AGENCY FOR INTERNATIONAL DEVELOPMENT
WASHINGTON, D. C. 20523

BIBLIOGRAPHIC INPUT SHEET

FOR AID USE ONLY
Batch 62

1. SUBJECT CLASSIFICATION

A. PRIMARY: Food production and nutrition — AS00-0000-G510

B. SECONDARY: Human nutrition--Argentina

2. TITLE AND SUBTITLE
Determinants of infant and child mortality: an econometric analysis of survey data for San Juan,Argentina

3. AUTHOR(S)
Russell,L.B.; Burke,C.S.

4. DOCUMENT DATE	5. NUMBER OF PAGES	6. ARC NUMBER
1975	187p.	ARC

7. REFERENCE ORGANIZATION NAME AND ADDRESS
NPA

8. SUPPLEMENTARY NOTES *(Sponsoring Organization, Publishers, Availability)*

9. ABSTRACT

10. CONTROL NUMBER
PN-AAE-021

11. PRICE OF DOCUMENT

12. DESCRIPTORS
Argentina
Children
Infants
Mortality
San Juan,Argentina?
Statistical analysis
Weight

13. PROJECT NUMBER

14. CONTRACT NUMBER
AID/CM/otr-C-73-212 GTS

15. TYPE OF DOCUMENT

AID 590-1 (4-74)

DETERMINANTS OF INFANT AND CHILD MORTALITY: AN ECONOMETRIC ANALYSIS OF SURVEY DATA FOR SAN JUAN, ARGENTINA

NATIONAL PLANNING ASSOCIATION

Louise B. Russell
Carol S. Burke

February, 1975

Prepared for the Agency for International Development under Contract No. AID/CM/otr-C-73-212, Work Order No.1

ACKNOWLEDGEMENTS

This project has been jointly sponsored by the Agency for International Development and the Pan American Health Organization. A.I.D. provided our financial support during the design stages and during the study itself. PAHO made the survey data available and supplied programming assistance and computer time. To facilitate this cooperative arrangement we were appointed unpaid consultants to PAHO.

We are grateful to the many people at A.I.D. and PAHO who cooperated in making this study possible. Dr. Robert J. Muscat, Chief Economist of A.I.D., and Dr. Charles L.Williams, Jr., worked out the cooperative arrangement. Dr. Ruth Rice Puffer, Principal Investigator of the Inter-American Investigation of Mortality in Childhood, and Miss Mary H. Burke, Chief of PAHO's Health Statistics Department, shared with us their firsthand knowledge of the survey. Dr. Carlos Serrano of PAHO provided expert advice in the development of the nutritional classification system.

We would especially like to thank Ms. Ruthann M. Piepenburg, a programmer-statistician with PAHO, who skillfully solved our software problems and persuaded the computer to produce the computations on which this report is based; and Dr. Barbara Herz, project monitor with A.I.D., for her help and encouragement during all phases of this project.

LBR
CSB

TABLE OF CONTENTS

Page

SUMMARY ..... i.

I. Introduction 1

II. The Determinants of Birthweight 4

Introduction 4
The Regressions: Specification and General Discussion 6
Sex of Child 20
Age of Mother 20
Reproductive History of the Mother (Number of Pregnancies and Outcome of Previous Pregnancy) 25
Complications of Pregnancy (Length of Pregnancy and Multiple Births) 28
Spacing 30
Medical Care Before and During Birth 31
Maternal Nutrition 33
Appendix Tables 36

III. Application of the Nutritional Classification System for Infants Under One Year 41

Introduction 41
Nutritional Classification System 44
Application of the System 45
Correlations Between Nutrition Measures 54
Determinants of Nutritional Status 58

IV. Determinants of Neonatal Mortality 61

Introduction 61
Birthweight 67
Length of Pregnancy 73
Sex of Child 74
Medical Care During Birth (Birthplace and Who Attended Birth) 75
Parity and Spacing 76
Environmental Factors 77
Appendix Tables 79

Page

V. Determinants of Mortality Among Infants 28 Days to 5 Months 84

Introduction 84
Birthweight, Length of Pregnancy and Sex of Child 88
Medical Care 95
Nutrition 97
Environmental and Social Conditions 99
Multicollinearity, A Hypothesis, and Directions for Further Research 103
Appendix Tables 107

VI. Determinants of Mortality in Children Six Months or Older ... 111

Introduction 111
The Regressions: Infants 6 to 11 Months 115
The Regressions: Children 1 through 4 Years 119
Appendix Tables 121

VII. Analysis of a Major Cause of Death: Diarrheal Disease 124

Introduction 124
The Regression Results 128
Appendix Tables 136

VIII. Conclusions and Some Implications for Policy 140

Introduction 140
The Conclusions of the Study 141
Recommendations for Further Research 149

Appendices

A. Selection of the Observations 153
B. Definition of Variables Used in the Regressions 155
C. Weighting the Sample Observations 166
D. Bibliography 173

SUMMARY

CHAPTER I. Introduction

This report describes an econometric analysis of household survey data on infant and child mortality. The survey was taken in San Juan Province, Argentina, one of 13 Latin American project sites included in the Inter-American Investigation of Mortality in Childhood, conducted by the Pan American Health Organization. The purposes of this analysis have been two-fold:

- To identify the determinants of infant and child mortality and obtain measures of their net effects, in order to provide the groundwork for an examination of alternative policies to combat child mortality.

- To test the ability of the survey to support this type of analysis and, by extension, to determine the desirability of carrying out similar analyses for other project sites.

Chapter II of the report presents the results of an analysis of the determinants of birthweight, a major factor in the survival of very young infants. Chapter III discusses the application of a system for summarizing the diet information requested for each child in the survey. Chapters IV through VII contain analyses of infant and child mortality at different ages. Neonatal mortality is examined in Chapter IV, mortality between 28 days and 5 months in Chapter V, from 6 months to five years in Chapter VI, and deaths between 28 days and 5 months from diarrheal disease in Chapter VII. Chapter VIII discusses the policy implications of the results and makes recommendations about the use of data from other survey sites.

CHAPTER II. The Determinants of Birthweight

Other studies have shown, and our study confirms (see especially Chapter IV), the importance of birthweight to the survival of very young infants. In general, the larger the infant, the better its chances of survival. Because of this importance birthweight was singled out for analysis in its own right. The regressions were based on the 2,084 children under two years of age at interview or death for whom birthweight information was available.

Of course, not all the factors that determine birthweight are amenable to the influence of programs aimed at reducing infant mortality. For example, both the sex of the infant and whether the birth is single or multiple affect birthweight. Our analysis agrees with studies done by others in showing that female babies are lighter at birth than males -- by about 195 grams (7 ounces), while multiple births average 700 to 800 grams (1 pound, 9 ounces to 1 pound, 12 ounces) less than single births.

The length of pregnancy is very important to the infant's weight. Pregnancies of six months or less produce babies about 2000 grams (4 pounds, 7 ounces) lighter than full-term babies. The differentials for seven and eight month babies are not consistent, probably reflecting inaccuracies in these data. But these infants are significantly lighter than full-term infants -- with the average differential around 400-450 grams (14-16 ounces).

The equations show that very young mothers have lower weight infants than older mothers. The difference is quite large for mothers under 17; a significant, but smaller, difference holds for mothers between 17 and 20. This result suggests one reason for the higher mortality rates of infants born to young mothers. The analysis did not reveal any consistent tendency for mothers over 35 to have lower weight babies.

In line with previous studies, this study found that first babies are lighter than subsequent babies; no differences were found among higher parity children. The spacing of pregnancies has some effect on birthweight. Babies born within 12 months of a preceding birth were lighter than more widely spaced infants by 140-150 grams (5 ounces). The optimal interval between births appears to be 19 to 24 months, as these infants were heavier than those that were either more closely or more widely spaced.

