sales are effective most of the years or whether 1974 was an exceptionally good year in terms of crop harvest enabling farmers to sell food crops. Other characteristics as land size, specialization on cash crops, and performance of non-farming activities can be seen as more permanent indices less dependent on the combined seasonal and cyclical influences of weather, amount of precipitation, and so on. The description below focuses essentially on the characterization of the population of the four villages in terms of their main occupation, sources of income, farming activities, and family variables as family size and education variables. It complements descriptions of the villages found in other sections of this report.

The collection of data available in the Rand-Rockefeller files concerns 818 families in the four villages. Information for farming activities or other income sources is available for 737 families only, and exclusion of cases with missing values in many farming variables reduced the number of total cases to 716 cases. Among them, 598 were agricultural families having access to land; the remaining cases were families whose members worked as wage labor or carried out other activities as small trade and craft work. Given our interest here in analyzing the influence of education in farm production, the present description will be mostly concerned with the 598 farm families.

Table C-16 below depicts some elements for the total sample and for each village considered here. As is seen, only 17.3 percent of

the families depend entirely on farming income. Salaried work is carried out by 70 percent of the families to complement farming income. On the average, wage income accounts for 40 percent of total income accruing to the average farmer in these villages.

TABLE C-16

INCOME SOURCES BY VILLAGES

Villages	All	#3	#6	#8	#14
STRUCTURE OF CASES (%)					
Agriculture only	17.3	14.7	27.7	5.4	17.3
Agric. & Salaries	48.2	65.9	56.4	13.8	48.2
Agric. & Activities	12.6	9.4	5.3	26.2	13.6
Agric., Sal., and Act.	21.9	10.0	10.6	54.6	20.9
TOTAL	100.0	100.0	100.0	100.0	100.0
STRUCTURE OF INCOME (%)					
Agriculture Income	45.5	32.3	53.6	53.9	36.7
(Consumed)	(24.4)	(24.8)	(21.6)	(26.8)	(26.6)
(Sold)	(21.1)	(7.4)	(31.7)	(27.1)	(10.1)
Non-Agriculture	54.5	67.8	46.4	46.1	63.3
Wage Income	39.8	55.6	34.6	31.2	36.7
Others	14.7	12.2	11.8	14.9	26.6
Number of cases (farmers)	598	170	188	130	110
% Total Cases in Villages	83.6%	85%	84.7%	74.3%	92.4%
Total Income	550.8	531.4	662.7	489.2	465.4
Income per Capita	131.9	136.7	167.5	111.4	88.7

The weight of non-farming income on total income varies among villages: Villages 6 and 8 show 54 percent of income coming from agricultural sources, while in Villages 3 and 14, corresponding values are 32.3 and 36.7 percent

respectively. This is related to other indices concerning land productivity and farm produce. On the one hand, one sees (Table C-17) that Villages 6 and 8 have higher levels of land productivity, use larger amounts of cash inputs per land unit, and the weight of food staples to total produce is considerably lower than in Villages 3 and 14. The percentage of sold farm output is consequently higher in Villages 6 and 8 as well.

TABLE C-17

INDICES FOR AGRICULTURE PRODUCTION BY VILLAGE

Village	Total	Vil. 3	Vil. 6	Vil. 8	Vil. 14
Average extension of land per family (cuerdas)	68.6	97.23	83.17	36.8	70.3
Average size of parcel (cuerdas)	27.5	28.70	37.7	15.3	23.5
Average number of plots	2.5	2.90	2.1	2.3	2.8
Average price of land (quetzales/cuerdas)	2.0	1.13	2.7	3.7	.9
Average productivity of land (quetzales/cuerdas)	4.79	3.0	4.81	8.85	3.16
% of corn and beans in total agriculture output	55%	84%	50%	33%	74%
% of corn and beans sold (out of respective output)	16%	14.4%	20.4%	7.7%	14.2%
% agriculture output sold	45.2%	23.7%	53.6%	54.9%	27.2%
% agriculture output not sold	54.8%	76.3%	46.4%	45.0%	62.8%
Labor-Land ratio	2.9	3.08	3.24	2.5	2.65
Fertilizer-Land ratio	.34	.16	.39	.55	.19
Hired Labor-Total Labor	21.1%	11.97%	19.9%	47.9%	10.1%
Productivity of Labor (Gross)	1.15	.95	1.14	1.64	.99
Productivity of Labor (Net)	1.04	.90	1.02	1.42	.92
% Families Using Hired Labor	38.6%	34.1%	43.0%	47.8%	28.4%
Average Wage	.84	.75	.85	.87	.84

The average size of land plots is quite small, 1.3 hectares. Families hold, however, on the average, more than one plot which makes the average landholding equal to 2.8 hectares (79.9 cuerdas).

Agriculture production consists mainly of five crops: corn and beans (food crops), tomato and chili (cash crops), and maicillo (for animal feeding). Other products as ayoit and yuca are less important. Table C-18 depicts the structure of production and sales by villages as observed in 1974.

TABLE C-18

CHARACTERIZATION OF FOUR GUATEMALAN VILLAGES
PRODUCTION AND SALES OF FIVE MAJOR CROPS

Village	Families (%)					Weight (%) Total				
	All	#3	#6	#8	#14	All	#3	#6	#8	#14
Crops Grown:	%	%	%	%	%	%	%	%	%	%
Corn	95.0	89.4	97.3	84.0	97.3	42.2	70.8	33.9	32.4	48.6
Beans	51.0	51.3	73.4	12.3	49.1	12.7	13.3	16.3	.5	25.3
Tomato	15.2	2.1	44.1	--	3.6	14.1	1.8	30.6	--	2.5
Chili	7.0	--	12.8	13.1	.9	20.0	--	16.7	48.5	1.8
Maicillo	19.1	30.7	2.1	.8	46.4	3.1	8.7		0.0	11.6
Others	21.7	13.2	14.4	41.5	21.8	7.9	5.4	2.2	18.5	10.1
Total	100.0 (598)	100.0 (189)	100.0 (188)	100.0 (139)	100.0 (110)	100.0 (300.2)	100.0 (201.2)	100.0 (419.5)	100.0 (355.2)	100.0 (184.2)
	Farmers[b] (%)				SALES	Sold/Produced[c] (%)				
Village	All	#3	#6	#8	#14	All	#3	#6	#8	#14
Crops sold:[a]										
Corn	32.0	37.0	42.6	32.1	7.0	12.2	14.6	14.4	8.1	6.2
Beans	25.5	9.5	34.8	12.5	35.1	28.5	11.5	33.3	6.3	31.1
Tomato	98.0	75.0	100.0	--	100.0	97.6	98.9	97.6	--	88.7
Chili	100.0	--	100.0	100.0	100.0	75.0	--	99.9	60.0	100.0
Maicillo	41.2	57.0	.7	--	27.0	32.9	61.8	6.7	--	18.9
Others	100.0	100.0	100.0	100.0	100.0	100.0	100.0	100.0	100.0	100.0
Total	71.4 (363)	56.8 (96)	73.0 (136)	62.0 (79)	100.0 (52)	46.4 (110.7)	20.6 (81.4)	29.4 (343.3)	50.3 (293.9)	27.4 (107.2)

a) Value of sales evaluated at the average selling prices.
b) Percentage of farmers selling the crop out of farmers growing it.
c) Percentage of sales on production of farmers selling the crop.

Most of the farmers grow corn, 32 percent of them selling, on the average, 12.2 percent of their production. Being the major source of food in these villages, corn production represents 42.2 percent of total agricultural production attaining its highest level in Village 3 where it represents 70.8 percent of total farm production. Beans, the major source of protein in villages, has the second place in agricultural production. Roughly half of the farmers grow some beans; among these, one fourth sell some bean production (on the average, 29 percent). Tomato and chili grown by fewer farmers (15 and 7 percent respectively) are essentially grown for sale. Maicillo grown by 19 percent of the farmers represents 3 percent of the total average produce only. It is, however, sold by 40 percent of those who cultivated it, who, on the average, sold 33 percent of their maicillo crops. Twenty percent of the farmers plant other crops for sale which represent around 8 percent of total produce.

Sales and salaried work yield a substantial amount of cash income which makes these populations very different from the subsistence level farmers of Nakajima-Fisk model. Adding sales, salary income, and income from other activities (such as small trade, arts and crafts) makes cash income represent, on the average, 75 percent of total income.

Hiring labor seems to be a common practice in most of the villages: 38.7 percent of the farmers hire some labor, Villages 6 and 8 being those where hired labor is more heavily used (43 and 48 percent of farmers respectively). In Village 8, total hired labor represents 48 percent of total labor input while, in the remaining villages, the average is less than 20 percent. This is consistent with the fact that Village 8 has the largest proportion of landless and non-farming families (25.7 percent).

Finally, education and family variables are depicted in C-19. As observed, both family size and average farmer are quite similar across

villages. Education differs considerably among villages, however; for all, the average of educational grades is very low. In Village 6, we see low rates of literacy and grade attainment, although these male household heads report more literacy. On the average, the number of grades of males is less than females. The different conclusions yielded by looking at educational grades and literacy indicate that there is not a consistent relation between literacy and grades; e.g., in Villages 3 and 6, male household heads have less schooling but are likely to read and write better than females; conversely, Village 3 is the one with the highest rate of male literacy, but it is Village 8 where males have the largest amount of

TABLE C-19

EDUCATION AND FAMILY VARIABLES BY VILLAGE

Village	All	Vil. 3	Vil. 6	Vil. 8	Vil. 14
Family	4.65	4.57	4.69	4.36	5.12
N. of people present	4.60	4.48	4.58	4.29	5.00
N. of minors	1.38	1.53	1.35	1.20	1.43
Age mother	35.5	37.9	32.8	36.7	34.6
Age father	34.2	34.1	35.8	31.2	35.5
Read and write:[a]					
Mother	.39	.57	.25	.35	.39
Father	.32	.98	.39	.29	.33
Av. Family[b]	.80	.88	.70	.91	.68
Education - grades					
Mother	.72	.90	.55	.75	.65
Father	.54	.43	.55	.76	.39
Av. Family[b]	1.27	1.35	1.04	1.68	1.06

[a]Average of points according to: 0 - if cannot read and/or write
1 - if can read and/or write a little
2 - if can read and write well

[b]Average computed for the members of the family aged 7 or more.

schooling (although less than one year of schooling). Ability to read and write was collected by asking whether a person could read and write rather than by testing. Indices were computed as an average of three values -- 0 if unable to read and write, 1 if reads and writes a little, and 2 if reads and writes well. When we regrouped into just two categories using 0 for the two first cases and 1 for the case when the person reads and writes well, the results are more striking as far as village differences are concerned.

TABLE C-19a

MALE LITERACY USING A 0, 1 SCALE

Villages	All	#3	#6	#8	#14
Male's Literacy	.366	.362	.164	.525	.258

This confirms the relatively higher level of literacy in Village 8 and the lowest position of Village 6. Recalling that Villages 6 and 8 were the ones where land and labor productivity levels were the highest, one concludes that it is impossible from these simple indices to reach any conclusions on the size and direction of the influence of education on agricultural production. More complete analysis, thus, needs to be carried out to disentangle the effects of education and other factors on the productivity levels of these farmers in Guatemala. Before turning to this more complete treatment, we shall present a short discussion on poverty in the villages as background for the analysis.

Appendix C-3

For example, for a production function as $y=f(l,t,i)$, if $f(.)$ is homogeneous of the first degree, $y=t.g(l/t, i/t)$ and marginal product of labor, dy/dl, is given by $dg/d(lt)$, holding i/t constant. Assuming land constant in the short run, optimality conditions for both farm groups imply that firms will operate at the same level of labor marginal product which, ultimately, will be the wage rate practiced in the village, and at the same labor-land ratio $(l/t)^*$, for $g(.)$ is the same across farms. The optimal labor-land ratio, $(l/t)^*$, will be such that

$$g'(l/t_i)^* = g'(l/t_j)^*$$

for all i and j firms, i and j denoting farms. With land fixed in the short run, $t_i = \bar{t}_i$, optimal labor inputs will be determined as

$$l_i^* = \bar{t}_i(l/t)^*$$

The amount of hired sold labor services would then depend on the amount of labor available in the family, l_i^f, and the optimal amount l_i^*. The farmers will hire labor if l_f is less than l^* and will sell otherwise.

Since groups differ in wage perceived in the labor markets, such as literate farmer members obtaining, on the average, wages 60 percent higher than the illiterate, the amount of family labor available for farming activities will be determined in the function of this wage w_i^*, $L_f^*(w)$: $dy/dl_f^* = w_f^*$. Optimal labor-land ratio will be a function of the minimum wage $\bar{w}$,

$$(l/t)^*: \quad dg(l/t)^* = \bar{w}$$

and labor inputs by $l^* = (l/t)^*.\bar{t}$. Labor will be hired or sold whenever labor family is less or greater than the optimal l^*:

$$l_h(s) = l^* - l_f^*(w^*) = t.(l/t)^* - l_f^*(w^*)$$

REFERENCES

Fisk, E. K. "Planning in a Primitive Economy: From Pure Subsistence to the Production of a Market Surplus." Economic Record 40 (1964).

Fisk, E. K. "The Response of Nonmonetary Production Units to Contact with the Exchange Economy." Agriculture in Department Theory. New Haven: Yale University Press, 1977.

Griliches, Z. "Estimating Returns of School: Some Economic Problems." Econometrica 45 (No. 1) (January 1977): 1-22.

Jamison, D., and Lau, L. J. Farmer Education and Farm Efficiency. Baltimore: John Hopkins University Press, forthcoming.

Lau, L., and Yotopoulos, P. "A Test for Relative Efficiency and Application to Indian Agriculture." American Economic Review 61 (1971): 94-109.

Lockeed, M.; Jamison, D.; and Lau, L. "Farmer Education and Farmer Efficiency -- A Survey." Economic Development and Cultural Change, forthcoming.

Nakajima, C. "Subsistence and Commercial Family Farms -- Some Theoretical Models of Subjective Equilibrium." Subsistence Agriculture and Economic Development. Chicago: Aldine, 1969.

Reynolds, L. G. "Agriculture in Development Theory: An Overview," in Agriculture in Development Theory. New Haven: Yale University Press, 1977.

Schultz, T. "The Value of the Ability to Deal With Development." Journal of Economic Literature 13 (1975): 872-876.

Sen, A. K. "Peasants and Dualism With and Without Surplus Labor." Journal of Political Economy 74 (October 1966): 424-450.

Yotopoulos, P. and Lau, L. J. "A Test for Relative Economic Efficiency: Some Further Results." American Economic Revue 63 (1973): 214-23.

II-D. Education and Fertility in Rural and Semi-Urban Guatemala.[1]

by Mari S. Simonen

Introduction

This report summarizes analyses of socio-economic and attitudinal determinants of fertility in rural and semi-urban Guatemala. The focus is on the effects of schooling and literacy of the female and male head of household on their fertility desires and actual parity.

The question examined is whether schooling and/or literacy of the female and male head of household have a significant effect on fertility, and, if so, how and why they affect fertility in the context of the communities under study. In addition, we will examine what other factors explain fertility at the individual level in the communities studied. The reason underlying this work is the need to understand better why people in less developed countries are having a given number of children, and, consequently, to be able to design more responsive and effective population policies. It is not assumed here, however, that decreasing fertility is automatically a desirable policy goal. Rather, decreasing fertility may be desirable only if it increases the welfare of the family.

Fertility is defined here as the number of children ever-born and as the number of desired children. Emphasis is thus on the 'stock' aspect of fertility, although fertility desires are also analyzed from the point of view of sequential fertility decision-making. The data used are from a series of cross-sectional household surveys administered

---

[1] Note that data used in this section are from the four rural villages of the INCAP data base as well as from two semi-urban villages (Rand/Rockefeller data base). For a more comprehensive report of results presented here, see Appendix 4.

in 1975-76 in the same four rural communities where INCAP (Instituto de Nutricion de Centro America y Panama) had implemented its nutrition experiment since 1969, and in two semi-urban communities nearby Guatemala City. All six communities are part of the Ladino, rather than the Indian, culture.[1]

Theoretical Framework and Methodology

The analytical work of this study is centered around estimation of a structural equation model of individual fertility determinants. The model is displayed in diagram 1 in which postulated causal relationships among variables are represented by unidirectional arrows from each determining variable to each variable dependent on it. Postulated non-causal correlations between exogeneous variables are represented by curvilinear lines, and residual variables are represented by unidirectional arrows.