The relationship between birthweight and the outcome of the previous pregnancy was explored, but the small number of infants with other than living previous birth mean that the findings are highly tentative. The results suggest that, when the previous pregnancy ended in fetal death, the current infant was lower weight than would otherwise be the case.

The analysis provides several interesting results about medical care before and during delivery. First, mothers who received prenatal care had significantly heavier babies than mothers who did not. The infants of mothers with 3 to 5 visits for prenatal care averaged 135 grams (5 ounces) heavier, while those whose mothers had six or more visits averaged 200 grams (7 ounces) heavier. Second, two measures of medical care during birth showed no association with birthweight. Neither the place of birth (home or hospital), nor the attendant at birth (physician, midwife, or other) appears to make any difference to the infant's weight at delivery.

Since the survey did not collect information about the mother's nutrition, we were unable to examine this factor directly. We tried to get around the statistical problems that arise from the omission of an important variable by testing a number of socio-economic factors as possible proxies for maternal nutrition. The results were reasonable and suggested that the proxies might be serving their intended purpose. The validity of the results for other factors is strengthened by the fact that they were unchanged by the addition of these proxies.

CHAPTER III. Application of the Nutritional Classification System for Infants Under One Year

The PAHO survey contains a wealth of information about each child's diet, specifically the types of foods eaten and the number of servings of each per week. Since nutrition is believed to be an important determinant of an infant's or child's survival, it was clearly desirable to make the best possible use of these data in the analysis. In the design stage of this study we developed a classification system to summarize each child's diet in terms of its adequacy or inadequacy in five major nutrients -- protein, calories, iron, Vitamin C, and other vitamins and minerals. Chapter III describes the results of applying this system to the diet information for infants less than a year old.

The classification system could not be applied to 31 percent of the 1,946 infants because of missing information. Most of these unclassifiable infants were less than 28 days old and for the majority, the reasons for the lack of information were clear: the infant either died within 48 hours after birth, was in the hospital continuously from birth to death, or both.

The chapter then discusses the characteristics of the system as applied to infants 28 days and older. In general, the classification showed fewer and smaller differences between living and dead infants in the age range 28 days to 5 months than in the 6 to 11 month group. Overall, about 60 percent of the live infants and 50 percent of the dead infants had diets that were probably adequate in protein; within the 6-11 month group, the percentages were 65 and 34, respectively. Similarly, about 65 percent of the live infants and 50 percent of the dead infants had diets adequate in calories -- while 31 percent of the live infants and 35 percent of the deaths had diets that were grossly inadequate in calories.

We considered the possibility that we could summarize the nutritional information still further by letting the measures for one or two nutrients represent the diet as a whole. Examination of the correlations between adequacy measures for different nutrients indicated that this was possible, in some cases even necessary. Specifically, the classification in terms of protein and/or calories could generally summarize the infant's diet adequately.

We finish the chapter by discussing the factors that other studies have shown are likely to affect nutritional status. Among these are family income, family size, ignorance and taboo -- all of which are associated with primary malnutrition, or inadequate food intake.

CHAPTER IV. Determinants of Neonatal Mortality

A very high proportion of all infant deaths occur during [illegible] first four weeks of life (the neonatal period) -- and especially during [illegible] first week. Of the 1900 infant deaths in San Juan between August 1968 and [illegible]ly 1970, 577 occurred during the first week after birth and another 329 during the next three weeks. The sheer numeric importance of these deaths indicated that they

should be studied separately. The results support the belief that environmental factors do not play a significant direct role until later: the infant's condition at birth is the most important predictor of its survival to the postneonatal period. The analysis in this chapter asks what determines an infant's ability to survive the neonatal period, and is based on a total of 1,946 infants -- 806 neonatal deaths, 936 infants who died after the neonatal period, and 204 infants who were living at the time of the survey.

The equations provide additional evidence of the importance of birthweight and, by extension, of the importance of understanding the underlying determinants of birthweight. The probability of death for babies 1400 grams (3 pounds) or less at birth is .71 -- .69 higher than the average probability for infants in the range 3500-3900 grams.[1/] The probability is less but still very high, .42, for infants between 1500 and 1900 grams. It declines still further for infants between 2000 and 2400 grams, to about .09. After 3000 grams (6.5 pounds) additional weight makes little difference and the probabilities of death at any of these weights are very low.

In addition to the effects of birthweight, the gestation period, or length of pregnancy, has a very important effect on an infant's chances of survival. The data are insufficient to provide an accurate picture of the relationship in detail, but they show clearly that infants born during the seventh month, or earlier, have much higher probabilities of death than full-term infants -- between .24 and .36 higher.

Our results agree with those of other investigators in showing that female infants have a lower probability of death than males during the neonatal period, even after factors like birthweight and length of pregnancy have been taken into account. But the differential, .03, is considerably smaller than those estimated for birthweight and length of pregnancy. Parity and spacing also show some direct effects on neonatal mortality. Infants of the seventh or higher parities have a somewhat higher probability of death than lower parities (.05), while children born more than four years after a previous birth have a somewhat lower probability (also .05). The analysis found no evidence of differences in the mortality experience of infants born at home rather than in the hospital, or of infants delivered by midwives rather than physicians.

Finally, the analysis provided further support for the notion that social and environmental factors have little direct bearing on mortality during the first weeks of life. Children born into rural, or very poor urban and suburban, neighborhoods had a somewhat poorer chance of survival, as did those born to single mothers. But a variety of other environmental factors showed no association with mortality.

CHAPTER V. Determinants of Mortality Among Infants 28 Days to 5 Months

Infant deaths drop off sharply after the neonatal period. In San Juan 906 of the 1900 survey deaths occurred during the first four weeks, in contrast to an average of 147 deaths during each of the next five months -- a high number

---

1/ The reader should recall that probabilities cannot be greater than one (the event is certain to happen) or less than zero (the event is certain not to happen).

when compared with the experience of developed countries, but much below the toll of the first month. The analysis of causes of mortality during this period was based on all infants between 28 days and 5 months at death or interview for whom sufficient information was available, a total of 689 deaths and 99 living infants.

The regressions show that both low birthweight and prematurity are associated with higher mortality between 28 days and 5 months, although their effects are much less important than during the neonatal period. Infants who weigh 2400 grams or less at birth, but who nevertheless manage to survive the first month, have a probability of death higher by .12-.17 than do heavier infants. The probability of death for infants who are less than full term at birth averages about .12 higher than the probability for full term infants. Female infants may continue to have an advantage in survival during this period -- but the estimated coefficient is unstable and not always statistically significant.

Infants who have been vaccinated have significantly better chances of survival than those who have not been. The estimated differential in the probability of death is about .07-.09. Although types of vaccination were requested on the survey, the information was seldom provided and it was impossible to test whether any particular type was primarily responsible for the result.

Breastfed babies are significantly more likely to survive beyond the fifth month than those who are not breastfed. The probability of death is estimated to be about .09 lower for breastfed infants. Variables representing the nutritional adequacy of the infant's diet did not produce useful results for this age group.

While the source of the family's water supply is not associated with mortality, there is evidence that the type of toilet facilities may be. In equations which included only those factors likely to be proximate causes of death, infants whose families had no toilet facilities had a probability of death higher by about .11 than those whose families had flush toilets, and the differential was statistically significant. But this result was quite sensitive to the addition of general social variables such as type of neighborhood and father's occupation. Infants living in one room dwellings also had significantly higher chances of death.

A number of social variables, which generally serve as proxies for the conditions that accompany poverty, showed associations with mortality -- among them, type of neighborhood, family size, father's occupation, and marital status of the mother. But these proxies introduced statistical difficulties -- they were correlated with some of the factors which seemed more likely to act as proximate causes of death and caused instability in the results for these other factors. This problem is especially likely to arise in relatively small observation sets, and both this and the other analyses in the report indicate that sample size is the source of the trouble here. Combining the data from two or more project sites to produce a larger set of observations would be a worthwhile extension of this research.

CHAPTER VI. Determinants of Mortality in Children Six Months or Older

In the San Juan project, most of the deaths of children under one year -- fully 86 percent -- occurred before the infant reached six months of age. Relatively few occurred during the next six months or in the following four years. This chapter analyzes these older children in two separate groups -- infants 6 to 11 months and children 1 through 4 years. The analysis of 6 to 11-month-olds was based on 247 deaths and 85 living infants in this age group; the observation set of 1 to 4-year-olds included 247 deaths and 674 living children.

The fewness of the deaths for each group was the primary reason for the undistinguished results. The regressions for 6 to 11-month-olds showed that generalizations were possible for this group, but that the observation set was too small to produce clear results significant at the conventional levels. Several of the results hinted that the determinants of death might be similar to those for infants 28 days to 5 months -- for example, the advantage shown for infants who were vaccinated was significant at the 80 percent level.

The results for children 1 to 4 years gave no evidence that it is possible to generalize about the causes of death for this group.

CHAPTER VII. Analysis of a Major Cause of Death: Diarrheal Disease

The analysis in preceding chapters is based on the assumption that factors like low birthweight and poor living conditions predispose a child to illness and death, and that these conditions are truly the underlying causes of death. This chapter considers infants for whom diarrheal disease was either an underlying or associated cause of death, in an attempt to determine whether certain conditions are more closely related to particular illnesses. The 363 deaths from diarrheal disease and 99 living children in the age group 28 days to 5 months made up the observation set.

The notable thing about the results is that they are simply a replay of those for all infants 28 days to 5 months -- with the important difference that the levels of significance and R^2's are lower. They give no evidence that the factors involved in diarrheal disease differ from those for general mortality in this age group. They do, however, give a very clear example of the effects of sample size since the equations are based on the same number of living children as were the regressions for all infants 28 days to 5 months, but only about half as many deaths.

CHAPTER VIII. Conclusions and Some Implications for Policy

To help them decide among possible programs aimed at reducing infant mortality, program planners need to know how much of their limited resources will be absorbed by each program, and its likely benefits in terms of lives saved.

One of the purposes of this study has been to provide some of the needed information. Chapter VIII reviews the conclusions of the study in light of the types of programs that planners might consider.

The study provides precise information about only one part of the two-part link between maternal nutrition and infant mortality. Chapter IV (see especially Chart IV.2 and accompanying discussion) describes the relationship between birthweight and mortality in detail. But the lack of any information about maternal nutrition from the survey made it impossible to measure the effect of maternal nutrition on birthweight. This additional information is needed before planners can weigh the probable benefits of programs to supplement the nutrition of pregnant women against the costs.

The study provides strong evidence that prenatal care is effective in reducing infant mortality. The results showing the effect of prenatal care on birthweight, in conjunction with those showing the effect of birthweight on mortality, allow us to estimate what would have happened if care had been provided to the nearly 5000 women in San Juan who had no prenatal visits. Of 163 infants who would have died without care, 41 could have been saved if such a program had been introduced.

The results related to care during delivery are of a negative sort -- delivery by midwives rather than physicians had no effect on mortality, either directly or through birthweight; similarly there was no difference in the mortality of infants born at home rather than in the hospital. Programs to promote one kind of care over the other are thus unlikely to reduce infant mortality.

The lower mortality of breastfed infants suggests that programs to encourage breastfeeding may be worthwhile. The estimated advantage of breastfed infants implies that, for every 11 women who could be persuaded to breastfeed their infants, one infant would survive who would not otherwise have survived.

Programs to vaccinate infants are also likely to be effective -- the estimates indicate that one infant could be saved for every 11 to 14 who were newly vaccinated.

A number of the conclusions suggest that family planning programs could, by focusing on high risk groups, help reduce mortality. Very young women, those who have just given birth, and single women should be encouraged to postpone pregnancy. If pregnant, they should be considered to be at greater than average risk.

The chapter closes by discussing several areas for further research based on the survey data collected for other sites. We believe that further econometric work based on these surveys could most profitably begin by using them to enlarge the set of observations for infants between 28 days and 5 months. The usefulness of replicating the results for birthweight and neonatal mortality and of extending the analysis of 6 to 11-month-olds is also considered.

CHAPTER I. Introduction

This report describes an econometric analysis of household survey data on infant and child mortality. The survey was taken in San Juan Province, Argentina, one of 13 Latin American project sites included in the Inter-American Investigation of Mortality in Childhood, conducted by the Pan American Health Organization. The purposes of this analysis have been two-fold:

- To identify the determinants of infant and child mortality and obtain measures of their net effects, in order to provide the groundwork for an examination of alternative policies to combat child mortality.
- To test the ability of the survey to support this type of analysis and, by extension, to determine the desirability of carrying out similar analyses for other project sites.

The San Juan survey was selected for two reasons. First, the sample was large enough to support a fairly detailed analysis, and one that analyzes sub-groups of children separately. Second, the PAHO staff recommended San Juan as one of the projects for which the data were most complete and of highest quality. These characteristics seemed desirable for both goals of the study. A large sample and good data offer the best chance for identifying causes of mortality with confidence. By the same token, if a good survey does not produce useful results, it is safe to conclude that it would not be worthwhile to extend the analysis to sites with poorer data.

San Juan Province is located in central-western Argentina, close to the border between Argentina and Chile, and was one of the few projects to include the surrounding rural area, in addition to the city of San Juan and its suburbs. The total population of the province was estimated at 384,020 as of 1 August 1969, of which 52,230 were estimated to be children under five years of age. Child death rates were among the highest found in the investigation. Deaths of children under one year of age, per 1000 live births, averaged 81.3 -- 50.7 in the city itself, 87.9 in the suburban departments, and 94.5 in the rural areas. Deaths of children one to four years of age averaged 3.0 per 1000 population.[1/]

All deaths of children under five that occurred in San Juan between August 1968 and July 1970 -- a total of 2,156 -- were included in the survey. In addition, one out of every 48.3348 households was selected for interview and these produced a sample of 891 living children under the age of five. A lengthy questionnaire was completed for each child. For the deaths, information was sought from various medical records as well as through an interview of the infant's family. The questionnaire for the living children, which differed in a few respects from the one for the infant deaths, was completed using only information obtained from the home interview.

Not all of the children originally included in the survey were retained for this analysis. For a number of the deaths, it had proven impossible to obtain a home interview and thus the child's record provided almost none of the information considered important for the analysis. An additional few of the deaths and of the living children[2/] had records so incomplete that they could not contribute to the analysis, and these were deleted. In all, 167 deaths and

---

1/ Data from Puffer and Serrano, 1973, pp. 9-12, 35, and 61.

2/ Because the living children are only a sample of the living population, they are sometimes referred to in this report as the sample.

13 living children were deleted, leaving 1,989 deaths and 878 living children for the analysis.

The regression analysis of the San Juan data is described in the following chapters. Chapters II and III provide some background for the subsequent chapters about mortality. Chapter II presents the results of an analysis of the determinants of birthweight, a major factor in the survival of very young infants. Chapter III discusses the application of a nutritional classification system that we developed for infants under one year in order to summarize the diet information available from the survey.

Chapters IV through VII contain analyses of the mortality of different groups of children. Chapter IV analyzes the causes of death during the neonatal period i.e., the first four weeks after birth. Chapter V takes up the determinants of mortality among infants 28 days to 5 months of age. Chapter VI continues the analysis by age group and considers infants 6 to 11 months of age and children between one and four years. Chapter VII separates out infants between 28 days and 5 months for whom diarrheal disease, a major cause of death in all the project sites, was a cause of death.

Chapter VIII reviews the results and discusses some of their implications for policy.