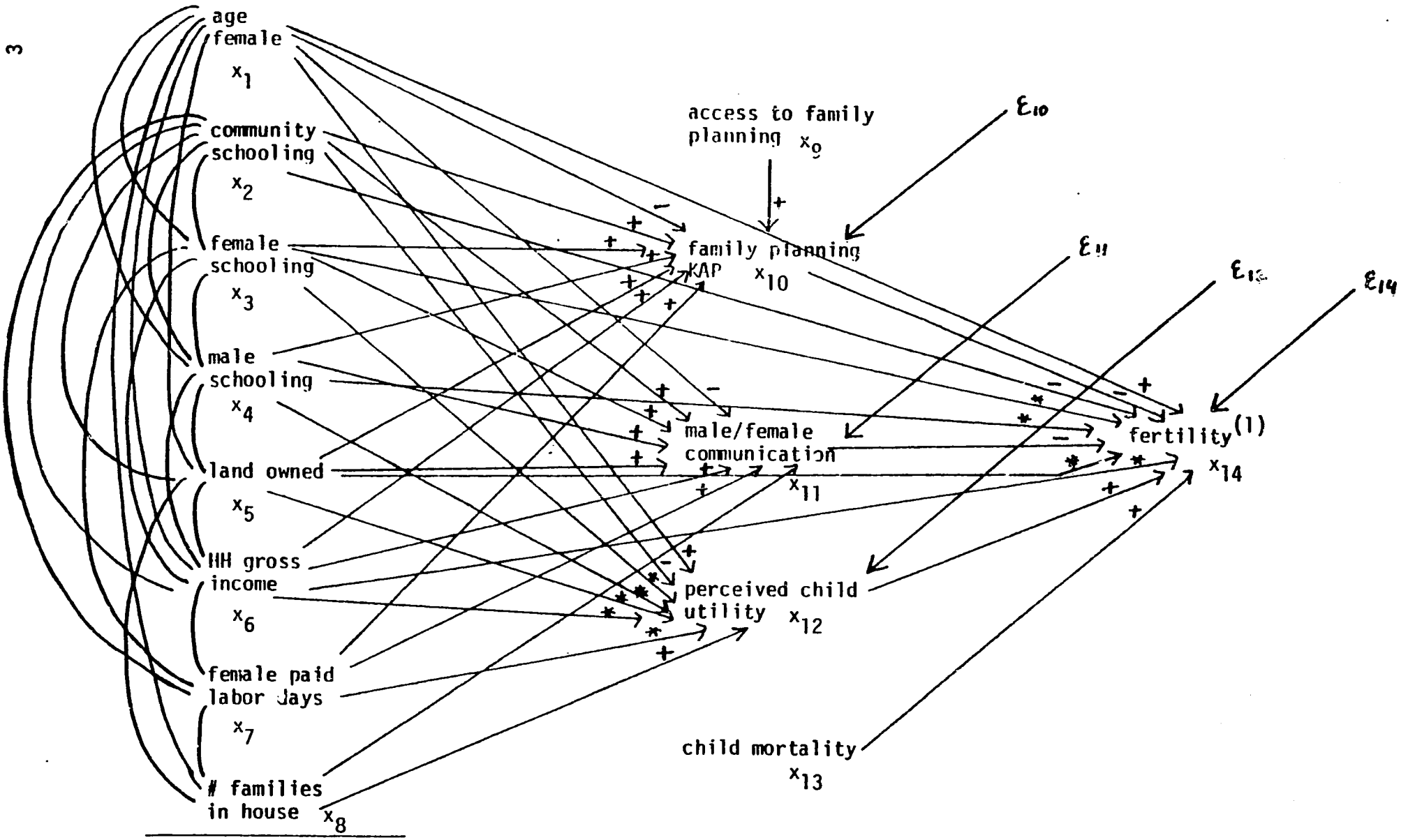
The model is based on extensive conceptual work with respect to different theories of fertility determinants. Only a brief summary of the model is presented here.[2] It is hypothesized in the model in diagram 1 that individual female and male schooling affect fertility, together with community level schooling, family planning access, landownership, income, female labor force participation, family type and

---

[1] They speak Spanish and identify more with the Spanish as compared with the indigenous culture.

[2] For a more complete review of the theoretical framework, see Simonen, M., "Schooling and Fertility in Rural and Semi-Urban Guatemala. A Cross-Sectional Study of Six Ladino Communities." Berkeley Project on Education and Nutrition, University of California, Berkeley, January 1980. For a comprehensive overview of empirical evidence linking education to fertility in LDC's, see Cochrane, S. (1978), Education and Fertility, What Do We Really Know?.

Conceptual Model of Fertility at the Family Level



child mortality, mainly through three intervening variables: (a) family planning knowledge, attitudes and practice (KAP), (b) male/female communication, and (c) child utility.

The theories and findings in the field of family planning emphasize that schooling along with other socio-economic variables, on one hand, improve access to family planning information and services, and, on the other hand, enable the person to make more informed decisions about family planning/contraception. Also, schooling may contribute to more positive attitudes toward the idea of limiting family size (Berelson 1966, Cochrane 1978). Moreover, some theoretical and empirical evidence support the hypothesis that family planning KAP has a negative effect on fertility (CELADE 1972, Sear 1975, Williams 1976, Cochrane 1978). However, the theoretical evidence is not well developed. The argument is simply that favorable attitudes toward knowledge about and use of contraception help parents control fertility (Williams 1976). Moreover, there is also accumulating evidence that family planning KAP may be a significant influence on fertility only, once the motivation to have no more children has been reached, i.e., generally after the family has the number of children (or boys) judged necessary to assure a given number of surviving children when parents attain retirement age (Wyon and Gordon 1971, McGreevey and Birdsall 1974, Cochrane 1978).

Social psychological theories of interpersonal decision-making emphasize the importance of social exchange processes within the family as a determinant of family size and fertility (Hill, et al. 1955, Bagozzi and Van Loo 1978). These theories assume further that schooling along with other socio-economic variables operate mainly through this intervening variable, and, although neither theory nor empirical evidence

yield definite clarification as to the direction of effect of schooling on social exchange within the family, both tend to support a positive relationship (Cochrane 1978, see Chapter 5).

Economic, sociological, anthropological, and psychological theories of fertility emphasize the role played by values of children in determining fertility (Davis 1955, Leibenstein 1957, Hoffman and Hoffman 1973, Cassen 1976, Nag in Marshall and Polgar 1976). While psychologists tend to emphasize the psychological value of children to parents, economists emphasize their economic/utilitarian value (contribution to income and as old age security), and sociologists emphasize how the perceived value of children reflects long-term societal value of children. These theories further hypothesize that societal level variables as well as individual schooling, wealth, and production pattern affect child utility, i.e., as the socio-economic factors improve, the values (utilities) of children are affected, in that economic benefits tend to decrease while both direct and indirect costs tend to increase. Theoretical work on the effects of type (mode) and relations of production on fertility sees fertility largely as a reflection of demand for family labor which, in turn, depends on mode and relations of production of family/household (McFarlane 1978, Deere and deJanvry 1978).

While it is hypothesized in the fertility model in diagram 1 that the effect of the socio-economic variables on family planning KAP and male/female communication and that of these, two intervening variables on fertility are in the same direction in the rural as in the semi-urban area, the effects of schooling, land-ownership, income and female labor force participation on perceived child utility and on fertility are hypothesized to vary by rural vs. semi-urban area.

It is hypothesized that in the rural area, the net direct and indirect effect on fertility of the socio-economic variables mentioned above is positive, whereas in the semi-urban area, it is negative. Reason for the differential effect is the expected greater increase in 'income effect' than in 'price effect' in the rural area (i.e., parents have an increased capacity to support children and may need them more for they have more land, whereas the reverse is true in the semi-urban area) (Williams 1976). The direct effects of the socio-economic variables on fertility represent these variables' potential effects on fecundity (not explicitly modeled in the diagram) and on attitudinal variables not explicitly modeled in the diagram (educational aspirations for children and consumption aspects of children, for example).

The effect of family type on fertility (nuclear vs. extended) is hypothesized to operate through perceived child utility and male/female communication. It is hypothesized that a move towards a nuclear family decreases the net value of children due to an increase in their costs and decrease in value (Leibenstein 1957, Schultz 1969, Simon 1974). The direction of effect of family type on male/female communication is left unspecified due to limited theoretical and empirical evidence.

And last, child mortality is hypothesized to have a positive direct effect on fertility following the theoretical work of economists and others on child survival hypothesis and replacement behavior (Schultz 1976, Preston 1978). Although a controversy exists about how much child mortality affects fertility, ample evidence supports a positive relationship (CICRED 1974, McGreevey and Birdsall 1974). The reason for modeling child mortality as an exogeneous rather than as an intermediate variable

is mainly a practical one, i.e., limits had to be set on the focus and number of relationships examined in this study.

Methodology

The main analytic technique used in this study is structural equation model estimation by means of ordinary least squares regression method, path analysis (Duncan 1975).[1] The model in diagram 1 can be written out as the following set of structural equations:

$$x_{10} = \alpha_{10} + \beta_{101}x_1 + \beta_{102}x_2 + \beta_{103}x_3 + \beta_{104}x_4 + \beta_{105}x_5 + \beta_{106}x_6 + \beta_{107}x_7 + \beta_{109}x_9 + \varepsilon_{10}$$

$$x_{11} = \alpha_{11} + \beta_{111}x_1 + \beta_{112}x_2 + \beta_{113}x_3 + \beta_{114}x_4 + \beta_{115}x_5 + \beta_{116}x_6 + \beta_{117}x_7 + \beta_{118}x_8 + \varepsilon_{11}$$

$$x_{12} = \alpha_{12} + \beta_{121}x_1 + \beta_{122}x_2 + \beta_{123}x_3 + \beta_{124}x_4 + \beta_{125}x_5 + \beta_{126}x_6 + \beta_{127}x_7 + \beta_{128}x_8 + \varepsilon_{12}$$

$$x_{14} = \alpha_{14} + \beta_{141}x_1 + \beta_{142}x_2 + \beta_{143}x_3 + \beta_{144}x_4 + \beta_{145}x_5 + \beta_{146}x_6 + \beta_{1410}x_{10} + \beta_{1411}x_{11} + \beta_{1412}x_{12} + \beta_{1413}x_{13} + \varepsilon_{14}$$

where α = constant term

x_i = nonstochastic variables

β = structural equation parameters

ε_i = stochastic disturbance term

---

[1] In addition, multiple classification (MCA) analysis and logit analysis (maximum likelihood estimation of qualitative choice models) are used (Andrews 1973, Nerlove and Press 1973).

The model is estimated for female respondents aged 14+ living in union with a male at the time of the interview.[1] It is estimated separately by rural and semi-urban stratified samples.

Structural equation methodology of path analysis relies on ordinary least squares estimation and thus assumes linearity and additivity among variables. Since OLS estimation is used, the basic OLS assumptions are made: variance of the disturbance term is constant, and exogeneous variables and the disturbance term of a given equation are uncorrelated. In addition, in order for OLS procedures to yield unbiased and efficient estimators of the path coefficients, it is assumed that the stochastic disturbance terms in a recursive system of equations are uncorrelated. Also, path analysis makes an explicit assumption of causal ordering among the variables. In order to obtain unbiased estimates for the structural equation parameters, the model has to be rightly specified.

Since the model to be estimated is recursive,[2] the parameters can be uniquely determined from observed data, i.e., the model is identified.

The major advantage of structural equation models is that they allow the estimation of <u>total</u> effects[3] of several variables on the main endogeneous variable while controlling or holding constant the effects of other variables in the model, whereas multiple regression reduced form, for example, allows examination of direct effects only.

Since cross-sectional data is used, one should be cautious about the interpretation of the observed effects. This is especially relevant

---

[1] Includes both consensual and legal unions.

[2] All causal linkages run 'one way', i.e., no two variables are reciprocally related in such a way that each affects and depends on the other, and no variable 'feeds back' upon itself (Duncan 1975).

[3] Total effect = combined direct and indirect effects.

in the case of the attitudinal variables in this study, for they were collected contemporaneously with the fertility information. In addition, a word of caution concerns the estimation of the hypothesized child mortality-fertility and the female labor force participation-fertility relationships. It has been pointed out that there may be two-way interactions in the child mortality-fertility relationship (CICRED 1975). Similarly, it is recognized that fertility may affect female labor force participation.

Sample and Variables Used

The sample consists of a total of 847 female respondents (578 rural and 269 semi-urban) aged 14+ at the time of interviews in 1975-76. These 847 respondents were living in union (consensual or legal) with a male at that time or had been in union previous to 1975-76. Note, however, that the models presented in this report are estimated only for females who were actually living in union with a male at the time of the interviews. The unit of analysis is the individual female head of household interacting with the male head of household, the rest of the family and household, and the community of residence.

Below, the main exogeneous and endogeneous variables are briefly described:

Female age is measured as a continuous variable. It is used as a control for the effect of age on the endogeneous variables. In the case of fertility, female age is used as a proxy for exposure to pregnancy since data are not available on duration of union.

Urbanization is measured as a dichotomy denoting rural vs. semi-urban residence.

Child mortality is measured in terms of the proportion of own children that have died out of the total number of live births. The data do not allow specification of infant vs. child mortality.

Female and male education is measured in terms of number of grades of formal schooling completed and literacy. Literacy is a self-report about whether one reads and writes well, a little, or not at all. It is used as a dichotomy where zero is assigned to the two first answer categories (reads and writes a little or not at all) and one to the last category (reads and writes well).

Income is per capita income derived from household total income. Household total income is the sum of (1) total agricultural production in average 1974 prices; (2) total salaries and wages earned by all household members in 1974; and (3) net income earned by all household members from activities in 1974. Unit of measurement is a quetzal (quetzal = 1 dollar).

Land-ownership is measured as total land owned by the nuclear and the extended family. Unit of measurement is a 'cuerda' (cuerda = .044 hectares).

Female labor force participation is measured in terms of total time (hours) worked for pay during 1975-76. It thus measures exclusively paid female work and is a measure of the degree to which a female actually participates in the labor market outside of home.

Family type is measured as the number of nuclear families living in and sharing the same house. It thus measures the territorial aspect of family type but excludes the functional aspect (i.e., families living apart but sharing income).

Type of economic production is measured as a categorical variable: (1) wage-laborer, (2) subsistence farmer, (3) semi-commercial farmer, and (4) commercial farmer. A family is defined as 'wage-laborer' if it has not

reported agricultural production and thus derives its income from salaries and wages and non-agricultural activities. Subsistence farmer has some agricultural production but does not sell any of it nor hire labor; semi-commercial farmer sells some of its agricultural production but does not hire labor, and commercial farmer both sells agricultural production and hires labor.

Family planning knowledge, attitudes, and practice (KAP) is measured as a simple additive scale of three items: (a) whether the respondent thinks that women should have all children that come or do something to limit the number; (b) knowledge of a specific way to limit the number of children coded as [0] no knowledge, [1] knowledge of a 'traditional' method, and [2] knowledge of a 'modern' method;[1] and (c) whether one practices a method or not. The scale ranges from 0 to 4, 0 indicating a low pro-family planning and 4 a higher pro-family planning position.

This family planning KAP scale is a measure of the respondent's overall family planning position. It does not measure the intensity of family planning KAP within the 'traditional' and 'modern' dimensions.

Male/female communication is measured as a simple additive scale of four items. Two are the female respondent's report on whether she has talked and agreed with spouse about the number of children desired, and the other two are comparisons of male and female respondents' answers to a question about how many additional children they want. This scale ranges from 0 to 4 with 0 indicating low and 4 higher male/female communication.

This scale focuses on one aspect of intra-couple communication -- namely the family size preference and family planning aspect.

---

(1) 'Traditional' = herbs, rhythm, lactation, withdrawal. . .
'Modern' = pill, condom, diaphragm, going to family planning clinic. . .

Perceived child utility is measured by means of questionnaire items on female respondent's perceptions about her children's utility in general and as old age security. The basic idea is to measure how many times the respondent mentions children during the interview,[1] and the assumption is that respondents who mention children to be useful in several different contexts (and questionnaire items) perceive and desire greater 'utility' from children than respondents who mention children fewer times or not at all.

The child utility scale is derived from 10 items and ranges from 0 to 10 in value, with 0 indicating a low and 10 a higher perceived utility.

Fertility is measured as desired parity and children ever-born. Desired parity is measured as a continuous variable, and it is derived from the female respondent's answer to the item: "In addition to the children you already have, how many more do you want?"

---

[1] This scale is adapted from a similar scale developed by Dr. Eva Mueller. See E. Mueller, "Economic Motives for Family Limitation. A Study Conducted in Taiwan," Population Studies, Vol. 26, No. 3, Nov. 1972.

Description of the Sample[1]

The four rural communities are located in a dry mountainous area in the Northeast of Guatemala City about two hours driving time from Guatemala City. These are primarily agricultural communities. Although they are fairly similar to each other, some important characteristics also differentiate the rural communities (see tables 1 and 3).

The two semi-urban communities are located at about 20 minutes driving time south from Guatemala City. Many inhabitants work in the capital city. Although both of these semi-urban communities share a lot in common with Guatemala City and each other, it should be noted that one of them is more agricultural than the other (see table 2).

Table 1 displays mean values of selected demographic and socio-economic variables describing the sample used in this study, and tables 2 and 3 describe selected community characteristics. Average female age in the rural sample is 37 years, while it is 34 years in the semi-urban sample. For males it is 41 years in the rural and 39 in the semi-urban sample. Average number of children-ever-born is 5.72 in the rural and 4.64 in the semi-urban sample while figures for completed parity[2] are 8.62 and 7.00 for rural and semi-urban, respectively. Child mortality rate is 19 percent in the rural and 14 % in the semi-urban sample.

About one fourth of the female respondents in the rural sample are

---

(1) For a more complete description, see : Berkeley Project on Education and Nutrition, "Education and Nutrition:Performance and Policy", Final Report, University of California,Berkeley, March 1980. Mari S. Simonen, "Education and Fertility in Rural and Semi-Urban Guatemala. A Cross-Sectional Study of Six Ladino Communities", University of California, Berkeley, Jan. 1980. Mejia Pivaral,"Caracteristicas Economicas y Socio-Culturales de Cuatro Aldeas Ladinas de Guatemala", Guatemala Indigena, vol.7,no.3, 1972.
(2)Children-ever-born for females aged 45-55.