A number of technical matters have been relegated to appendices. Appendix A discusses the selection of the observations. Appendix B describes the way variables were defined for the analysis and gives a complete list of the variables used. Appendix C discusses the weighting procedure and sets out the weights for each subgroup that was analyzed. Appendix D contains the bibliography.

CHAPTER II. The Determinants of Birthweight

Introduction

An infant's birthweight is an important measure of its condition at birth. It is, in a sense, a summary of the infant's prenatal history, and as such it is a widely accepted indicator of its chances for survival. For both of these reasons, multivariate regression analysis was used to investigate the determinants of birthweight in the San Juan sample.

Of the 2,867 records used in this study,[1/] valid information about birthweight was available for 2,591 -- 1,789 deaths (89.9 percent of the deaths) and 802 living children (91.3 percent). Birthweight information was available from two sources for the deaths, a medical record and the home interview; the value taken from the medical record was given preference on the assumption that it was more accurate.[2/] The home interview was the only source of information for the living children, for this as for all other variables.

The weighted mean of the birthweights is 3,397 grams, or about 7.5 pounds. Comparison of this average with mean birthweights for the United States and Britain suggests that the San Juan birthweights may be a little inflated.

1/ See Appendix A for a discussion of the selection of the observations.

2/ As a test of the validity of the birthweights reported on the home interviews, the weights reported by both sources were correlated. The correlations, presented and discussed in the feasibility study (Russell and Burke, 1974, Part Two), were quite high.

For example, Pethybridge, et.al. (1974) calculated a mean of 3,383 grams for single births in the "primary distribution" (those over 2,636 grams in weight) for southwest England in 1965. A study of all live births in New York City in 1968 found a mean birthweight of 3,161, with whites averaging 3,213 and blacks averaging 3,007 (Kessner, et.al., 1973). St. George, et.al. (1970) report a mean of 3,163 grams for a group of births in Trinidad, and compare it to averages reported by other studies of 3,200 for the United States and 3,320 for England and Wales.

The distributions of birthweights for four age groups of children who died are compared with the distribution for the living sample in Charts II.1 through II.4. Perhaps the most noticeable thing about the distributions is that, instead of forming smooth curves, they follow a choppy up-and-down pattern, peaking at intervals of about 200-250 grams. This is apparently a very common result and is due to the fact that people tend to round off the weight actually registered by the scales.[1/] In their study of England, Pethybridge, et.al. (1974) found a strong tendency for birthweights to cluster at the zero and eight ounce points, and to a lesser extent at the even ounces in between. The pattern was present in both home and hospital births, although it was more marked for the home births.

The charts show a series of clear and logical differences between the birthweight distributions for the dead infants and the distribution for the living ones. For example, the distribution for infants who died within a week of birth peaks considerably to the left of the normal distribution, showing a heavy concentration of very low birthweights (Chart II.1). Interestingly, it also shows a second, lesser peak well within the normal range. This suggests that some of the very young deaths may have been infants with some serious anomaly at birth,

---

1/ In the present case, the pattern may be further affected by the fact that birthweights were entered only to the nearest hundred grams on the survey questionnaire.

unrelated to their weight. The mode of the distribution for infants who died between 7 and 27 days is again at a birthweight lower than that for the living children, but somewhat higher than the one for deaths under a week; it too is skewed to the right (Chart II.2). Finally, the distribution for deaths between 28 days and 5 months is very similar to the normal distribution, except that it is displaced slightly to the left, while the distribution for deaths between 6 and 11 months exactly reflects the normal distribution (Charts II.3 and II.4). Taken together, the charts suggest that birthweight is unlikely to be a major determinant of mortality after the first few months of life.

The Regressions: Specification and General Discussion

The literature review in the feasibility report (Russell and Burke, 1974, Part One) discussed a number of factors that are believed to affect birthweight, or that are associated with high infant mortality and may have their effects indirectly through birthweight. The following categories of factors were discussed: sex of the child; age of the mother; reproductive history of the mother; complications of pregnancy; spacing between births; medical care before and during the birth; and maternal nutrition.

A series of regressions were specified to test the importance of these factors and measure their effects. The regressions are based on the records of all children in the San Juan survey who were under two years of age at death or interview, and for whom information about birthweight was available. There is a total of 2,084 such children in the survey -- 1,701 deaths and 383 from the sample of living children. The set of observations is limited to children under two years because this group most nearly matches the conceptually desirable group -- the population of all live births (which is equivalent to the population of birthweights) during a given time period. The full survey, which includes children up to five years of age, is incomplete because children

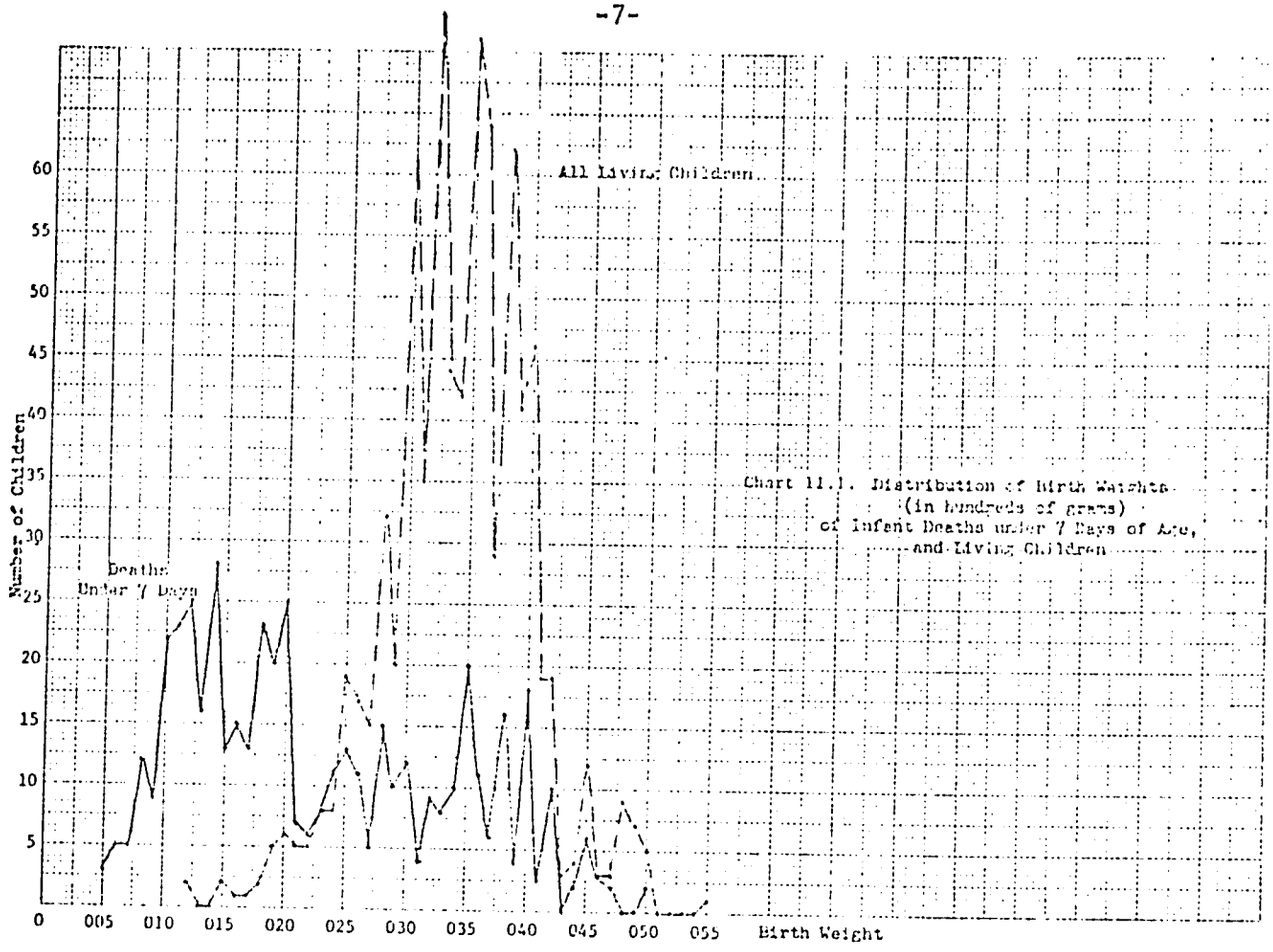


Chart II.1. Distribution of Birth Weights (in hundreds of grams) of Infant Deaths under 7 Days of Age, and Living Children