TABLE 1

MEAN AND STANDARD DEVIATION OF SELECTED DEMOGRAPHIC AND SOCIO-ECONOMIC VARIABLES DESCRIBING SAMPLE

		RURAL COMMUNITY #					SEMI-URBAN COMMUNITY #		
		3	6	8	14	All	31	32	All
Female age	$\bar{x}$	39	35	38	36	37	34	35	34
	SD	15	12	15	14	14	10	14	13
	N	172	167	137	102	578	114	155	269
Male age	$\bar{x}$	42	39	42	41	41	38	40	39
	SD	12	11	15	13	13	9	128	12
	N	134	140	101	89	464	107	14	235
Family size	$\bar{x}$	5.11	5.25	4.84	5.54	5.17	5.60	4.90	5.20
	SD	2.28	2.28	2.28	2.30	2.28	2.10	2.02	2.08
	N	161	159	128	100	548	113	155	268
Children ever-born	$\bar{x}$	5.70	5.92	5.01	6.39	5.72	4.74	4.57	4.64
	SD	3.38	3.91	3.14	3.91	3.61	2.58	2.97	2.81
	N	172	167	137	102	578	114	155	269
Completed parity (age 45-55)	$\bar{x}$	8.13	8.86	8.11	9.71	8.62	6.50	7.29	7.00
	SD	3.37	3.55	3.85	4.16	3.67	2.37	2.14	2.22
	N	30	22	18	17	87	10	17	27
Desired parity	$\bar{x}$	4.56	4.74	4.09	5.81	4.65	3.86	3.57	3.69
	SD	2.87	3.22	2.52	3.87	3.07	2.06	1.82	1.92
	N	127	113	111	54	405	106	143	249
Child mortality rate	$\bar{x}$	17	23	17	20	19	13	15	14
	SD	22	23	22	21	22	19	20	20
	N	172	167	137	102	578	114	153	269

Table continued on next page

TABLE 1 (cont'd.)

		RURAL COMMUNITY #					SEMI-URBAN COMMUNITY #		
		3	6	8	14	All	31	32	All
Female schooling[1]	$\bar{x}$	1.27	1.03	1.09	.87	1.09	2.87	2.75	2.80
	SD	1.56	1.51	1.54	1.18	1.48	3.11	2.42	2.73
	N	160	155	128	100	543	110	150	260
Female literacy % literate	%	31	23	24	15	24	57	59	58
	N	50/162	36/158	31/128	15/100	132/548	63/110	89/151	152/261
Male schooling[1]	$\bar{x}$	1.18	1.13	2.05	.75	1.28	3.70	3.82	3.77
	SD	1.74	1.57	1.94	1.25	1.71	3.39	2.85	3.10
	N	119	126	95	84	424	102	125	227
Male literacy % literate	%	39	39	53	28	40	76	78	77
	N	52/133	54/139	54/102	25/90	185/464	81/106	98/126	179/232
Income per capita 1974-75	$\bar{x}$	110	130	95	82	107	251	180	210
	SD	106	135	119	78	115	191	148	171
	N	145	148	124	93	510	108	146	254
Land-ownership (cuerdas)	$\bar{x}$	57	50	10	64	45	.76	2.51	1.76
	SD	123	99	29	101	97	5	9	8
	N	147	150	126	94	517	112	150	262
Household economic production type %[2]	1	26	18	30	13	22	96	72	82
	2	25	18	23	37	25	2	12	8
	3	27	28	12	29	24	1	13	8
	4	22	35	34	19	28	1	2	2
Female paid labor hours 1975-76	$\bar{x}$	334	547	1352	327	635	868	751	800
	SD	949	1116	968	918	1077	1367	1214	1279
	N	172	167	137	102	578	112	154	266

[1] Number of grades of schooling completed.

[2] 1 = wage labor, 2 = subsistence farmer, 3 = semi-commercial farmer, and 4 = commercial farmer.

Table 2

Community Characteristics[1]

	RURAL	SEMI-URBAN	SEMI-URBAN	
			1	2
Median household income (quetzales) 1974-1975	555	1425	1473	1396
% income from agriculture	61	21	4	22
% income from wages	38	74	80	70
% income from activities	15	17	18	17
Principal occupation of male (%): agriculture	96	28	9	36
other	4	72	81	64
Principal occupation of female (%): none	68	62	62	62
domestic	3	12	12	11
manufacturing[2]	16	0	0	0
merchant	6	12	11	13
specialized laborer, factory, blue, and white collar	3	9	9	8
other	4	4	4	5
Community level schooling (grades)	2.53	6.65	6.49	6.79

(1) From census survey for whole community and not only sample of this study.

(2) Cottage-industries at home (weaving palm products, for example).

Table 3

DESCRIPTION OF SELECTED COMMUNITY CHARACTERISTICS BY DUMMY VARIABLES USED

	Community Dummy #					
	1	2	3	4	5	Reference Category
Area and Community #	Rural 3	Rural 6	Rural 8	Semi-Urban 31	Semi-Urban 32	Rural 14
Community level schooling[1]	2.53	2.08	2.95	6.49	6.76	1.72
Family planning access	no	yes	no	yes	yes	no
% females in cottage industry	0	0	68	1	0	0
% families growing cash crops[2]	2	57	13	---	---	4

[1] Average of grades of schooling, completed by male and female heads of household.

[2] Tomato and chile

literate, compared with over half of semi-urban females. The literacy figures are higher for males in both samples: 40 percent in the rural and 77% in the semi-urban sample report to be literate. For both females and males in both samples the average number of grades of schooling completed is low: 1.09 for females and 1.28 for males in the rural and 2.80 for females and 3.77 for males in the semi-urban sample. The average income per capita in 1974-75 just about doubles from rural to semi-urban sample: 107 vs. 210 quetzales.(1)

Tables 1,2, and 3 point to some important differences between the communities within the rural and semi-urban samples. For example, in table 1 the completed parity, child mortality and female literacy figures show differences between the rural communities while completed parity, family size, income and economic production type show differences between the two semi-urban communities.

Note in table 3 that community #6 differs from the other rural communities in that it is the only rural community having access to family planning, and, in it, over half of the families grow labor intensive cash crops (tomato and chili) as compared with about 5 percent in the other rural communities. Rural community #8 differs from the others in that it has the highest level of community schooling(2) and nearly two-thirds of female heads of household are engaged in cottage-

(1) 1 quetzal=1 dollar

(2) Measured as the community mean sum of male and female head of household's grades of schooling completed.

industries, mostly in the form of weaving palm products (Mejia Pivaral 1972).

Results

Path Analysis

Tables 4 through 8 display OLS regression results of estimating the fertility model in diagram 1 with female literacy as the measure of schooling. The model is estimated separately by rural and semi-urban samples, and two measures of fertility are used: children-ever-born and desired parity. Diagrams 2 and 3 display the results in the form of path diagrams.

Prior to the models presented here, analyses were undertaken to assess the effect of grades of schooling completed by female and male of fertility. Results of these analyses showed that, although the relationship between grades of schooling of female and male and fertility is negative, it is curvilinear, especially in the case of female schooling in the rural sample. Consequently, the fertility model in diagram 1 was estimated with a quadratic function of female schooling on children-ever-born. The results showed that female schooling measured in terms of grades of schooling completed has a statistically non-significant direct effect on children-ever-born in both areas, whereas its indirect effect--through perceived child utility--is significant

Diagram 2A

RURAL

Children Ever-Born Model
N=308

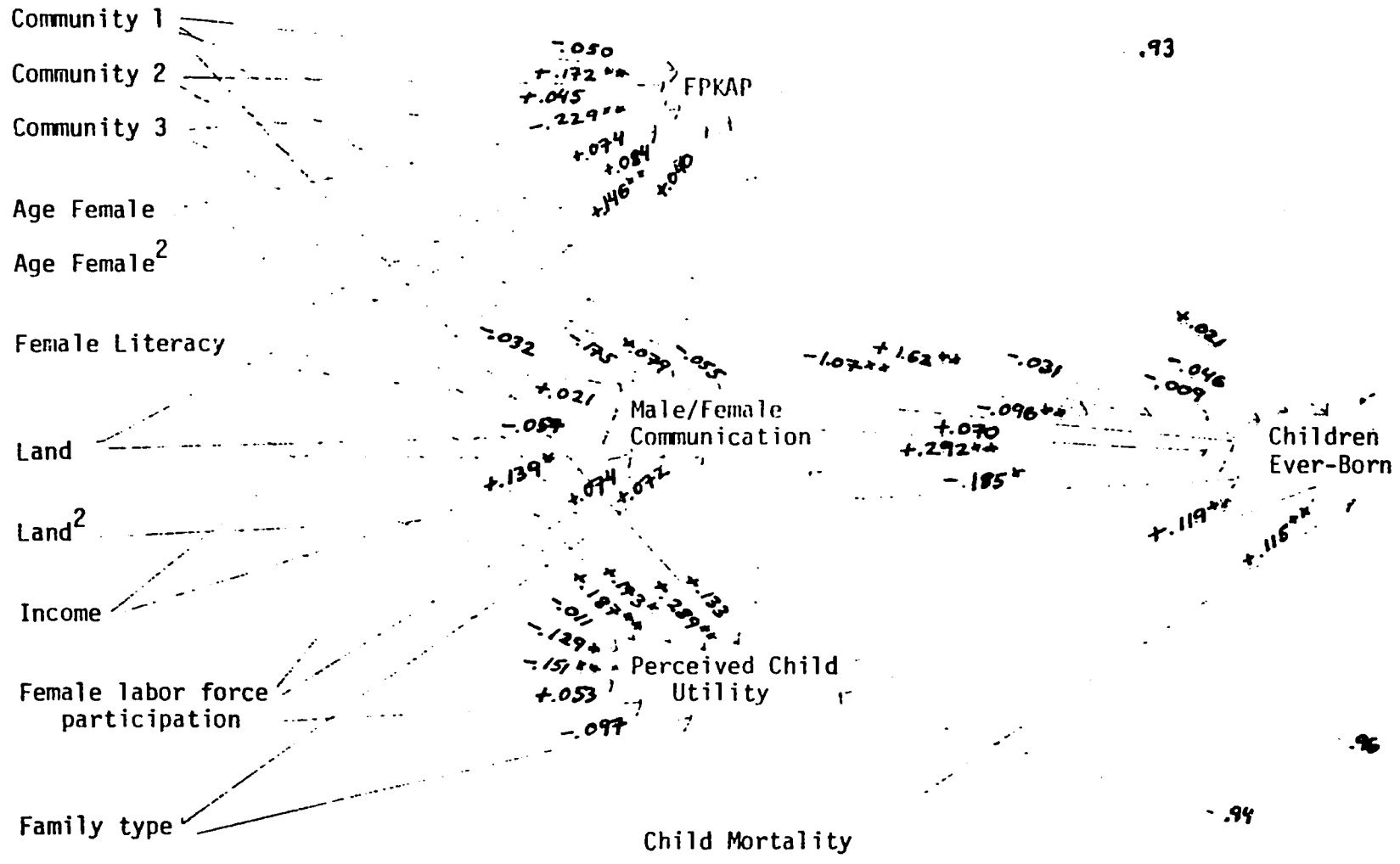


Diagram 2B

SEMI-URBAN

Children Ever-Born Model
N=91

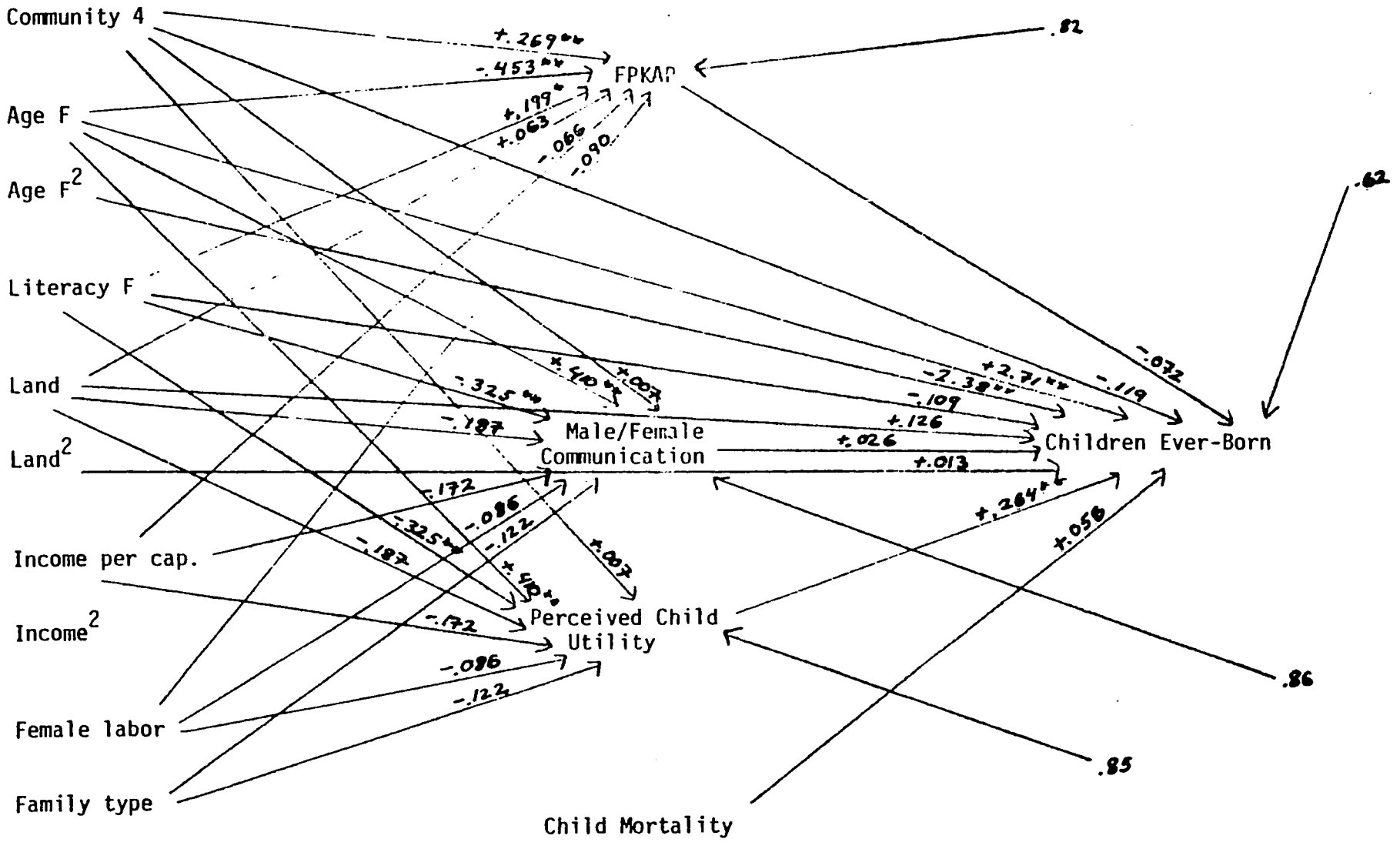
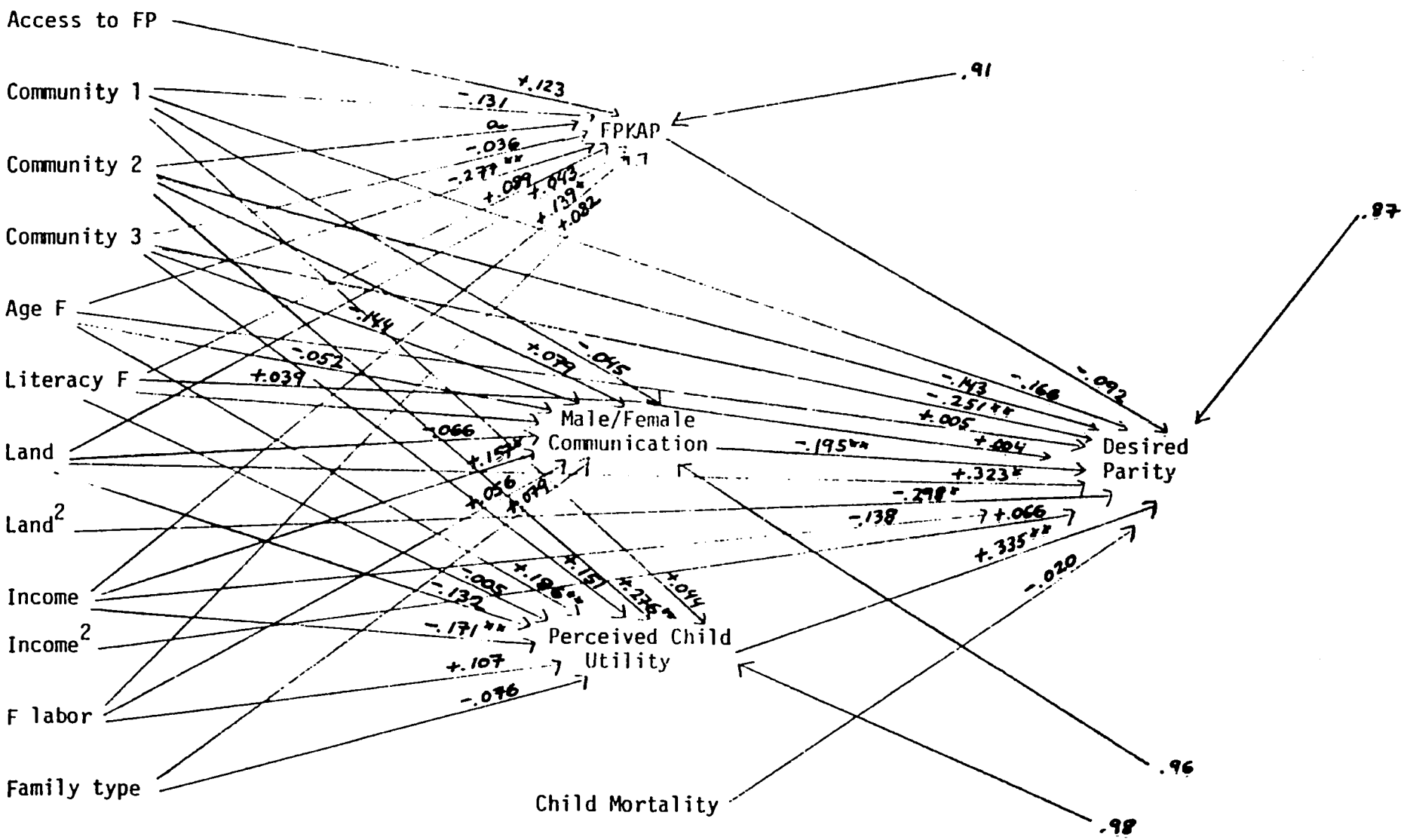


Diagram 3A[1]

RURAL

Desired Parity Model
N=214



(1) Parity equation includes both family planning KAP and male/female communication variables.