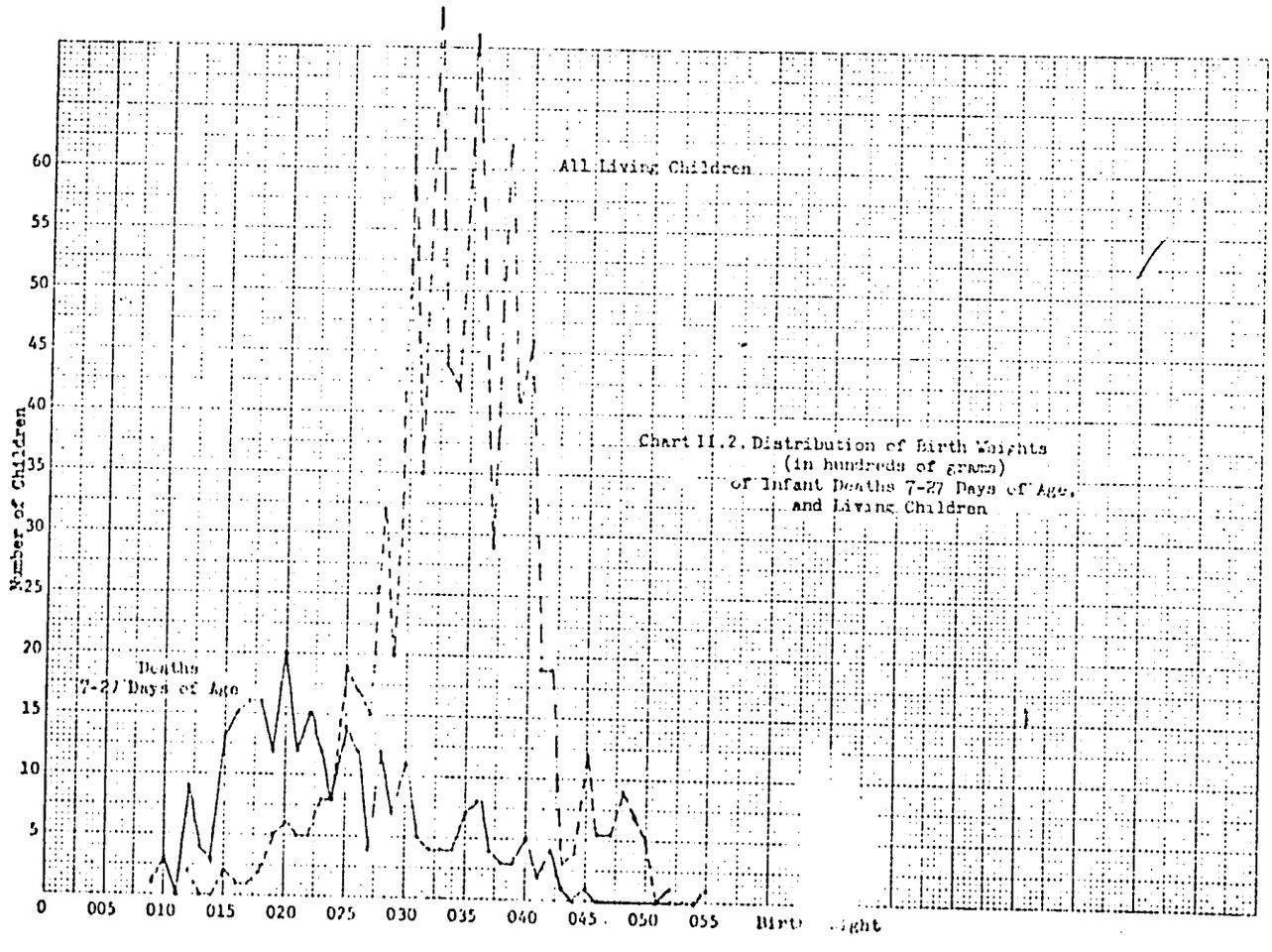


Chart II.2. Distribution of Birth Weights (in hundreds of grams) of Infant Deaths 7-27 Days of Age, and Living Children

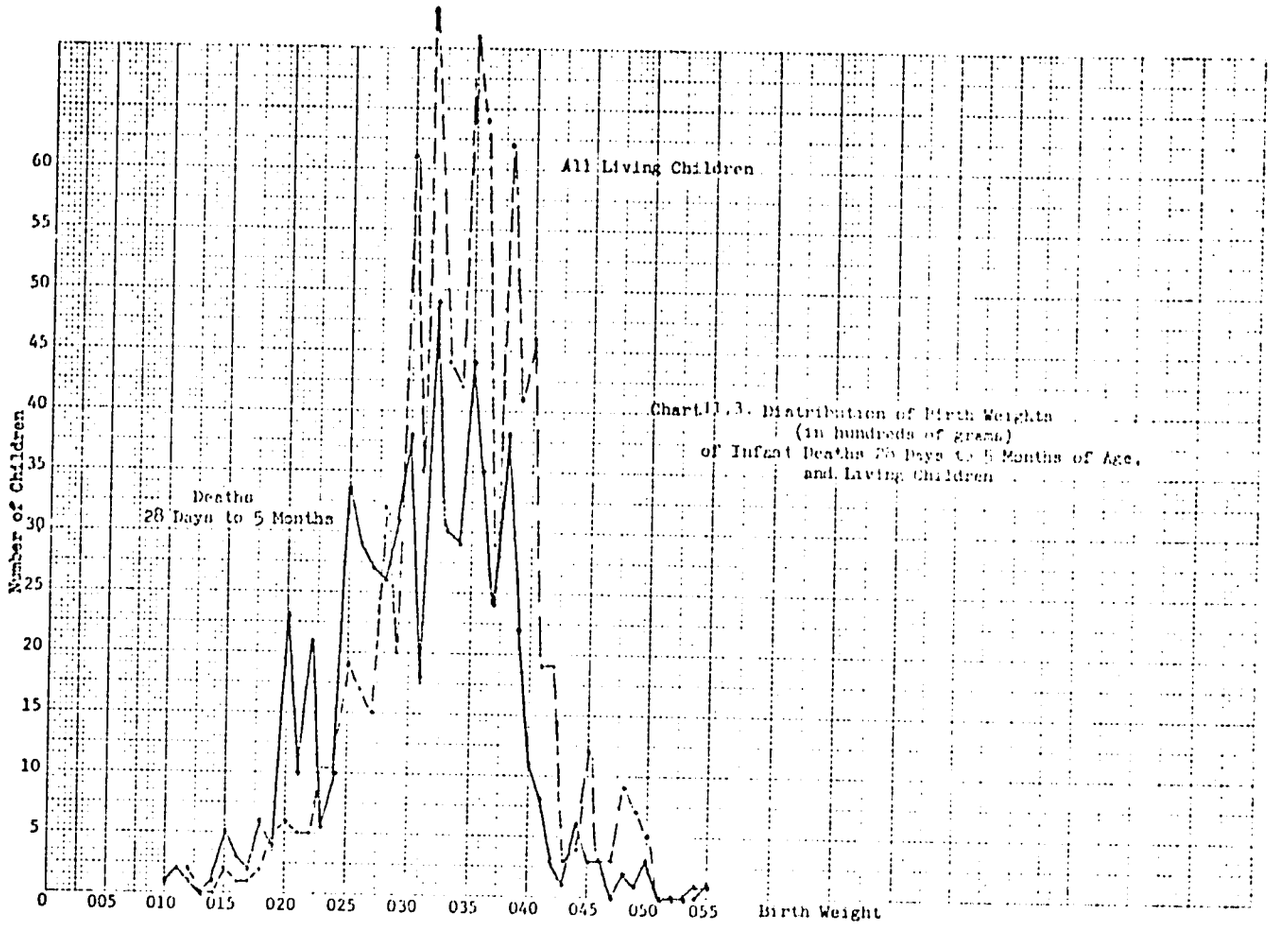


Chart II.3. Distribution of Birth Weights (in hundreds of grams) of Infant Deaths 28 Days to 5 Months of Age, and Living Children

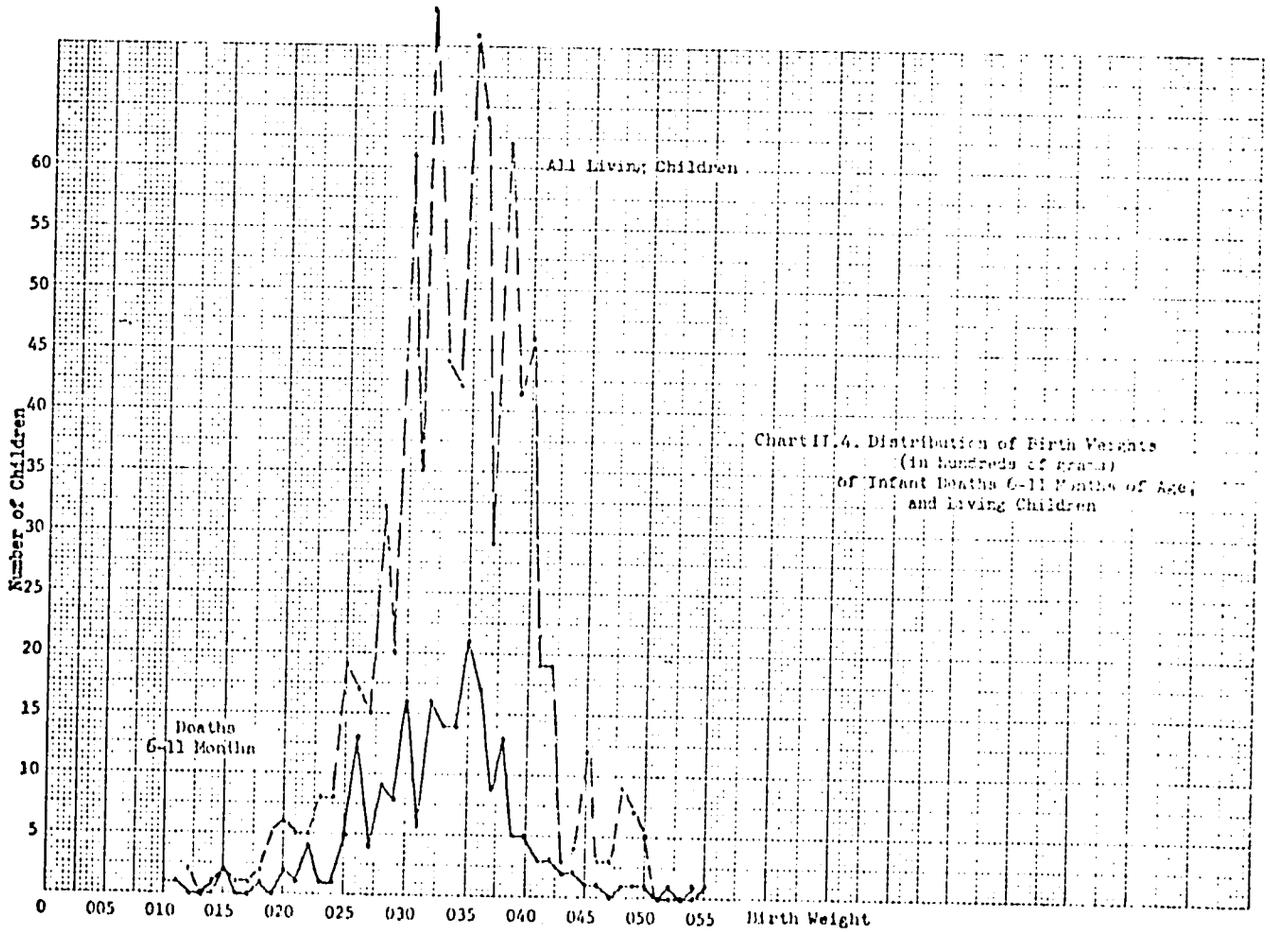


Chart II.4. Distribution of Birth Weights (in hundreds of grams) of Infant Deaths 6-11 Months of Age, and Living Children

who were born and died during the three years preceding the two year survey period are omitted.[1/]

Equations II.1 through II.6 present the regression results. Equation II.1, which may be thought of as the basic equation, contains variables representing all the factors for which good measures were available. Equations II.2 and II.3 differ only in that the former contains an alternate representation of the prenatal care factor, while the latter incorporates information about the outcome of the previous pregnancy. Since no information about maternal nutrition was available from the survey, several socio-economic factors were tested as possible proxies, in the hopes that the net effects of the other factors would be measured more accurately when these proxies were included. Equation II.4 presents the most interesting of these regressions.

For obvious reasons it was not possible to include in the regressions variables representing the interval between births and the outcome of the previous pregnancy, when the set of observations included first births. Thus a second set of observations was defined -- all children under two years of age for whom birthweight data were available and who were also of the second or higher birth order.[2/] Equation II.5 is the basic equation for this set of observations. It resembles Equation II.1 except that variables representing the interval between births and the outcome of the previous pregnancy are also included. Equation II.6 examines the effects of some proxies for maternal nutrition on the specification of II.5.

The variables in the equations are in the binary form; they are defined

---

1/ See Appendix C for a fuller discussion of this point as well as for the observation weights used in the regressions.

2/ The weights used for this set of observations are given in Appendix C.

in Appendix B. In this form a factor,[1/] such as the age of the mother, is divided into several mutually exclusive classes -- for example, under 17, 17-20, 21-25, etc., and a series of variables are defined for each record which take on the value one if the observation falls in that class and the value zero if it does not. If there are n mutually exclusive classes, variables are defined for n-1 of the classes. The omitted class is uniquely represented by values of zero for each of the n-1 variables. The important thing to keep in mind is that the regression coefficients show the effect (in this case, the difference in birthweight[2/]) associated with membership in the particular class, <u>as compared to membership in the omitted class</u>.[3/]

The reliability of the result for a particular variable depends in part on the number of observations that fall in the interval represented. If the coefficient is counter-intuitive and it turns out that relatively few observations fall in that category, we are justified in viewing the result as doubtful. The variables were defined to avoid this problem but, in some cases, a particular class or interval was of such interest that a variable was created for it even though the number of observations was relatively small. For the reader's information, tables showing the distributions, across the classes defined for each factor, of the two sets of observations used for the regressions are given in the appendix to this chapter.

With this background established, we can turn to the examination of the regression results. As the equations demonstrate, most of the factors tested are statistically significant. The coefficients of determination (R^2) are quite

---

1/ Since several variables are often necessary to represent the classes of a given factor, the terms will be given separate meanings here. A factor will always refer to something like the age of the mother or the length of pregnancy, while the term variables will refer specifically to the binary or binaries used to represent the factor.

2/ Since birthweight is measured in hundreds of grams a coefficient of, say, 1.297 represents 129.7 grams.