Diagram 3B

SEMI-URBAN

Desired Parity Model
N=85

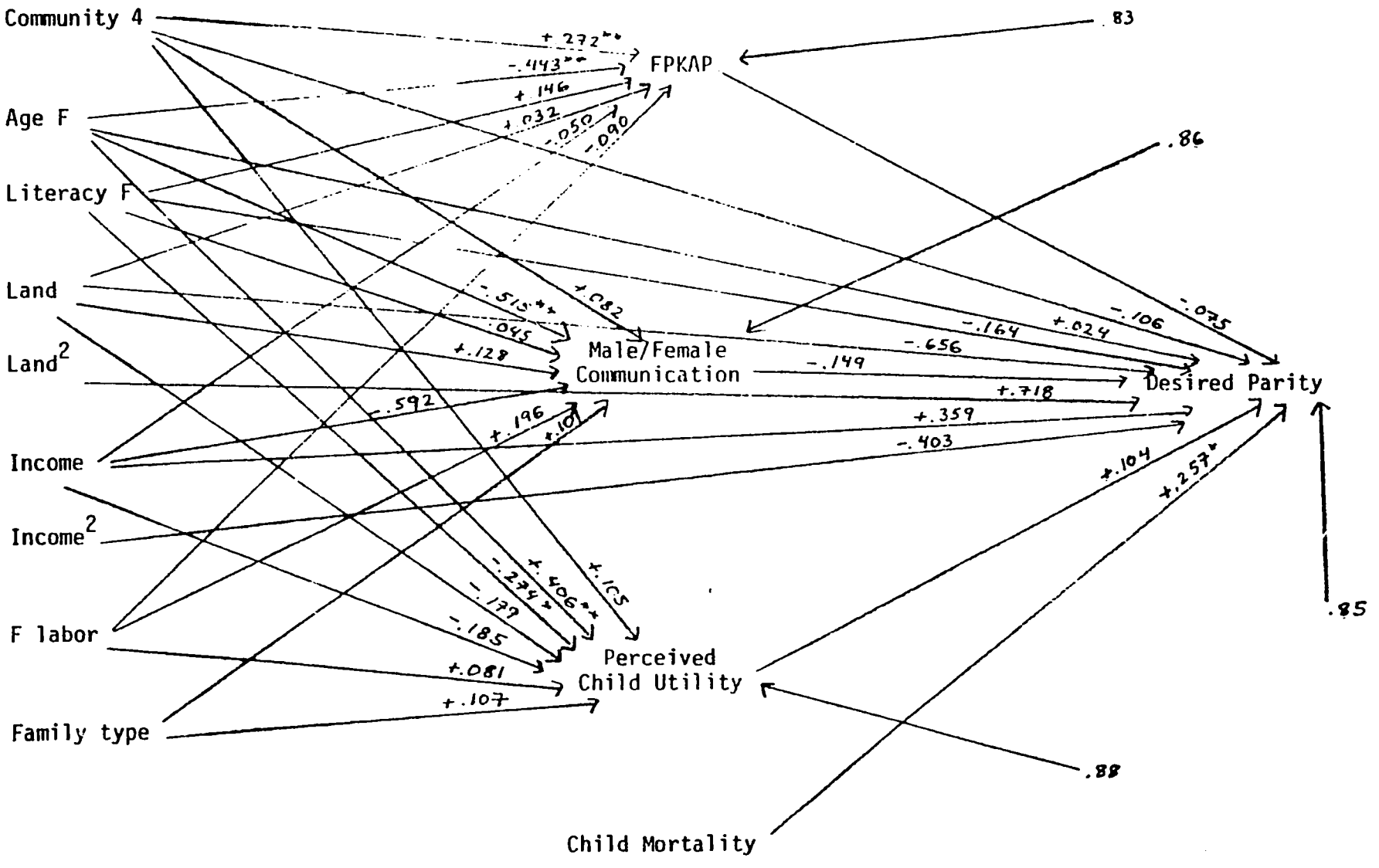


TABLE 4

MEAN AND STANDARD DEVIATION OF VARIABLES IN OLS ESTIMATION OF THE FERTILTIY MODEL:

	Desired Parity Model				Children Ever-Born Model			
	Rural N=214		Semi-Urban N=85		Rural N=308		Semi-Urban N=91	
	$\bar{x}$	SD	$\bar{x}$	SD	$\bar{x}$	SD	$\bar{x}$	SD
Age female	34	10	35	13	34	10	35	13
Literacy female	.26	.44	.61	.49	.25	.43	.58	.49
Income per capita	126	155	240	188	115	140	237	183
Land ownership	50	106	2.49	9	50	100	2	9
Female paid labor hours	592	1055	616	1190	535	1025	612	1192
Family type	1.31	.55	1.35	.67	1.29	.53	1.33	.65
Parity	---	---	---	---	5.90	3.35	4.92	2.56
Desired parity	4.98	3.24	3.76	1.68	---	---	---	---
Family planning KAP	1.68	1.40	2.89	1.29	1.43	1.36	2.80	1.33
Male/female communication	1.95	1.17	2.59	1.16	1.88	1.13	2.49	1.20
Perceived child utility	3.27	2.19	3.01	2.09	3.35	2.21	3.13	2.13
Child mortality	1.30	1.75	.89	1.20	1.31	1.71	.89	1.19
Village Dummy 1	.28	.45	---	---	.26	.44		
Village Dummy 2	.32	.47	---	---	.32	.47		
Village Dummy 3	.22	.41	---	---	.19	.39		
Village Dummy 4	---	---	.40	.49	---	---	.38	.49

Table 5

OLS ESTIMATION RESULTS. T-STATISTICS IN PARENTHESES.

Dependent Variable: Desired Parity(1)

	#1				#2				#3				Dependent Variable: Children Ever-Born			
	Rural N=214		Semi-Urban N=85		Rural N=214		Semi-Urban N=85		Rural N=214		Semi-Urban N=85		Rural N=308		Semi-Urban N=91	
	β	Beta	β	Beta	β	Beta	β	Beta	β	Beta	β	Beta	β	Beta	β	Beta
Age female	.002	.005	.003	.024	.012	.035	.005	.044	-.007	.022	.008	.068	.561**	1.62	.522**	2.71
	(.06)		(.16)		(.48)		(.31)		(.28)		(.47)		(7.30)		(5.84)	
Age female^2	---	---	---	---	---	---	---	---	---	---	---	---	-.005**	-1.07	-.005**	-2.38
													(4.70)		(4.91)	
Literacy F	.032	.004	-.564	-.164	-.015	-.002	-.618	-.180	.006	.000	-.510	-.149	-.740**	-.096	-.566	-.109
	(.07)		(1.45)		(.03)		(1.65)		(.02)		(1.32)		(2.52)		(1.39)	
Land	.010*	.323	-.122	-.656	.009*	.299	-.119	-.640	.011*	.003	-.127	-.686	.009**	.262	.039	.126
	(2.83)		(1.21)		(2.22)		(1.20)		(2.20)		(1.27)		(3.22)		(.36)	
Land^2	-.000*	-.298	.003	.718	-.000*	-.279	.003	.700	-.000*	-.258	.003	.733	-.000*	-.185	.000	.013
	(2.28)		(1.37)		(2.15)		(1.34)		(2.10)		(1.39)		(2.38)		(.04)	
Income per capita	-.003	-.138	.003	.359	-.003	-.137	.003	.361	-.005	-.239	.004	.421	---	---	---	---
	(.86)		(1.04)		(.86)		(1.06)		(.90)		(1.24)					
Income^2	.000	.066	-.000	-.403	-.000	.055	-.000	-.404	.000	.060	-.009	-.467	---	---	---	---
	(.42)		(1.21)		(.35)		(1.22)		(.55)		(1.42)					
FPKAP	-.211	-.092	-.097	-.075	---	---	---	---	-.410**	-.177	-.187	-.144	.052	.021	-.139	-.072
	(1.26)		(.55)						(2.57)		(1.18)		(.50)		(.75)	
M/F Communication	-.540**	-.195	-.216	-.149	-.626**	-.226	-.259	-.179	---	---	---	---	-.209	.070	.057	.026
	(2.80)		(1.19)		(3.46)		(1.59)						(1.70)		(.30)	
Child utility	.495**	.335	.083	.104	.482**	.326	.078	.097	.512**	.346	.081	.101	.180**	.119	.318**	.264
	(5.08)		(.90)		(4.96)		(.86)		(5.16)		(.88)		(2.94)		(3.28)	
Child mortality	-.003	-.020	.022*	.257	-.005	-.028	.023*	.265	.002	.001	.022*	.254	.026*	.155	.007	.056
	(.28)		(2.35)		(.40)		(2.46)		(.19)		(2.31)		(3.74)		(.76)	
Village d1	-1.191	-.166	---	---	-1.13	-.158	---	---	-1.15	-.159	---	---	-.351	-.046	---	---
	(1.87)				(1.78)				(1.61)				(.93)			
Village d2	-.995	-.143	---	---	-1.06	-.153	---	---	-1.24*	-.179	---	---	-.069	-.009	---	---
	(1.56)				(1.66)				(1.96)				(.19)			
Village d3	-1.959**	-.251	---	---	-1.99**	-.255	---	---	-1.70**	-.215	---	---	-.262	-.031	---	---
	(2.86)				(2.90)				(2.54)				(.63)			
Village d4	---	---	-.362	-.106	---	---	-.418	-.122	---	---	-.348	-.102	---	---	-.587	-.119
			(.96)				(1.16)				(.93)				(1.48)	
Constant	5.80**	---	4.01**	---	5.41**	---	3.80**	---	4.78**	---	3.41**	---	-7.50**	---	-6.47**	---
	(5.33)		(3.58)		(5.18)		(3.62)		(4.56)		(3.39)		(5.66)		(3.55)	
R^2	.24		.27		.24		.27		.20		.26		.58		.61	
R^2 adj.	.19		.16		.19		.17		.17		.16		.57		.56	
SSE	1697		173		1711		174		1756		176		1433		229	
D.F.	200		73		201		74		201		74		295		80	

(1) Desired parity equation is estimated in three ways:
- #1 - equation includes both family planning KAP and male/female communication variables.
- #2 - equation excludes family planning KAP variable.
- #3 - equation excludes male/female communication variable.

Table 5 (cont'd.)

Dependent Variable = Perceived Child Utility

Variables	Desired Parity				Children Ever-Born			
	Rural N=214		S-U N=85		Rural N=308		S-U N=91	
		Beta		Beta		Beta		Beta
Age female	.042**	.186	.062**	.406	.043**	.187	.065**	.410
	(2.67)		(3.40)		(3.26)		(3.74)	
Literacy female	-.027	-.005	-1.17*	-.274	-.056	-.011	-1.40**	-.325
	(.08)		(2.46)		(.20)		(3.16)	
Per capita income	-.002**	-.171	-.002	-.185	-.003**	-.151	-.002	-.172
	(2.46)		(1.66)		(2.65)		(1.66)	
Land-owned	-.003	-.132	-.041	-.179	-.003*	-.129	-.045	-.187
	(1.92)		(1.45)		(2.29)		(1.65)	
F labor force participation	.000	.107	-.000	.081	.000	.053	-.000	-.086
	(1.48)		(.75)		(.88)		(.86)	
Family type	-.306	-.076	-.335	.107	-.405	-.097	-.401	-.122
	(1.12)		(.99)		(1.77)		(1.23)	
Village D1	.213	.044			.666	.133		
	(.47)				(1.91)			
Village D2	1.29**	.276			1.36**	.289		
	(2.93)				(4.08)			
Village D3	.801	.151			.978*	.173		
	(1.60)				(2.41)			
Village D4			.063	.015			.029	.007
			(.14)				(.07)	
Constant	1.92**	---	2.65**	---	2.01**	---	2.84**	---
	(2.35)		(2.98)		(3.28)		(3.32)	
R^2	.13		.23		.12		.28	
R^2 adj.	.09		.16		.09		.22	
SSE	896		283		1316		294	
D.F.	204		77		298		93	

Table continued on next page

TABLE 5 (cont'd.)

DEPENDENT VARIABLE: MALE/FEMALE COMMUNICATION

	Desired Parity				Children Ever-Born			
	Rural		S-U		Rural		S-U	
Age female	-.006	-.052	-.044**	-.515	-.004	-.032	-.045**	-.501
	(.74)		(4.23)		(.55)		(4.33)	
Literacy female	.104	.039	.105	.045	.055	.021	.272	.113
	(.57)		(.40)		(.37)		(1.06)	
Per capita income	.001*	.157	-.004	-.592	.002*	.139	-.004*	-.638
	(2.21)		(1.80)		(2.38)		(2.04)	
Income^2			.000	.566			.000*	.605
			(1.76)				(1.98)	
Land owned	-.001	-.066	.016	.128	-.001	-.057	.019	.139
	(.94)		(1.03)		(.98)		(1.17)	
F labor force participation	.000	.056	.000	.196	.000	.074	.001	.192
	(.90)		(1.84)		(1.19)		(1.87)	
Family type	.169	.079	.176	.101	.153	.072	.235	.128
	(1.14)		(.95)		(1.27)		(1.25)	
Village D1	-.118	-.045	---	---	-.140	-.055	---	---
	(.48)				(.78)			
Village D2	.197	.079	---	---	.190	.079	---	---
	(.81)				(1.08)			
Village D3	-.406	-.144	---	---	-.503*	-.175	---	---
	(1.49)				(2.37)			
Village D4	---	---	.191	.082	---	---	.234	.096
			(.79)				(.96)	
Village D5	---	---	---	---	---	---	---	---
Constant	1.81**	---	4.12**	---	1.72**	---	3.97**	---
	(4.06)		(6.99)		(5.39)		(6.65)	
R^2	.08		.26		.08		.26	
R^2 adj.	.04		.18		.05		.18	
SSE	268		84		360		96	
D.F.	204		76		298		82	

Table continued on next page

Table 5 (cont'd.)

DEPENDENT VARIABLE: FAMILY PLANNING KAP[1]

	Desired Parity				Children Ever-Born			
	Rural		S-U		Rural		S-U	
Age female	-.040** (4.16)	-.277	-.042** (3.95)	-.413	-.032** (4.04)	-.229	-.045** (4.31)	-.453
Literacy female	.287 (1.38)	.089	.385 (1.40)	.146	.233 (1.36)	.074	.538* (2.02)	.199
Per capita income	-.001* (2.07)	.139	-.000 (.49)	-.050	.001** (2.58)	.146	-.000 (.67)	-.066
Land owned	.001 (.65)	.043	.004 (.27)	.032	.001 (1.45)	.083	.001 (.59)	.063
Female labor force participation	.000 (1.17)	.082	-.000 (.90)	-.090	.000 (.68)	.040	-.000 (.93)	-.090
Access	.369 (1.39)	.123			(a)			
Village D1	-.409 (1.49)	-.131			-.156 (.73)	-.050		
Village D2	a				.503** (2.45)	.172		
Village D3	-.122 (.39)	-.036			.157 (.63)	.045		
Village D4			.712** (2.78)	.272			.737** (2.89)	.269
Village D5								
Constant	2.73** (6.40)	---	3.98** (9.36)	---	2.06** (6.45)	---	3.96** (9.29)	---
R^2	.17		.31		.13		.33	
R^2 adj.	.14		.25		.10		.28	
SSE	348		97		498		108	
D.F.	205		78		299		84	

a = not forced into equation due to insufficient tolerance.

(1) KAP = knowledge, attitudes, and practice.

TABLE 6

DIRECT AND INDIRECT EFFECTS OF FEMALE LITERACY AND LAND-OWNERSHIP ON FERTILITY:

	Rural		Semi-Urban	
	Desired Parity	Actual Parity	Desired Parity	Actual Parity
Female Literacy				
Direct effect	.004	-.096*	-.164	-.109
Indirect 1[1]	-.008	.001	.011	.014
Indirect 2	-.008	.004	-.007	.003
Indirect 3	-.002	-.001	-.028	-.086*
Indirect total	-.018	.004	-.035	-.069
Total effect	-.014	-.092	-.199	-.178
Land Ownership				
Direct effect	.323*	.262*	-.656	.126
Indirect 1	-.004	.002	-.002	.004
Indirect 2	.013	-.004	-.019	.004
Indirect 3	-.044	-.015*	-.019	-.049
Indirect total	-.035	-.017	-.040	-.041
Total effect	.288	.245	-.696	.085
Land^2				
Direct effect	-.298*	-.185*	.718	.013

[1] Indirect 1 = through family planning KAP
Indirect 2 = through male/female communication
Indirect 3 = through perceived child utility

in the semi-urban area. In the case of desired parity, grades of schooling completed by female showed a statistically significant direct effect in the semi-urban area only.

Comparing the estimated OLS coefficients of grades of schooling compelted by females vs. males on children ever-born, it was found that, although both are non-significant, the magnitude of the coefficient and t-statistics are stronger on male schooling in the rural sample and on female schooling in the semi-urban sample (see Table 7).

In the structural equation model estimation results reported here, community schooling and family planning access/availability (of services) are measured in terms of community dummy variables. The use of community dummy variables allows control for community effects but avoids multicollinearity between individual schooling and community schooling variable, for example. Per capita income is excluded from the fertility equations (but not from the intermediate variable equations) in order to avoid a possible reverse causation bias from fertility (i.e., family size) to income. Also, a quadratic term of land-ownership is included in the fertility equations to test for a curvilinear relationship between land-ownership and fertility. Note also that, in the desired parity model, desired parity equation is estimated separately with male/female communication and family planning KAP in the equation. The reason for this is that in this model only, including both of these intermediate variables in the parity equation in the same time decreases the magnitude of KAP coefficient significantly. In the children ever-born model, coefficients on family planning KAP and male/female communication do not change significantly from when they are included both in the parity equation to when they are included in the equation separately, one at a time.

Table 7

COMPARING THE EFFECTS OF FEMALE AND MALE SCHOOLING[1] AND LITERACY ON DESIRED AND ACTUAL PARITY:

		RURAL				SEMI-URBAN			
		Female		Male		Female		Male	
		b	Beta	b	Beta	b	Beta	b	Beta
Children Ever-Born	Schooling	.084	.036	-.240	-.122	-.293	-.298	-.261	-.304
		(.36)		(.73)		(1.62)		(1.56)	
	$Schooling^2$	-.032	-.062	.041	.100	.025	.215	.024	.301
		(.66)		(.95)		(1.23)		(1.52)	
Desired Parity	Schooling	-.331	-.151	.434	.231	-.380*	-.597	-.124	-.225
		(.97)		(1.28)		(2.18)		(.74)	
	$Schooling^2$	.076	.160	-.051	-.130	.036	.486	.006	.123
		(1.03)		(.74)		(1.86)		(.40)	
Children Ever-Born	Literacy	-.740**	-.096	-.307	-.045	-.566	-.109	-.812	-.140
		(2.52)		(1.11)		(1.39)		(1.84)	
Desired Parity	Literacy	.032	.004	.216	.033	-.564	-.164	-.492	-.129
		(.07)		(.50)		(1.45)		(1.11)	

[1] Grades of schooling completed.

The significant results in Tables 4 through 8 can be briefly summarized as follows:

Children Ever-Born Model

- Female literacy has a direct negative effect on fertility in the rural sample and an indirect (through perceived child utility) negative effect in the semi-urban sample.
- The direct effect of land on fertility was found to be significant in the rural but not in the semi-urban sample; in the rural sample, it is first positive, and beyond a threshold level[1] becomes negative.
- The indirect effect of land was also found to be significant in the rural sample and negative. Thus, it is interesting to note that, while these results show that land-ownership's direct effect on fertility is initially positive, its indirect effect is negative from the start on.
- Child mortality (mortality of own children) was found to have statistically significant positive effect in the rural but not in the semi-urban sample.
- Of the intervening variables, perceived child utility has a significant positive effect on children ever-born in both samples. Note that the effects of family planning KAP and male/female communication on children ever-born are non-significant.
- The significant explainers of child utility in the rural sample are: land-ownership (-) as mentioned earlier, per capita income (-), and residence in community #6 and in #8 (+). That is, these variables affect fertility indirectly, through perceived child utility, in the rural area. It is interesting that residence in rural community #6 and #8 has a positive effect on perceived child utility since these two communities are characterized with labor-intensive economic production activity requiring relatively little physical strength (in community #6 it is

[1] The threshold level equals 556 cuerdas in the rural sample.

growing tomatoes and chili, and in community #8 it is cottage-industries, see Table 3). The negative effects of land-ownership and income on perceived child utility suggest that even in the context of the rural communities, an increase in wealth and income affect child utility so as to increase perceived costs (direct and opportunity costs) and decrease perceived benefits of children. In the semi-urban sample, female age (+) and literacy (-) have a significant effect on perceived child utility. Although the effects of land and income are non-significant, they have the right sign (-) and are approaching significance. Perhaps the reason for the negative effect of literacy on perceived child utility in the semi-urban sample is that, in the urban context, literate females identify more costs associated with child bearing and rearing than illiterate women.

Desired Parity

- The effects -- direct and indirect -- of female literacy were found to be non-significant in both samples. This should be contrasted with the results regarding the effects of grades of schooling completed by female on desired parity which showed a significant effect in the semi-urban sample. Also, it is somewhat surprising that female literacy was found to have a significant effect on actual parity but not on desired parity in the rural sample.
- The significant effects on desired parity in the rural sample are those of income (indirect negative effect through male/female communication and family planning KAP), land-ownership (curvi-linear effect, first positive and, after a threshold level, negative), age of female (positive indirect effect through perceived child utility), residence in community #6 (indirect

positive effect through perceived child utility and residence in community #8 (direct negative effect).

It is surprising that while residence in community #8 has a positive effect on perceived child utility (significant in the children ever-born model), it also has a negative effect on desired parity. The direct effect might reflect things such as educational aspirations of parents for their children, and, in this respect, it should be pointed out that community #8 has the highest community level schooling of the rural communities.

- In the semi-urban sample, the only significant explainer of desired parity is child mortality. Of the other variables, female literacy (-) is the closest to approaching significance.
- Note that both the effect of male/female communication and that of family planning KAP on desired parity are significant in the rural sample.

Comparing the effect of female and male literacy on fertility, Table 7 shows that the significant difference is with respect to female literacy having a statistically significant effect on children ever-born in the rural sample, as compared with the non-significant coefficient on male literacy.

Table 8 shows the percentage change in parity with respect to an additional year of schooling (elasticities of fertility with respect to schooling): we see that, in the rural area, actual parity decreases by 4.5 percent as a result of an additional year of female schooling and by 1.8 percent as a result of male schooling. The respective figures are 2.2 percent (female schooling) and 2.8 percent (male schooling) in the semi-urban area.

Table 8

PERCENTAGE CHANGE IN PARITY WITH RESPECT TO AN ADDITIONAL YEAR OF SCHOOLING[1]:

	RURAL		SEMI-URBAN	
	Female	Male	Female	Male
Children Ever-Born	-4.5%	-1.8%	-2.2%	-2.8%
Desired Parity	0.2	1.4	-3.0	-2.5

[1] Calculated from the literacy variable:

Formula used is $(d/\bar{y})/D$
where d = estimated coefficient on schooling variable
$\bar{y}$ = mean parity when literacy = 1
D = mean grades of schooling completed when literacy = 1

The Effects of Type of Economic Production

Additional analyses of the fertility model revealed important effects of economic production type of family on fertility. In summary, it was found that (see tables 9, 10, 11):

- wage laborers in both areas have lower mean parity and perceived utility than farmers, and type of economic production of family has a significant effect (through child utility) on fertility in the rural sample (i.e., semi-commercial farmers perceive significantly more utility than other categories of economic production).
- analysis of variance and multiple classification analysis (MCA) showed that female schooling is a significant explainer of children-ever-born among wage-laborers and among commercial farmers in rural and among wage-laborers in semi-urban sample; moreover, female schooling is a significant explainer of perceived child utility among wage-laborers in the semi-urban area but not among farmers in the semi-urban area.
- regression results show that, in the semi-urban sample, the indirect effect of literacy on fertility (negative effect through perceived child utility) is upheld among wage-laborers but not among farmers. This is an important finding, in that it suggests that it is not the urban vs. rural environment, per se, that 'explains' the negative indirect effect of literacy on fertility but, more so, the interaction of being wage-labor in the urban context.

Qualitative Choice Model of Fertility Desires

And, last, table 12 shows results of logit analysis (non-linear maximum likelihood estimation methodology for dichotomous dependent variables)[1] of fertility desires. This analysis of fertility desires

---

[1] See Nerlove, M. and Press, S.J., Univariate and Multivariate Log-Linear and Logistic Models (1973) Rand Corporation, Santa Monica.

Table 9

MEAN AND STANDARD DEVIATION FOR VARIABLES IN OLS REGRESSION BY TYPE OF ECONOMIC PRODUCTION. CHILDREN EVER-BORN MODEL.

VARIABLES		RURAL				SEMI-URBAN	
		Wage Labor	Subsis-tence Farmer	Semi-Commercial Farmer	Commercial Farmer	Wage-Labor	Farmers
		N=51	N=77	N=76	N=109	N=70	N=21
Age female	$\bar{x}$	30	35	33	34	33	43
	SD	9	9	11	9	12	15
Parity	$\bar{x}$	4.78	6.44	5.71	6.16	4.38	6.71
	SD	2.85	3.19	3.32	3.54	2.19	2.93
Child mortality rate	$\bar{x}$	19.58	19.23	16.31	17.15	14.53	22.12
	SD	24	17	20	19	18	21
Perceived child utility	$\bar{x}$	3.25	3.17	3.83	3.25	2.78	4.33
	SD	2.25	2.29	2.07	2.19	2.15	1.55
Male/female communication	$\bar{x}$	1.82	1.57	1.92	2.07	2.56	2.28
	SD	1.09	1.00	1.10	1.21	1.12	1.42
FPKAP	$\bar{x}$	1.27	1.09	1.37	1.75	2.98	2.19
	SD	1.33	1.23	1.37	1.40	1.18	1.63
Female literacy	$\bar{x}$	.41	.15	.21	.27	.61	.48
	SD	.50	.36	.41	.45	.49	.51
Land ownership	$\bar{x}$	17	37	40	79	1.24	6.03
	SD	31	75	54	143	6	14
Income/cap.	$\bar{x}$	118	54	69	185	244	213
	SD	114	62	46	190	188	169
Female labor hours	$\bar{x}$	716	557	171	662	536	861
	SD	1250	1099	430	1086	1074	1526
Number of families in house	$\bar{x}$	1.51	1.26	1.30	1.23	1.30	1.43
	SD	.67	.47	.57	.51	.57	.87
Access to FP	$\bar{x}$	.20	.26	.32	.39	1.00	1.00
	SD	.40	.44	.48	.49	0	0
Village dummy 1	$\bar{x}$	.29	.27	.30	.21	----	----
	SD	.46	.44	.46	.41		
Village dummy 2	$\bar{x}$	.20	.26	.34	.39	----	----
	SD	.40	.44	.48	.49		
Village dummy 3	$\bar{x}$	.31	.14	.06	.25	----	----
	SD	.47	.35	.25	.43		
Village dummy 4	$\bar{x}$	---	---	---	---	.48	.05
	SD					.50	.22

TABLE 10

OLS COEFFICIENTS. CHILDREN EVER-BORN MODEL WITH DUMMY VARIABLES TO CONTROL FOR EFFECT OF TYPE OF ECONOMIC PRODUCTION OF FAMILY

	Parity Equation		Perceived Child Utility Equation		M/F Communication Equation		Family Planning KAP Equation	
	Rural	S-U	Rural	S-U	Rural	S-U	Rural	S-U
Age female	.557** (6.47)	.514** (5.47)	.048** (3.52)	.061** (3.18)	-.001 (.19)	-.048** (4.05)	-.029** (3.41)	-.041** (3.37)
Age female^2	-.005** (4.32)	-.005** (4.64)						
Literacy female	-.748* (2.24)	-.399 (.92)	-.016 (.05)	-1.54** (3.38)	.021 (.13)	.297 (1.07)	.285 (1.52)	.481 (1.68)
Land	.009** (2.95)	-.013 (.12)	-.003* (2.34)	-.029 (.81)	-.001 (1.44)	.022 (1.01)	.001 (.99)	.003 (.15)
Land^2	-.000* (2.22)	.001 (.31)						
Income/Cap.			-.003** (2.50)	-.002 (1.36)	.001 (1.05)	-.004 (1.90)	.001* (1.95)	-.001 (.77)
Income^2						.000* (1.97)		
Female labor			.000 (1.78)	-.001 (1.45)	.001 (1.39)	.000 (1.77)	.000 (1.13)	-.000 (1.02)
Family type			-.504* (2.09)	-.396 (1.24)	.141 (1.07)	.244 (1.25)		
Perceived child utility	.189** (2.74)	.311** (2.99)						
M/F Communication	-.188 (1.41)	.014 (.07)						
FPKAP	.054 (.47)	-.136 (.71)						
Child mortality	.028** (3.67)	.007 (.70)						
Econ. type 1	-.409 (.68)	-.529 (.80)	.647 (1.22)	-.851 (1.21)	.429 (1.47)	-.253 (.59)	.242 (.73)	.307 (.69)
Econ. type 2	-.301 (.77)	.839 (.84)	.755* (2.25)	.449 (.43)	.360* (1.95)	.159 (.25)	.259 (1.21)	.589 (.89)
Econ. type 3	-.293 (.79)	-.414 (.33)	.172 (.49)	.652 (.47)	.489** (2.57)	-.762 (.90)	.332 (1.51)	.992 (1.13)
Village dummy 1	-.358 (.86)		.725* (2.03)		-.216 (1.10)		-.138 (.61)	
Village dummy 2	-.017 (.04)		1.45** (4.26)		.092 (.49)		.498* (2.29)	
Village dummy 3	-.084 (.17)		1.25* (2.82)		-.594** (2.44)		.227 (.81)	
Village dummy 4		-.394 (.89)		.369 (.80)		.319 (1.13)		.692* (2.38)
Access								
Constant	7.38** (4.84)	5.89** (2.86)	1.51* (2.39)	3.53** (3.17)	1.51** (4.22)	4.18** (5.38)	1.79** (4.88)	3.59** (5.67)
R^2	.57	.62	.17	.35	.09	.26	.15	.31
R^2 adj.	.55	.55	.13	.26	.05	.15	.11	.23
SSE	1372	212	1096	256	332	93	445	103
D.F.	263	73	266	76	266	75	267	77
N rural	279							
N semi-urban	87							

Table 11

COMPARING PARITY EQUATIONS BY TYPE OF ECONOMIC PRODUCTION. CHILDREN EVER-BORN MODEL.
Dependent Variable = Children Ever-Born

	RURAL				SEMI-URBAN	
	Wage Labor	Subsistence Farmer	Semi-Commercial Farmer	Commercial Farmer	Wage Labor	All Farmers
Age female	.310** (1.85)	.464** (2.08)	.616** (4.35)	.339 (1.62)	.599** (6.77)	.211 (.53)
Age female^2	-.001 (1.21)	-.004 (1.17)	-.006** (3.56)	-.002 (.73)	-.007** (6.29)	-.001 (.23)
Literacy F	-.384 (.75)	-.364 (.57)	-1.26 (1.89)	-.735 (1.26)	-.731 (1.76)	.014 (.01)
Land	.019 (.55)	.021* (2.53)	.008 (.61)	.012** (2.59)	.176 (1.30)	.058 (.24)
Land^2	.000 (.25)	-.000* (2.38)	-.000 (.12)	-.000* (2.07)	-.001 (.28)	-.002 (.37)
Perceived child utility	-.094 (.72)	-.061 (.59)	.231 (1.45)	.366** (2.89)	.243** (2.59)	.811 (1.40)
Child mortality	.018 (1.59)	.057** (3.59)	.025 (1.53)	.018 (1.31)	.006 (.56)	-.023 (.62)
M/F communication	-.189 (.72)	.022 (.09)	-.095 (.35)	-.268 (1.15)	-.060 (.31)	.615 (1.10)
FPKAP	.316 (1.45)	-.054 (.26)	.047 (.20)	-.005 (.02)	-.133 (.66)	-.281 (.60)
Village dummy 1	-.561 (.73)	-.467 (.75)	.130 (.18)	.469 (.46)	---	---
Village dummy 2	-1.29 (1.61)	.261 (.38)	-.298 (.42)	.989 (1.14)	---	---
Village dummy 3	-.612 (.85)	.266 (.35)	-.747 (.65)	.704 (.71)	---	---
Village dummy 4	---	---	---	---	-.278 (.74)	.768 (.23)
Constant	3.17 (1.21)	6.14 (1.66)	-8.68** (3.57)	-4.96 (1.44)	-6.98** (3.90)	-3.89 (.47)
R^2	.74	.71	.63	.52	.64	.65
R^2 adj.	.65	.65	.55	.46	.58	.30
SSE	107	227	309	647	120	60
D.F.	38	64	63	96	59	10
N	51	77	76	109	70	21

Table continued on next page

Table 11

COMPARING PERCEIVED CHILD UTILITY EQUATIONS BY TYPE OF ECONOMIC PRODUCTION
(Children Ever-Born Model)

	RURAL				SEMI-URBAN	
	Wage Labor	Subsistence Farmer	Semi-Commercial Farmer	Commercial Farmer	Wage Labor	All Farmers
Age female	.039	.044	.063**	.025	.058*	.070**
	(1.02)	(1.37)	(2.74)	(1.07)	(2.44)	(3.24)
Literacy F	-.148	-.627	.914	-.243	-1.83**	-.416
	(.22)	(.84)	(1.63)	(.53)	(3.44)	(.77)
Land	.025**	.002	-.008	-.[illegible]2	-.114**	-.023
	(2.14)	(.52)	(1.85)	(1.69)	(2.34)	(1.15)
Income/Cap.	.002	.000	-.010*	-.002*	-.001	.001
	(.618)	(.03)	(1.98)	(2.33)	(.74)	(.26)
Female labor	-.000	-.000	.002	.000	-.000	-.000
	(.27)	(.72)	(1.49)	(1.46)	(.63)	(1.94)
Family type	.184	-.786	-.166	-.401	-.369	-.534
	(.35)	(1.40)	(.41)	(.99)	(.87)	(1.66)
Village dummy 1	.279	-.224	.702	.990	---	---
	(.27)	(.32)	(1.16)	(1.47)		
Village dummy 2	-.379	1.76**	.610	1.85**	---	---
	(.35)	(2.60)	(1.04)	(3.08)		
Village dummy 3	.775	1.32	-.606	1.69*	---	---
		(1.42)		(2.45)		
Village Dummy 4	---	---	---	---	.252	-.779
					(.52)	(.33)
Constant	.95	1.39	2.26**	2.08	2.83*	2.70
	(.57)	(1.55)	(2.10)	(1.87)	(2.67)	(1.92)
R^2	.17	.18	.26	.21	.28	.69
R^2 adj.	.00	.07	.16	.14	.20	.52
SSE	221	325	237	407	232	15
D.F.	41	67	66	99	62	13
N	51	77	76	109	70	21

Table 12

CONDITIONAL LOGIT MODEL OF FERTILITY CHOICE.
DEPENDENT VARIABLE EQUALS THE LOG ODD OF CHOICE TO HAVE AT LEAST ANOTHER CHILD. BINARY LOGIT MAXIMUM LIKELIHOOD ESTIMATES, T-STATISTICS IN PARENTHESES. RURAL SAMPLE.