3/ Thus we can estimate the birthweight for infants with given characteristics by adding or subtracting the differentials associated with those characteristics from the constant term for the equation. Using Equation II.1 as an example, the estimate for female infants whose mothers had 1 or 2 visits for prenatal care but who, in all other respects, have the characteristics represented by the omitted class, is 31.360 (the constant term 32.885 minus 1.943, the coefficient for SEX, plus 0.418, the coefficient for VST1).

EQUATION 11.1: Basic Equation

Dependent Variable: Birthweight, in hundreds of grams.

Set of Observations: All records with valid birthweights and under 2 years of age at the time of the interview or death.

Name of Factor	Name of Variables	Coefficient	t-Score	F for Factor (See Note)
Sex of Child	SEX	-1.943	7.99[b]	63.79[b]
Age of Mother	AGM1	-4.062	4.28[b]	8.55[b]
	AGM2	-1.351	3.42[b]	
	AGM3	0.840	2.38[a]	
	AGM4	1.069	2.68[b]	
	AGM5	-0.551	0.91	
	AGM6	1.836	2.43[a]	
Number of Pregnancies (Parity)	TPRG1	0.976	2.67[b]	4.52[b]
	TPRG2	1.297	3.55[b]	
	TPRG3	0.838	1.58	
	TPRG4	2.069	3.47[b]	
Number of Prenatal Visits	VST1	0.418	0.88	10.36[b]
	VST2	1.356	3.05[b]	
	VST3	1.972	4.61[b]	
Length of Pregnancy	PLG1	-20.259	10.80[b]	67.82[b]
	PLG2	-5.608	4.74[b]	
	PLG3	-3.506	8.62[b]	
Place of Birth	BPLC1	-0.352	0.73	0.53
Who Attended Birth	ATND1	0.123	0.43	0.61
	ATND2	0.726	1.11	
Single or Multiple Birth	MULT1	-7.034	8.74[b]	76.31[b]
Constant Term	-	32.885	-	-

R^2 for the equation: .220
F for the equation: 27.72
a. significant at the 95 percent level of confidence.
b. significant at the 99 percent level of confidence.

NOTE: Each F-value indicates whether the factor as a whole is statistically significant. When a factor is represented by a single variable the F is simply the square of the t-score. When a factor is represented by more than one variable it turns out that, in general, the factor is significant whenever at least one of the variables has a significant t-score. Limitations on computer time made it impossible to calculate these F's for the equations other than the basic equation in each chapter, but the reader can tell when a factor is significant by examining the t-scores.

EQUATION II.2: Basic Equation with an Alternate Variable for Number of Prenatal Visits

Dependent Variable: Birthweight, in hundreds of grams.

Set of Observations: All records with valid birthweights and under 2 years of age at the time of the interview or death.

Name of Factor	Name of Variables	Coefficient	t-score
Sex of Child	SEX	-1.969	8.09[b]
Age of Mother	AGM1	-4.232	4.45[b]
	AGM2	-1.445	3.65[b]
	AGM3	1.018	2.89[b]
	AGM4	1.236	3.10[b]
	AGM5	-0.300	0.50
	AGM6	2.047	2.70[b]
Number of Pregnancies (Parity)	TPRG1	0.874	2.39[a]
	TPRG2	1.225	3.35[b]
	TPRG3	0.649	1.23
	TPRG4	1.950	3.26[b]
Number of Prenatal Visits	VSTX	1.415	3.53[b]
Length of Pregnancy	PLG1	-20.376	10.83[b]
	PLG2	-5.923	5.00[b]
	PLG3	-3.437	8.43[b]
Place of Birth	BPLC1	-0.331	0.69
Who Attended Birth	ATND1	-0.176	0.63
	ATND2	0.452	0.69
Single or Multiple Birth	MULT1	-7.023	8.71[b]
Constant Term	-	33.140	-

R^2 for the equation: .213
F for the equation: 29.42
a. significant at the 95 percent level of confidence.
b. significant at the 99 percent level of confidence.

EQUATION 11.3: Basic Equation with Alternate Variables for Number of Pregnancies

Dependent Variable: Birthweight, in hundreds of grams.

Set of Observations: All records with valid birthweights and under 2 years of age at the time of interview or death.

Name of Factor	Name of Variables	Coefficient	t-score
Sex of Child	SEX	-1.719	6.99[b]
Age of Mother	AGM1	-3.517	3.74[b]
	AGM2	-0.966	2.46[a]
	AGM3	0.825	2.35[a]
	AGM4	1.230	3.09[b]
	AGM5	-0.644	1.06
	AGM6	1.578	2.09[a]
Number of Pregnancies cross-classified by outcome of previous pregnancy	TPC1	6.559	6.21[b]
	TPC2	1.463	3.75[b]
	TPC3	-0.492	0.37
	TPC4	-0.199	0.29
	TPC5	1.869	5.10[b]
	TPC6	2.923	3.42[b]
	TPC7	-1.849	1.76
	TPC8	1.749	3.26[b]
	TPC9	-2.115	1.01
	TPC10	5.238	4.32[b]
	TPC11	2.165	3.52[b]
	TPC12	-0.556	0.35
Number of Prenatal Visits	VST1	0.361	0.76
	VST2	1.439	3.26[b]
	VST3	2.104	4.94[b]
Length of Pregnancy	PLG1	-20.304	11.00[b]
	PLG2	-5.528	4.74[b]
	PLG3	-3.473	8.66[b]
Place of Birth	BPLC1	0.130	0.27
Who Attended Birth	ATND1	0.047	0.17
	ATND2	0.498	0.77
Single or Multiple Birth	MULT1	-7.579	8.89[b]
Constant Term	-	32.361	-

R^2 for the equation: .249
F for the equation: 23.45
a. significant at the 95 percent level of confidence.
b. significant at the 99 percent level of confidence.

EQUATION II.4: Basic Equation with Fathers' Occupation and Type of Neighborhood

Dependent Variable: Birthweight, in hundreds of grams.

Set of Observations: All records with valid birthweights and under 2 years of age at the time of interview or death.

Name of Factor	Name of Variables	Coefficient	t-score
Sex of Child	SEX	-1.994	8.12[b]
Age of Mother	AGM1	-4.095	4.16[b]
	AGM2	-1.069	2.69[b]
	AGM3	1.019	2.86[b]
	AGM4	0.980	2.34[a]
	AGM5	-0.813	1.34
	AGM6	1.699	2.20[a]
Number of Pregnancies (Parity)	TPRG1	0.732	1.96[a]
	TPRG2	1.141	3.04[b]
	TPRG3	0.613	1.13
	TPRG4	2.098	3.36[b]
Number of Visits	VST1	0.368	0.77
	VST2	1.352	3.02[b]
	VST3	1.948	4.46[b]
Length of Pregnancy	PLG1	-20.086	10.89[b]
	PLG2	-5.421	4.61[b]
	PLG3	-3.596	8.62[b]
Place of Birth	BPLC1	-0.566	1.16
Who Attended Birth	ATND1	0.681	2.26[a]
	ATND2	1.421	2.13[a]
Single or Multiple Birth	MULT1	-6.687	8.09[b]
Neighborhood	NBR1	1.073	2.39[a]
	NBR2	-0.679	1.91
	NBR3	-0.522	0.65
	NBR4	-0.172	0.50
Occupation of Father	OCF1	0.395	0.76
	OCF2	1.551	2.64[b]
	OCF3	2.297	3.[illegible][b]
	OCF4	2.137	3.[illegible]
	OCF5	-0.369	0.[illegible]
	OCF6	-1.288	1.[illegible]
	OCF7	-0.469	0.[illegible]

EQUATION II.4, continued

Name of Factor	Name of Variables	Coefficient	t-score
Occupation of Father, continued	OCF8	-0.444	0.81
	OCF9	-0.008	0.00
	OCF10	-0.403	0.60
	OCF11	-2.353	2.27[a]
	OCF12	2.390	4.74[b]
	OCF13	0.091	0.12
	OCF14	-1.317	8.72[b]
	OCF15	1.328	2.35[a]
Constant Term	-	32.387	-

R^2 for the equation: .253
F for the equation: 17.32
a. significant at the 95 percent level of confidence.
b. significant at the 99 percent level of confidence.