Variables	PARITY		
	3, 4	5, 6	7+
Literacy female	.078	3.12	-.312
	(.13)	(1.42)	(.31)
Perceived child utility	.267*	.765	.741**
	(1.79)	(1.49)	(2.53)
Land-ownership	.002	.036	-.000
	(.45)	(.97)	(.08)
Family planning KAP	.239	1.04	-.877*
	(1.12)	(1.28)	(1.71)
Child mortality	.092	-1.34	.586*
	(.17)	(.81)	(1.81)
Age female	-.031	-.309	-.169
	(.67)	(1.25)	(1.18)
Income per cap.	.001	-.024	.001
	(.61)	(1.17)	(.20)
Village D1	-.201	.743	-2.21
	(.26)	(.27)	(1.32)
Village D2	-.592	.637	-1.91
	(.75)	(.18)	(1.43)
Village D3	-.124	2.42	-3.05
	(.12)	(.65)	(1.51)
Constant	-.017	4.23	3.40
	(.01)	(.54)	(.62)
Percent correctly classified	56	94	81
Likelihood ratio	9	22	46
D.F.	60	21	54
N	71	32	65
$\bar{y}$	.592	.375	.185

differs from the path models above in not only methodology[1] but in that it looks at fertility decisions sequentially at given parities. The purpose of this analysis was to examine the consistency of the path analysis results, especially with respect to the effects of perceived child utility, for it could be argued that perceived child utility is a result, at least partially, of high fertility and not the reverse.

Due to limitations on sample size, the analysis is done on groups of adjacent parities (i.e., group of third and fourth parity women, group of fifth and sixth parity women, and group of 7+ parity women). The results are reported only for the rural sample. The dependent variable is derived from the question whether the respondent wants additional children (in addition to the children you already have, do you want more children?) and coded as zero if no more children were wanted and as one if at least one more child was wanted.

The results indicate that, even when we analyze fertility (desires) sequentially by given parities, perceived child utility is consistently the strongest predictor of desire to have/not have another child. Note also that female literacy is non-significant here as in the OLS results in the case of desired parity. The results indicate, moreover, that at parity 7+, family planning KAP becomes a significant predictor (negative as expected) of desired fertility. This supports the hypothesis that family planning KAP has a significant effect on fertility, but only after the motivation to have no more children already exists.

---

[1] Maximum likelihood estimation vs. ordinary least squares regression.

Conclusion

This short report has presented results from analyses of determinants of fertility using cross-sectional household survey data from four rural and two semi-urban Ladino communities in Guatemala.

The focus has been on the role of female and male heads of household's education on actual and desired fertility. The effect of education on fertility has been examined in the context of a path model of socio-economic and attitudinal variables and fertility.

Table 13 summarizes the effects that were found to be statistically significant[1] by actual vs. desired parity and rural vs. semi-urban residence. It can be seen that while female literacy has a significant effect on actual parity (children ever-born) in both samples, its effects on desired parity are non-significant.

As expected, age of female is the most consistent predictor of actual and desired parity.

The models explain fertility better in the rural than semi-urban area, largely because of the important role played by perceived child utility and variables affecting it in the rural sample.

In the rural areas, land-ownership, income, and community characteristics have more of an effect on fertility than female literacy. In the semi-urban area, female literacy is the only variable with a significant effect on actual parity. It should be reminded that this effect is upheld among wage-laborers but not among farmers when the semi-urban sample is stratified by type of family economic reproduction.

Comparing the signs of estimated coefficients with those hypothesized

---

[1] $p \leq .05$

Table 13

Summary of Statistically Significant[1] Effects on Fertility

	Rural	Semi-Urban
Children-ever-born	age (+), age^2 (-) direct literary female (-) direct land (+), land^2 (-) direct child mortality (+) direct Community #2 (+) indirect through utility Community #3 (+) indirect through utility age female (+) indirect through utility land-owned (-) indirect through utility income (-) indirect through utility	age (+), age^2 (-) direct age (+) indirect through utility literary female (-) indirect through utility
Desired parity	Community #3 (-) direct land (+), land^2 (-) direct income (-) indirect through male/female communication and family planning KAP Community #2 (-) indirect through utility age (+) indirect through utility income (-) indirect through utility	child mortality (+) direct

(1) $p < .05$

in the fertility model in diagram 1, the most important difference between the two is the negative effect of female literacy on perceived child utility and fertility in the rural sample. In our theoretical model, it was hypothesized that, along with other socio-economic variables, literacy would have a positive effect on fertility due to expected prevalence of the 'income effect' over 'price effect.' This is, in fact, what takes place in the case of land-ownership, although we were able to show that it has a threshold level beyond which the initial positive effect becomes negative. In the case of female literacy, however, even the initial effect on fertility is negative and significant in both areas. Since, in the same time, the indirect effect of literacy on fertility is non-significant, the direct effect represents attitudes and values affected by literacy but not well-measured by the three intervening variables in the fertility model such as educational aspirations for children, for example.

The policy implications of these findings are important. The findings indicate that even in rural areas of a low literacy LDC, literate females tend to have fewer children ever-born than illiterate females. Extending literacy among rural populations and especially among rural women is, thus, expected to have an effect in lowering fertility.

Moreover, in rural areas such as the ones of this study, wealth of the household and income per capita show as strong an effect on fertility as literacy. Although wealth initially increases fertility, beyond a threshold level it will decrease fertility. In the long-run it is the improved socio-economic condition of the household which will contribute significantly to a decrease in fertility. These results

indicate the importance of indirect rather than direct population policies.

The policy implications related to the significance of perceived child utility as a determinant of fertility are also important. Whether perceived child utility represents actual child utility accurately or not, it does suggest that, in this sample, respondents are having children largely because they are perceived to be useful now and later in life and not so much for other reasons, such as lack of family planning knowledge, for example. Consequently, population policies should take this variable into account and try to alter it. How to alter perceived child utility is not without controversy. However, our results suggest that land-ownership (and perhaps, more generally, wealth), income, and community characteristics have a significant effect on utility in the rural sample and female literacy among wage-laborers in the semi-urban sample.

Our results indicate that perceived child utility is related to the type of economic production one is engaged in as a family, interacting with the overall context of the community. Laborers perceive less utility than farmers in both areas, and semi-commercial farmers perceive significantly more utility than other categories of farmers in the rural areas. The solution, with respect to population policy is not, however, to advocate changing farmers into wage-laborers. Instead, different forms of farm production should be examined in detail with respect to reliance on family labor for increased production. For example, cooperative forms of farm production may contribute to decreased reliance on family labor and may, therefore, tend to decrease fertility. Moreover, sources of old age security and 'social' security other than one's own children should decrease perceived child utility and, consequently, fertility.

In general, the findings of the present analyses indicate that high fertility in rural and semi-urban areas of LDCs may be better explained as a rational adaptation, given the limited set of alternatives available, than as an irrational and ignorant behavior (Carvajal and Geithman 1976, Folbe 1976, Williams 1976). Thus, it can be altered more effectively by means of changing the socio-economic conditions underlying high fertility motivation coupled with a family planning program than by simply providing knowledge about and access to contraception (Tabarrah 1971, Mamdani 1972, McGreevey and Birdsall 1974, Wolfson 1978).

REFERENCES

Andrews, F. M., et al. Multiple Classification Analysis. Institute for Social Research, University of Michigan, 1973.

Bagozzi, R. P., and Van Loo, M. F. "Toward a General Theory of Fertility: A Causal Modeling Approach." Demography (August 1978b).

Berelson, B. Family Planning and Population Programs. A Review of World Developments. University of Chicago Press, 1966.

Carvajal, M. J., and Geithman, D. T. Family Planning and Family Size Determination. The Evidence from Seven Lation American Cities. University of Florida Press, 1976.

Cassen, R. H. "Population and Development: A Review." World Development 4 (October 11, 1976).

CELADE and CFSC. Fertility and Family Planning in Metropolitan Latin America. Chicago: Community and Fertility Study Center. University of Chicago, 1972.

CICRED. Seminar on Infant Mortality in Relation to the Level of Fertility. Bangkok: 1974.

Cochrane, S. Education and Fertility: What Do We Really Know. Washington, D.C.: World Bank, 1978.

Davis, K. "Institutional Patterns Favoring High Fertility in Underdeveloped Areas." Eugenics Quarterly 2 (1955).

Deere, C. D., and deJanvry, A. "A Conceptual Framework for the Empirical Analysis of Peasants." Giannini Foundation of Agricultural Economics, Division of Agricultural Sciences, University of California, December 1978.

Duncan, O. D. Introduction to Structural Equation Models. New York: Academic Press, 1975.

Folbe, N. "Economics and Population Control." Science for the People 8 (November-December 1976).

Hill, R., et al. "Intra-Family Communication and Fertility in Puerto Rico." Rural Sociology XX (September/December 1955).

Hoffman, L. W., and Hoffman, M. L. "The Value of Children to Parents" in Fawcett, J. T. (ed.) Psychological Perspectives on Population. New York: Basic Books, 1973.

Leibenstein, H. Economic Backwardness and Economic Growth. New York: John Wiley & Sons, 1957.

MacFarlane, A. "Modes of Reproduction." Journal of Development Studies 14, No. 4 (July 1978).

Mamdani, M. The Myth of Population Control: Family, Caste and Class in an Indian Village. New York: Monthly Review Press, 1972.

McGreevey, W. P., and Birdsall, N. The Policy Relevance of Recent Social Research on Fertility. Washington, D.C.: Smithsonian Institution, 1974.

Marshall, J. F., and Polgar, S., eds., Culture, Natality, and Family Planning. Carolina: University of North Carolina Press, 1976.

Mejia Pivaral, V. "Caracteristicas Economicas y Socioculturales de Cuatro Aldeas Ladinas de Guatemala." Guatemala Indigena VII, 3, 1972.

Mueller, E. "Economic Motives for Family Limitation: A Study Conducted in Taiwan." Population Studies 26, 3, November 1972.

Nerlove, M., and Press, S. J. Univariate and Multivariate Log-Linear and Logistic Models. Santa Monica: Rand Corporation, 1973.

Preston, S. The Effects of Infant and Child Mortality on Fertility. New York: Academic Press, 1978.

Schultz, T. P. "Interrelationships Between Mortality and Fertility" in Ridker, R. G., op. cit., 1976.

Sear, A. M. "Predictors of Contraceptive Practice for Low Income Women in Cali, Colombia." Journal of Biosocial Science 7 (1975): 171-188.

Simon, J. L. The Effects of Income on Fertility. Carolina Population Center, University of North Carolina at Chapel Hill, 1974.

Simonen, M. S. Education and Fertility in Rural and Semi-Urban Guatemala. A Cross-Sectional Study of Six Ladino Communities. Berkeley Project on Education and Nutrition, University of California, Berkeley, January 1980.

Tabbarah, R. B. "Toward a Theory of Demographic Development." Economic Development and Cultural Change 19(2) (January 1971).

Williams, A. D. "Determinants of Fertility in Developing Countries -- Review and Evaluation of Literature" in Keeley, M. C. (ed.) Population, Public Policy and Economic Development. New York: Praeger Press, 1976.

Wolfson, M. Changing Approaches to Population Problems. Paris: Development Centre, OECD, 1978.

Wyon, J. B., and Gordon, J. E. The Khana Study: Population Problems in the Rural Punjab. Cambridge: Harvard University Press, 1971.

PART III

SYNTHESIS OF FINDINGS AND POLICY IMPLICATIONS
by Judith B. Balderston

Introduction

In this section, we summarize findings of the Berkeley Project on Education and Nutrition, identifying linkages between parts of the study and indicating those results which may be most important for policy purposes. Because of the wide range of data collected by INCAP and Rand on the families and individual children in the four villages, we have been able to go beyond the original questions of nutritional effects on school performance to investigate relationships connecting nutrition, schooling, work, family size, and agricultural production. Analytical results so obtained make possible the integration of findings, providing information necessary in the formulation of consistent nutritional, educational, population, or rural development policies.

In Part II of this report, we presented four related but separate studies which utilized information of the INCAP and Rand data bases collected in the four villages. The most important results of these four studies will be summarized in Section III-A. In themselves, each set of statistically significant findings is, we believe, important and robust, for both research and planning purposes.

In addition to the separate findings of each of the four studies, (the longitudinal models on child nutrition and growth, and the three cross-sectional models on schooling and work, literacy and agricultural efficiency, and literacy and fertility), we present the results of the four studies as an integrated whole. Not only are the four studies related because they are based on the same four villages, but the variables themselves measure conditions that are interrelated in the lives of the village inhabitants.

We must caution that, although we shall integrate the findings of the four studies, we could not combine variables from the four parts of the analysis into one complete and comprehensive model. Limitations imposed by available computer space and time would have precluded computation. Besides, limitations imposed by missing data would have made it impossible to construct one inter-generational model that would include a full set of economic, biological, psychological, and educational variables. We relied instead on carrying out separate studies based on sets of data that were similar in important characteristics. Following these separate analyses, we were then able to relate each set of results to the other sets.

III-A. Summary of Findings

From the series of estimations using longitudinal models carried out by Alan Wilson and presented in Section II-A, we have seen that:

1. Villages differ in farm activity and in nutritional status. Pre-intervention diet is shown to be better in the two Atole villages than in the Fresco villages. Mothers and fathers in the Atole villages have had better dietary intake before intervention, and this is confirmed by their larger head circumferences. Higher infant mortality existed in the Atole villages prior to the project.

2. The total calorie intake in home diets of children, after intervention, in the Fresco villages is consistently larger than in the Atole villages and contains almost the same proportion of protein. Because of supplementation, the aggregate proportion of protein in children's diet is substantially higher in the Atole villages.

3. At birth, the children in the four villages are at the 25th percentile for weight and the 10th percentile for height compared to American children. By 36 months, the village children are considerably below the 5th percentile of American children.

From three to seven years, children's growth follows below the 5th percentile level the slope of American children's growth patterns. The differences in growth between atole and fresco villages is especially marked for girls; girls are probably less well nourished at home than boys, but the sex disparity in home diets is less in the Atole villages.

4. There is a significant effect of supplementation, not home diet, on growth in weight and height. The most significant effect of supplementation was to increase the proportion of diet from protein sources. Protein

supplementation shows a strong positive effect on growth, especially up to 48 months. Diarrhea affects annual increments of growth negatively, as do other forms of illness, especially in the Fresco villages.

5. Combining calories and protein from both home diet and supplementation sources shows that there is a significant effect of protein on growth which appears between ages 12 and 24 months. It subsequently declines and appears again with smaller effects between 48 and 60 months. Here again is a consistent negative effect of diarrhea upon growth.

6. A highly significant positive effect of height on verbal development shows up at all ages. This is the most important result of this section: taller children do better on verbal tests, controlling for wealth (CONSUMP is a factor score of family possessions). CONSUMP itself is a significant predictor. Verbal test scores are negatively related to sex (girls do better than boys), morbidity, and family structure (children from nuclear families do worse than others) and positively to parental literacy and modernity backgrounds, and family size (with a higher proportion of older family members at the time of the subject's birth).

7. School enrollment is positively affected by verbal factor scores, by sex, by height, by the number of younger siblings at the time the subject is 6, by parental literacy and modernity, occupation of mother, but not by father's occupation.

8. Verbal factor scores and father's occupation have a highly significant positive effect on teacher assessment while mother's modernity and vocabulary, and teacher assessment of achievement influence negatively. Concurrent diet has a substantial, though not quite significant, effect upon teacher assessments.

9. Among the factors that appear most to influence growth in height, weight, and verbal development, there is a strong sequential influence of height, weight, and verbal development from one period to the next. These measures show stability from one period to another.

10. Mother's attendance at the supplementation center is a strong determinant of child's intake of supplementation, particularly in the Atole villages.

From cross-sectional analysis carried out by Judith Balderston and presented in Section II-B, it was seen that:

11. Taller, healthier children are more likely to enroll and attend a full year of school than smaller, less healthy children. School enrollment is related to the need for the child's work, but other things equal, the effect of prior nutrition on height and health determine school attendance.

12. Children from more affluent farming families work more (and also go to school more) than children from less affluent ones. This may be the result of the need for children's work because of larger land holdings, but it may also reflect the ability of more vigorous children to work more.

13. Elder siblings are more likely to attend school than younger siblings. Children in larger families are less likely to attend than children from smaller ones.

14. Parental literacy affects school attendance of children and their school achievement. Parental literacy is also related to affluence.

15. The opportunity for children to engage in paid work in commercial cash crop production is an important factor in determining high participation in the labor market and low participation in school. This is especially

evident in village 6 where over a several year period school participation is seen to be relatively low compared to the other villages.

From the analysis carried out by Maria Freire and presented in Section II-C, we saw that:

16. Farmers who are literate are more productive than those who are not. Literate farmers used more chemical inputs and raised relatively more cash crops. Even after controlling for chemical inputs and kinds of crops grown, education appeared to be a significant factor in explaining variations in output.

17. The impact of education on productivity appeared to differ among farming groups. For the middle range of farmers, literacy made the greatest difference. For commercial farmers, productivity did not appear to be affected by the farmer's education.

18. Production of traditional crops appeared to be less affected by literacy of the farmer than was the output of crops where innovation in planting and fertilizing is of importance.

From Mari Simonen's analysis presented in Section II-D, it was seen that:

19. Female literacy was found to have a significant direct negative effect on fertility (children ever-born) while male literacy was found to have no effect.

20. Land ownership (wealth) was found to have a significant direct curvilinear (first positive and after a threshold level, negative) effect on fertility (children ever-born). In addition, both land-ownership and income per capita were found to have a significant indirect (through perceived child utility) negative effect on fertility.

21. Of the three intervening variables considered in the fertility determinants' analyses, perceived child utility (i.e., the perceived benefits of children to parents in farm work, and general and old age security) showed a consistent positive effect on fertility (children ever-born), while family planning knowledge, attitudes, and practice and male/female communication showed no effect.

22. Important differences in the effects of the variables considered on fertility were found by different types of economic production of family. For example, belonging to a semi-commercial familial production unit shows a significant positive effect on fertility (children ever-born) through perceived child utility, while belonging to the other types of production units shows no effect.

III-B. Integration of Results and Policy Recommendations

Results

The strongest and most important conclusions that emerge from our analyses are as follows:

(1) Protein supplementation during early childhood has a positive effect on growth. Height, verbal development, school enrollment, and achievement are all affected positively by protein supplementation. Diarrhea has negative effects on all the same variables.

(2) Children's school enrollment and achievement are also affected by parental affluence and the need for the children's work. Village differences appear to affect patterns of work and school participation differentially. In one village, school participation is consistently low and affluence of family appears to affect school enrollment. In the other villages, enrollment is affected by child size and health. Family occupation, sex, and family size also affect child's activities and school participation.

(3) Schooling of farmers relates to agricultural production. Literate farmers accept innovation more readily and are able to bring higher returns than illiterate farmers in their farming activities.

(4) Schooling of female heads of household affects perceptions of the economic utility of children and the number of children ever-born. It is expected that because of the lower perceived utility of children, women will have fewer children as family economic conditions improve.

From these conclusions, we see from (1) that with improved nutrition and improved health conditions (medical care, potable water, improved

sanitation), children will have greater opportunity to participate and achieve better in school. With improved schooling of males (assuming that adult men continue as farming decision makers), we see from (3) that farmers will achieve higher productivity, increased affluence of the families, and less need for children's work. Improved schooling of females will lead from (4) to lowered perceptions of the need for children's help and then smaller families. Higher production and lower family size leads to higher per capita income and better nutrition for the members of the family. This leads to better school performance resulting from better health, physical and verbal development, and parental literacy.

We see that education is in a pivotal position. Literacy is one of the instruments through which farmers innovate and by which attitudes concerning the economic need for children are affected. But, educational performance is dependent on children's well-being and the family's need for their work. Children with poor health and chronic malnutrition may not realize their full physical and psychological potential. To improve their chances of school success, the early health and nutrition of children should be improved. Figure III-1 presents research results of the Berkeley project for the families in the four Guatemalan villages studied by INCAP and Rand. It portrays the interaction among interventions, conditions, and perceptions from the perspective of the child's development and is intended to show the interrelationship of family economic conditions, parental literacy, family size, health, diet, school, and work. Relationships are centered upon the child's well-being and the outcomes of well-being in terms of school and work. Arrows represent the significant and important relationships that were found in the course of the study in Berkeley.

From this diagram, which combines results from the four parts of

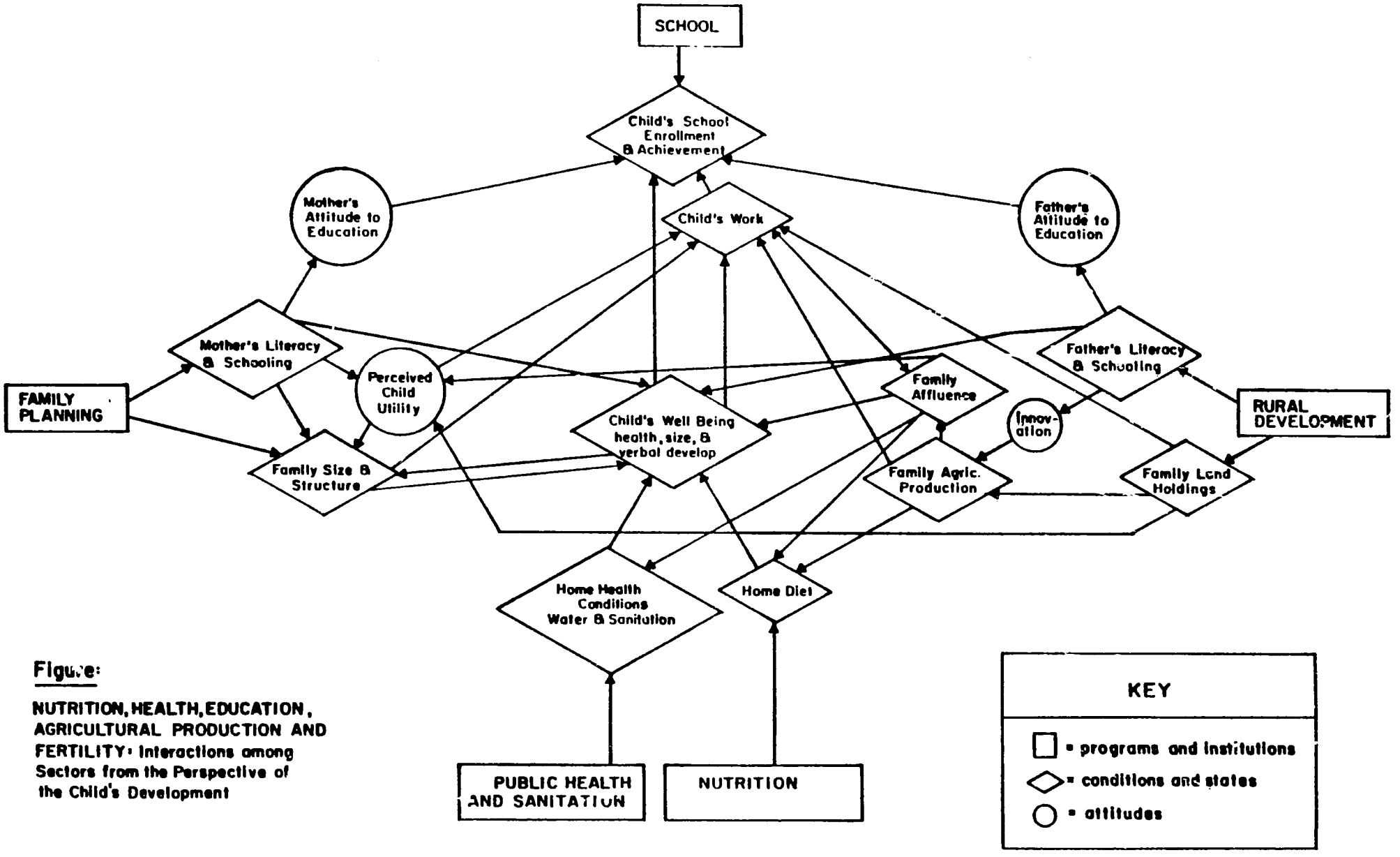


Figure:

NUTRITION, HEALTH, EDUCATION, AGRICULTURAL PRODUCTION AND FERTILITY: Interactions among Sectors from the Perspective of the Child's Development

our study, we note that the child's development is affected by family affluence and work, family size and structure, nutritional intake, quality of family diet and health conditions, and parental literacy. Improving conditions for the child at an early age has inter-generational benefits as the child reaches adulthood. At that point, the man's literacy and schooling appear to influence his productivity as a farmer, and the woman's literacy and schooling appear to influence her attitudes about desirable family size. Education, then, in combination with other development efforts, can be an effective method for increasing per capita income through its effects on increasing productivity and lowering family size.

Policy Implications

Out of these results, we shall now try to develop implications for policy planning that could improve the well-being of the child and the ultimate productivity and welfare of the adult. A great many combinations of program interventions occur (thirty possible combinations of the five interventions can be made, as shown graphically in Figure III-2), but some of these combinations of programs may be less feasible or of lower priority than others. We will suggest combinations of interventions because single interventions and poorly coordinated multiple interventions are inefficient or do not reach the poorest and neediest people. We will not offer cost-benefit estimates, as this is beyond the scope of the Berkeley Project's work. Moreover, costs are specific to countries and regions within countries; they also change over time and depend in part on the skill of policy interveners. Cost estimation would be a logical next step of analysis, along with the measurement of need for the specific context in which policy interventions are planned.

The criteria that we shall use in examining possible program interventions will be as follows: (1) we think it desirable to intervene with efficient programs, i.e., ones that are as economical as possible and that take advantage of complementarities between programs; (2) we think it desirable to intervene with programs that achieve greater equity be being directed particularly to those families in greatest need.

Since school success has been shown to depend partly on adequate prior nutritional intake and health, by intervening within these two areas (nutrition and health), children may have a better chance of achieving literacy. Thus, we recommend that educational interventions be accompanied by nutritional and public health interventions for, in terms of efficiency, the nutritional and health interventions will improve the effectiveness of the educational intervention. And, in terms of equity, the children who will be reached through improved nutrition and health will be just those children who previously would have failed to attend school.

School participation has also been found to depend on family affluence and the need for children's work. Moreover, food intake and health status of children have been found to depend partly on family affluence. Consequently, we recognize the efficiency and equity of interventions to improve incomes and family affluence that will improve children's nutrition, health status, and school success.

In order to improve nutritional intake, we do not perceive that a nutritional intervention would be necessary or advisable for the long-run. A short-term program of nutritional intervention can assist the present generation and will increase their long-run opportunities. In the long-run, an intervention of this kind would not be beneficial because of the increased dependency of families on food aid. Instead, in the long-run, measures to increase agricultural production and incomes are recommended.

Public health interventions need to be provided for the long-run, but such investments are relatively cheap and highly effective. A well, supplying potable water to a community, is not costly. Latrines are also of low cost. In both cases, the benefits of providing these services are seen to have very valuable positive externalities since the utilization of nutrients is substantially improved in the absence of infection.

Thus, it is recommended that public health interventions over the long-run and improvement of nutritional intakes both over the short-run (supplementation, family food allotments, etc.) and long-run (measures to improve family affluence) be instituted immediately. We expect that if these are undertaken, children's school performance will also improve, thereby increasing the efficiency of investments in schools.

Preconditions for the acceptance of innovation in agricultural production and more positive attitudes with respect to family planning are the literacy and grade attainment levels of adults; literate farmers tend to be more accepting of rural innovation,* and literate women tend to be more accepting of family planning information.** To increase the investment in either rural development or family planning without increasing the stock of educated and literate adults would be wasteful in the long-run since acceptance of measures would tend to be limited. Thus, the efficient and equitable approach would be to increase the education of children so that increased benefits of rural development and family planning interventions would occur as they reach adulthood. For the present generation of adults, intervention for rural development and family planning either need to be combined with adult literacy programs or should emphasize outreach to non-literate adults

---

*Rural innovation also depends on the quality and quantity of land available.

**Attitudes favorable to family planning depend also on the economic conditions of the family.

In the long table that follows, we present the list of potential interventions, taken one at a time and in all possible combinations of the five sectors: nutrition, public health, education, rural development, and family planning. For each potential policy intervention, we discuss briefly the expected outcome in terms of feasibility, efficiency, and equity.

By "interventions," we intend the following:

1. Nutrition: Direct interventions such as improvement of intake by supplementation, family food allotments, food stamps, etc., which would be directed to specific target groups or to the whole community. Indirect interventions such as improved access to knowledge and means of production so as to increase agricultural production and family incomes.

2. Public Health: The provision of potable water, sanitation, public health clinics.

3. Education: Assuming that existing public schools continue, this intervention involves improved school facilities, classes, and materials.

4. Rural Development: Agricultural extension services, loans, and/or direct provision of irrigation equipment, fertilizer, and seed.

5. Family Planning: Provision of information and methods (possibly integrated with the health clinic services).

Since a wide range of programs can take place within each type of intervention and cost will vary with the amount of service provided, it is beyond the scope of this report to suggest the level of intervention.

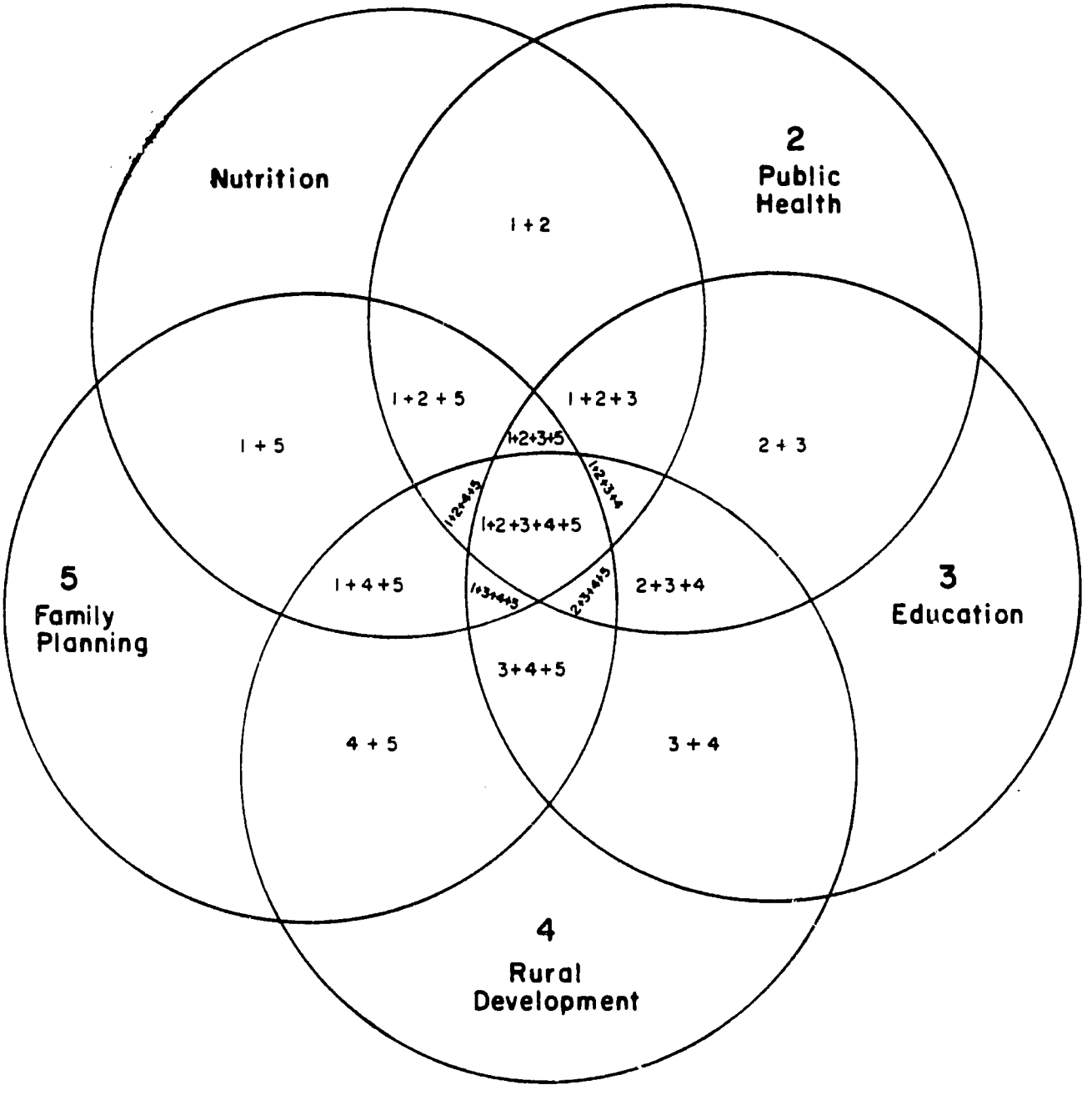
It is our intention in the remaining discussion to indicate the combinations of interventions that are relatively efficient and equitable. In the list that follows, an asterisk (*) will be found beside those interventions that, taking advantage of program complementarities, are able to save resources and meet the needs of the poorest populations.

We recognize, of course, that while our recommendations are drawn from

the results of a particular study and are, therefore, somewhat specific to that study, that in other contexts, it would be possible to conduct short cross-sectional studies that would identify needs, estimate costs, and could yield intersectoral policy results.

Figure III-2

REPRESENTATION OF ALL POSSIBLE INTERVENTIONS COMBINING SECTORS



Decision rules from the research findings:

If 1 then 2 ⟶ Improvements in nutritional intake should be accompanied by improvements in public health.

If 3 then 1 & 2 ⟶ Educational investments should be accompanied by improvements in health and nutrition.

If 4 then 3 ⟶ Rural development is more accepted by literate adults.

If 5 then 3 ⟶ Family planning is more accepted by literate adults.

Intervention	Discussion
1 Nutrition	A nutritional intervention alone without improving public health conditions would benefit those children who are already relatively free of diarrheal disease. Our results indicate that children's growth is not affected by nutritional intake alone but depends also upon the ability of the body to utilize nutrients. This intervention would be inefficient when undertaken alone.
2 Public Health	When potable water and sanitation are provided, they reduce the incidence of diarrheal disease. This intervention would reach the poorest children and would, therefore, appear to be equitable. It would be more efficient if combined with a nutritional intervention over the short-run. Over the long-run, decreased infant mortality, leading to larger families, might have adverse effects on per capita income, unless attempts were made to increase income and decrease family size.
3 Education	Investments in education alone in the absence of improved health and nutritional intake will tend to assist the children already able to attend and profit from schooling. This is inequitable since it benefits children from relatively more affluent families. It is inefficient because children with poor health and nutrition will tend to perform poorly in school. We therefore recommend that there be investments in nutrition and health before increasing investments in education.
4 Rural Development	Programs directed at improving agricultural practice appear to be most beneficial to literate farmers, who tend to accept innovation, and those farmers with access to adequate land. To reach other farmers, increased education through adult literacy programs, plus access to better land, would increase the effectiveness of the rural development programs. In the absence of such additional interventions, rural development alone would tend to be inequitable, because it would not reach the poorest people.

Intervention	Discussion
5 Family Planning	Family planning interventions appear to influence most the attitudes of women who are literate and for whom children are perceived as having lower economic utility. To increase the acceptability of family planning methods, the need for children as a source of labor and security to parents will have to be removed. This intervention alone tends to be relatively inefficient.
1+2 Nutrition and Public Health	The combination of public health interventions with nutritional improvement would improve the utilization of food intake through the lowered incidence of diarrhea and would improve the health and physical growth of children. As a result, children's potential school performance would also benefit. It would not be desirable to continue nutritional supplementation in the long-run, however, because of its high cost and increased dependency on outside aid. We recommend that, in the short-run, nutritional intervention be given to small children in order to improve their chances of success in school and, therefore, greater productivity as adults.
1+3 Nutrition and Education	Nutritional intervention with increased investments in education are wasteful since the public health intervention, necessary to improve utilization of food, is missing. Increasing expenditures for education would tend to benefit most those already able to attend school. This combination of interventions is therefore inefficient and inequitable.
1+4 Nutrition and Rural Development	Again, improving nutritional intake without improved health is wasteful. Rural intervention alone will tend to benefit those already receptive to innovation, most likely, literate farmers. Besides, since nutritional assistance tends to increase the dependency on food relief while rural development activities are intended to increase independence, these activities appear to be incompatible.
1+5 Nutrition and Family Planning	As before, nutritional improvement without a public health intervention is inefficient. Family planning interventions are inefficient alone also because of the low level of acceptability of the information unless accompanied by change in economic conditions and female literacy. This combination is not recommended because of its inefficiency.

Intervention	Discussion
2+3 Public Health and Education	Improvements in health conditions by the provision of sanitation and potable water would improve the health and physical development of children, making it possible for them to utilize more efficiently their nutritional intake from family diet. Improved growth and health will assist children to attend and perform in school while increased investment in school may increase school participation. Therefore, this is more efficient and of lower cost than combining nutrition and education. It is inequitable, however, in that children with poorest home diets will not receive the benefits of either public health or improved quality of schooling.
2+4 Public Health and Rural Development	Rural development plus the delivery of public health services would improve the health and vigor of adults and children while assisting farmers to adopt more efficient methods. Under 4, rural development alone, it was seen that literate farmers would benefit more from information provided than would others. To improve the effectiveness of this approach, rural deveopment would be especially directed to illiterate farmers or combined with adult literacy programs.
2+5 Public Health and Family Planning	Public health and family planning can be combined in a synergistic delivery system with mutually enhancing benefits. With public health improvements, it is expected that more infants will survive, and family planning interventions are directed to discouraging the higher family size. However, it was seen that family planning information does not tend to influence the attitudes of illiterate women and those who perceive children to be useful in work and in old age security. It is therefore advisable to combine family planning with increased education for females and enhancement of family economic production over the long run.
3+4 Education and Rural Development	Increased investments in education with rural development intervention will tend to improve the productivity of families with literate members and will increase the level of performance of children already able to attend school. It is an inequitable intervention, however, because the poorest families are not reached.

Intervention	Discussion
1+2+5 Nutrition with Public Health and Family Planning	Better nourished children with better health will do better in school. More children who survive will lead to bigger families. Increased family planning interventions may help to encourage limiting births. Eventually, increased literacy and schooling will lead to more productive adults who may desire fewer children.
2+3+4 Public Health with Education and Rural Development	The combination of public health, education, and rural development interventions will tend to decrease morbidity; healthier children will be more likely to attend school and perform well. Rural development will enhance agricultural production. For farmers who are literate and innovative, this will be especially beneficial. Eventually, through increased production, diet might improve at home. Because of the absence of family planning programs, family size may increase with improved health conditions, and, thus, per capita food consumption might not improve. Lack of a nutritional intervention detracts froms its impact on the malnourished poor.
2+3+5 Public Health with Education and Family Planning	Public health with improved education and family planning will lead to higher school attendance. Family planning and public health were seen to be complementary programs. Investments in education may lead to increased school participation with beneficial effects on economic productivity. This appears to be a slow but efficient combination of programs, but inequitable because of the absence of a nutritional intervention, which would reach the malnourished poor.
3+4+5 Education with Rural Development and Family Planning	Education with rural development and family planning will benefit most those children already able to attend school because of reasonably good health and robustness and because their families are able to spare their work. Literate adults will tend to benefit from the rural development and family planning interventions. This appears to be an inefficient and inequitable combination of programs, as it tends to benefit those who already have been able to take advantage of schooling.
2+4+5 Public Health with Rural Development and Family Planning	Public health with rural development and family planning would tend to improve chances of infant survival while attempting to limit family size. However, the latter program will most benefit literate mothers with less need for the children's

Intervention	Discussion
1+2+4+5 Nutrition with Public Health, Rural Development, and Family Planning	This set of interventions includes all sectors except additional investments in education. Nutritional improvement plus health would increase the participation and achievement of children who would probably otherwise fail to attend or to perform well. The interventions of agricultural assistance and family planning would be beneficial in the long-run especially if geared to illiterate adults in the short-run. By having nutritional improvement and health, it is likely that increased spending on schools might not be necessary, that benefits from these programs would have positive externalities for school achievement. This is a recommended set of interventions that appears efficient and equitable. Nutritional interventions would not be necessary for the for the long-run, as eventually increased productivity would assist in improving family diet.
1+3+4+5 Nutrition with Education, Rural Development, and Family Planning	This includes all sectors except public health. As a result, the nutritional intervention would be wasted since those with diarrheal infection would not be able to utilize additional nutrients effectively. The additional investment in education would inequitably tend to reach children already able to attend school. The rural development and family planning information would tend to benefit farmers already innovating, as well as literate mothers. We recommend that this set of interventions not occur, as it would be inefficient and wasteful without the introduction of potable water and sanitation.
2+3+4+5 Public Health, Education, Rural Development, and Family Planning	Since this set of interventions does not include nutrition, it is lacking in immediate assistance to the children of the poorest and most malnourished families. These children will be only partially assisted by the decreased morbidity for public health measures. The combination of education, rural development, and family planning would appear to be effective in improving economic productivity and decreasing family size in the long-run. The absence of nutritional assistance for the poorest families does detract from its advantages.
1+2+3+4+5 Nutrition, Public Health, Education, Rural Development, and Family Planning	This is the most complete, most ambitious, and therefore most expensive set of interventions. The nutritional intervention could be used only in the short-run to benefit those who need it most. In combination with improved health and decreased morbidity and infant mortality, children

Intervention	Discussion
	work. Rural development programs will assist in increasing economic productivity but will benefit most the literate farmers. This appears to be a feasible combination of programs but one that inequitably tends to reward those families with education. In the long-run, as children's school performance improves with better health, higher levels of educational attainment will enable them to participate more fully in innovation. In the long-run, this appears to be efficient, but it is slow, and effects could be felt sooner with the short-term introduction of improved nutrition for the most malnourished children.
1+2+3+4 Nutrition with Public Health, Education, and Rural Development	Increased nutritional intake with public health, education, and rural development will, in the short-run, assist children to be better nourished, healthier, bigger, and better able to perform in school. In the long-run, farmers who are able to acquire literacy will achieve higher productivity. There is the danger with this set of interventions that family size will grow larger because of decreased morbidity and that, although income and food may increase, per capita income and intake may not. It is therefore recommended that a family planning intervention be introduced to curtail such increases, with a special attempt to reach illiterate mothers. The nutritional intervention would only be needed for a short period, as eventually rural development programs would help achieve higher productivity and, therefore, improve home diets. We believe that this is a feasible combination of programs and is relatively efficient and equitable.
1+2+3+5 Nutrition with Public Health, Education, and Family Planning	This combination of improved nutrition, health, education, and family planning would be efficient and equitable in the long-run, as it would tend to assist the most needy children to achieve better health, growth, and school performance while helping to limit the growth in family size due to decreased infant mortality. We recommend this as a combination of interventions where nutritional assistance would be provided in the short-run and where family planning would be combined with special attempts to reach uneducated women.

Intervention	Discussion
3+5 Education and Family Planning	Education and family planning would be mutually beneficial eventually. Education of females would result in more receptivity to family planning information. However, this is inefficient as there is a delay between generations. It is also inequitable becuase the education does not tend to reach children whose health and nutritional status prevent them from attending school.
4+5 Rural Development and Family Planning	Rural development with family planning would result in some efficiency gains in agriculture and, with some decline in family size due to the acceptability of family planning information, might improve per capita income. Interventions would reach already literate adults and would, therefore, be inequitable unless focused on illiterate adults or combined with adult literacy program. If the rural development innovations would decrease reliance on family labor and increase old age security of families, then it would result in families using family planning more efficiently.
1+2+3 Nutrition with Public Health and Education	Better nourished, healthier children would benefit from better schools and eventually would lead to more productive adults. This is a recommended set of interventions since it provides assistance to children whose health and nutrition would limit their growth and development and likelihood of participating in school. In the longer run, family size would increase with improvements in diet and health. Family planning interventions would be recommended. Nutritional interventions would, in the long-run, create dependence on food assistance. This is a desirable set of interventions if the additional nutrition is offered as a short-term intervention until its long-term benefits can be realized from the higher productivity of educated adults.
1+2+4 Nutrition with Public Health and Rural Development	Increased nutritional intake for the short-run with enhanced rural assistance in the long-run will help food intake. Increased public health measures will help in the utilization of food. Higher infant survival will probably occur from introduction of public health measures which may prevent per capita income from increasing. On the other hand, increased efficiency of farmers may lead to the increased opportunity of children to attend school instead of working. In the long-run, children who are better educated will become more efficient farmers and women with lower expectations of desired family size. This appears to be an equitable and efficient combination of interventions.

Intervention	Discussion
	would function better in school, and in the long-run, educated adults would tend to be more productive and more accepting of family planning information. In the short-run, due to increased child survival, it would be important to focus family planning programs on those mothers not yet reached because of lack of education. We recommend this program if resources are available.

APPENDIX I

GLOSSARY OF VARIABLES
encountered in the text

ACHCAL3	Estimated total home diet in kcal units age 42 months
ACHHDC3	Home diet calories age 42 months
ACHHDC5	Home diet calories age 66 months
ACHHDP3	Home diet protein age 42 months
ACHHDP5	Home diet protein age 66 months
ACTGY	Gross income from non-agriculture activities (excluding wages)
ATTAGE	Child's age in months at the beginning of year assessed
BCHCAL2	Estimated total home diet in kcal units age 30 months
BCHHDC2	Home diet calories age 30 months
BCHHDC4	Home diet calories age 54 months
BCHHDP2	Home diet protein age 30 months
BCHHDP4	Home diet protein age 54 months
BYR	Birth year
CHDIAR20	Child's incidence of diarrhea age 24-36 months
CHDIAR21	Child's incidence of diarrhea age 36-48 months
CHDIAR14	Child's incidence of diarrhea age 48-60 months
CHDIAR15	Child's incidence of diarrhea age 60-72 months
CHHLTH6	Child's health (proportion of days healthy) age 72-84 months
CHSP18	Average daily total supplementation in kcal age 30 months
CHSP19	Average daily total supplementation in kcal age 42 months
CHSP13	Average daily total supplementation in kcal age 54 months
CHSP14	Average daily total supplementation in kcal age 66 months
CHSPC18	Supplementation calories age 30 months
CHSPC19	Supplementation calories age 42 months
CHSPC13	Supplementation calories age 54 months
CHSPC14	Supplementation calories age 66 months
CHSPP18	Supplementation protein age 30 months
CHSPP19	Supplementation protein age 42 months
CHSPP13	Supplementation protein age 54 months
CHSPP14	Supplementation protein age 66 months
COHORT	Age in months on 31 March
COM1	= 1 if subsistence farm, (sales=0)
COM2	= 1 if semi-subsistence farm (cash crops = 0)
COM3	= 1 if commercial farm (cash crops > 0)
Community 1 (village d1)	Dummy variable for village, takes the value of 1 if village=3
Community 2 (village d2)	Dummy variable for village, takes the value of 1 if village=6
Community 3 (village d3)	Dummy variable for village, takes the value of 1 if village=8
Community 4 (village d4)	Dummy variable for village, takes the value of 1 if village=31 (semi-urban #1)
CONSUMP	House possessions and salary
D6	Dummy variable (Atole - large)
D8	Fresco - small
D14	Atole - small
DIARRHEA	
DIED3	Number elder siblings deceased at age 36 months
DIED6	Number elder siblings deceased at age 72 months

DSTRUCT3	Family structure: nuclear vs. all other at age 36 months
DSTRUCT6	Family structure: nuclear vs. all other at age 72 months
EDASP	Mother's educational aspiration for child in number of grades
ENROLL	School enrollment; (0, 1)
ESC22	How many years enrolled in school
FAMILY TYPE	Number of nuclear families living in and sharing the same house
FAMSZ	Family size
FMSZ6	Number household members at age 72 months
FPKAP	Family planning knowledge, attitudes, and practice scale
F LABOR	Female labor force participation (number of hours for pay 75-76)
FOR13	Mother's expectation of child's grade attainment
HSSIZ	Number of people in household
HSTYPE	Type of dwelling (1, 2, 3, 4)
HT12	Height in centimeters age 36 months
HT14	Height in cenitmeters age 48 months
HT15	Height in centimeters age 60 months
LAND	Land ownership of nuclear and extended family (measured in cuerdas, 1 cuerda = .044 hectares)
MAATT3	Mother's attendance at the supplementation center at age 24-36 months
MAATT4	Mother's attendance at the supplementation center at age 34-48 months
MAATT5	Mother's attendance at the supplementation center at age 48-60 months
M/F COMMUNICATION	Male/female communication scale
MDVOC	Mother modernity and vocabulary
MORBOTH2	Child's incidence of ailments other than diarrhea age 24-36 months
MORBOTH3	Child's incidence of ailments other than diarrhea age 36-48 months
MORBOTH4	Child's incidence of ailments other than diarrhea age 48-60 months
NAMING1	Naming psychometric test age 36 months
NAMING2	Naming psychometric test age 48 months
NAMING3	Naming psychometric test age 60 months
NSLDCAL	Calorie value (in kcal) of nonsold produce (5 major crops)
NSLDPRO	Protein value (in kcal) of nonsold produce (5 major crops)
OLDSIB	Number of older siblings of same sex as subject
ORDR6	Proportion of older siblings at age 72 months
PAHT	Father's height in centimeters
PARSTIM3	
PASTAT	Father's occupation, ownership of land, animals
PERUTIL	Perceived child utility by female head of household (summative scale of items measuring perceived utility of child(ren) as old age security and as having economic value in general)
PLB1	Proportion of family labor in total agricultural labor
PLB2	PLB1 standardized by value of land planted

RCHHOC11	Estimated total home diet (kcal) age 54 months
RECOG1	Recognition psychometric test age 36 months
RECOG2	Recognition psychometric test age 48 months
RECOG3	Recognition psychometric test age 60 months
RECOG4	Recognition psychometric test age 72 months
RDWRTMA	Mother's literacy (0 = illiterate, 1 = reads a little, 2 = reads well)
RDWRTPA	Father's literacy (0 = illiterate, 1 = reads a little, 2 = reads well)
RWMA	Mother's literacy (0, illiterate or reads a little; 1 = reads well)
RWPA	Father's literacy (0, illiterate or reads a little; 1 = reads well)
S21	Average math scores
S22	Average language scores
SALALL	Total wage income accruing to household
SCHOOL	School enrollment (0, 1)
SCHOOL AND WORK	School and work together (0, 1)
SEX	Male=1, female=0
SEX1	Male=1, female=2
2	Male=1, female=2
3	Male=1, female=2
4	Male=1, female=2
SEXSUBJ	Male=1, female=2
STRUC3	Family structure: nuclear vs. all others
TREAT	Atole=1, Fresco=0
V2	Village 6 = 1
V3	Village 8 = 1
V4	Village 14 = 1
VGRAS	Gross value total agriculture production
VILSUP3	TREAT + CHSP19
WORK	Work only = 1
WT12	Weight in kilograms age 36 months
WT14	Weight in kilograms age 48 months
WT15	Weight in kilograms age 60 months
Y62 to Y70	Dummy variable for birth year
YEAR	Birth Year
YONG3	Number of younger siblings at age 36 months
YONG6	Number of younger siblings at age 72 months
ZMAOCC	Mother's occupation factor scores
ZMARD	Mother's literacy and schooling
ZMARD	Father's literacy and schooling

CODES

Village Codes

Original Code	Code Used by A. Wilson	Code Used Elsewhere in Report		
3	FR03 (Fresco - big)	1	Rural	V1
6	AT06 (Atole - big)	2		V2
8	FRD8 (Fresco - small)	3		V3
14	AT14 (Atole - small)	4		V4
31 semi-	---	1	Rural	
32 urban	---	2		

Sex Codes

Original Code	Code Used by A. Wilson	Code Used Elsewhere in Report
1 = male	1	1
2 = female	0	2