EQUATION II.5: Basic Equation for Children of Parity Two or Higher

Dependent Variable: Birthweight, in hundreds of grams

Set of Observations: All records of children with valid birthweights who were under two years of age at the time of interview or death, and were not first children.

Name of Factor	Name of Variable	Coefficient	t-score	F for Factor (see note)
Sex of Child	SEX	-1.962	6.85[b]	46.92[b]
Age of Mother	AGM1	-3.427	2.42[a]	4.15[b]
	AGM2	-0.937	1.79	
	AGM3	0.781	1.85	
	AGM4	1.105	2.48[a]	
	AGM5	-0.494	0.73	
	AGM6	1.744	2.14[a]	
Number of Pregnancies (Parity)	TPRG2	0.436	1.27	1.85
	TPRG3	-0.085	0.16	
	TPRG4	1.069	1.82	
Outcome of Previous Pregnancy	CL1	0.118	0.24	0.06
	CL2	0.177	0.27	
Interval Since Last Birth	SPC1	-1.524	2.52[a]	4.33[b]
	SPC2	0.434	0.79	
	SPC3	0.669	1.32	
	SPC4	1.466	3.02[b]	
	SPC5	0.790	1.51	
	SPC6	0.342	0.64	
Number of Prenatal Visits	VST1	0.891	1.62	8.13[b]
	VST2	1.200	2.37[a]	
	VST3	2.221	4.49[b]	
Length of Pregnancy	PLG1	-21.356	8.73[b]	53.18[b]
	PLG2	-4.431	3.32[b]	
	PLG3	-4.243	8.85[b]	
Place of Birth	BPLC1	0.102	0.18	0.03
Who Attended Birth	ATND1	0.064	0.19	0.47
	ATND2	0.677	0.96	
Single or Multiple Birth	MULT1	-8.649	8.44[b]	71.23[b]
Constant Term		32.368		

R^2 for the equation: .218
F for the equation: 15.41
a. significant at the 95 percent level of confidence.
b. significant at the 99 percent level of confidence.

EQUATION II.5, continued

NOTE: Each F-value indicates whether the factor as a whole is statistically significant. When a factor is represented by a single variable the F is simply the square of the t-score. When a factor is represented by more than one variable it turns out that, in general, the factor is significant whenever at least one of the variables has a significant t-score. Limitations on computer time made it impossible to calculate these F's for equations other than the basic equation in each chapter but the reader can tell when a factor is significant by examining the t-scores.

EQUATION II.6: Basic Equation for Children of Parity Two or Higher with Father's Occupation and Type of Neighborhood

Dependent Variable: Birthweight, in hundreds of grams.

Set of Observations: All records of children with valid birthweights who were under two years of age at the time of interview or death, and were not first children.

Name of Factor	Name of Variables	Coefficient	t-score
Sex of Child	SEX	-2.209	7.71[b]
Age of Mother	AGM1	-2.913	1.99[a]
	AGM2	-0.688	1.30
	AGM3	1.201	2.81[b]
	AGM4	1.240	2.70[b]
	AGM5	-0.425	0.63
	AGM6	1.938	2.34[a]
Number of Pregnancies (Parity)	TPRG2	0.546	1.57
	TPRG3	-0.313	0.60
	TPRG4	1.096	1.81
Outcome of Previous Pregnancy	CL1	-0.284	0.57
	CL2	0.328	0.50
Interval Since Last Birth	SPC1	-1.370	2.26[a]
	SPC2	-0.148	0.27
	SPC3	0.934	1.82
	SPC4	1.479	3.01[b]
	SPC5	0.099	0.19
	SPC6	0.028	0.05
Number of Prenatal Visits	VST1	1.094	2.00[a]
	VST2	1.432	2.84[b]
	VST3	2.475	4.96[b]
Length of Pregnancy	PLG1	-21.040	8.83[b]
	PLG2	-3.865	2.92[b]
	PLG3	-4.454	8.93[b]
Place of Birth	BPLC1	-0.351	0.62
Who Attended Birth	ATND1	0.588	1.62
	ATND2	1.281	1.79
Single or Multiple Birth	MULT1	-8.612	8.19[b]
Type of Neighborhood	NBR1	2.290	4.54[b]
	NBR2	0.537	1.25
	NBR3	0.808	0.94
	NBR4	1.254	2.99[b]

EQUATION II.6, continued

Name of Factor	Name of Variables	Coefficient	t-score
Occupation of Father	OCF1	1.377	2.17[a]
	OCF2	1.789	2.60[b]
	OCF3	2.100	2.74[b]
	OCF4	2.447	3.05[b]
	OCF5	-1.041	1.65
	OCF6	-0.357	0.22
	OCF7	-0.813	0.99
	OCF8	-0.224	0.33
	OCF9	-0.701	1.15
	OCF10	-0.824	1.08
	OCF11	-2.308	2.14[a]
	OCF12	3.449	6.16[b]
	OCF13	0.999	1.19
	OCF14	-0.699	0.72
	OCF15	1.081	1.78
Constant Term	-	31.521	-

R^2 for the equation: .269
F for the equation: 11.96
a. significant at the 95 percent level of confidence.
b. significant at the 99 percent level of confidence.

respectable for regressions based on large household surveys and indicate that the included variables explain between 21 and 28 percent of the variation in birthweights.

The results shown in Equations II.1 through II.6 will be taken up in detail in the following sections.

Sex of the Child

The review of the literature revealed that female infants have lower mortality rates than males, particularly during the neonatal period (0-27 days), and these findings initially led us to hypothesize that females have higher birthweights. But our reading since then has indicated that females are generally lighter at birth (St. George, et.al., 1970; Pethybridge, et.al., 1974). This is borne out by the coefficient for sex in the regressions: for example, the coefficient in Equation II.1 shows that female infants in San Juan are about 194 grams lighter than male infants.[1/] The coefficient for this factor is highly significant and fairly stable in value across equations and sets of observations, ranging from a low of -1.719 (172 grams) in Equation II.3 to a high of -2.209 (221 grams) in Equation II.6.

Age of the Mother

Studies of neonatal and infant mortality show that these death rates vary with the age of the mother. There is general agreement that infant mortality is higher for older mothers -- those over 35. There is also evidence that very young mothers experience high infant mortality rates, although

---

1/ The reader should recall that a regression coefficient gives the net effect of the variable, holding constant for the effects of all the other variables included in the equation.

the exact age that marks the danger line is not clear. After careful review of past studies the National Academy of Sciences (1970) concluded that mothers under 17 show sharply higher infant mortality rates, while the risks for mothers 17 to 19 are similar to those for mothers 20 to 24 years of age.

We hypothesized that birthweight might be the mechanism through which the age of the mother influences her infant's risk of death. Accordingly, we entered maternal age in the regressions. Following the National Academy of Sciences we specified the age classes as under 17 (AGM1), 17 to 20 (AGM2), 26 to 30 (AGM3), 31 to 35 (AGM4), 36 to 40 (AGM5), and over 40 (AGM6), with 21 to 25 the omitted class.

As Equation II.1 shows, younger mothers in the San Juan survey do have significantly lighter babies than older mothers. (The age coefficients from Equation II.1 are graphed in Chart II.5.)[1/] The babies of mothers under 17 are 406 grams -- more than 14 ounces -- lighter than the babies of women in the 21 to 25 year age bracket. Mothers 17 to 20 years of age are still at a significant disadvantage, although the differential is much less -- about 135 grams, or just under 5 ounces.

The babies of mothers 26 to 35 years of age (AGM3 and AGM4) are significantly heavier than those of younger mothers. But the advantage disappears for mothers 36 to 40 -- the coefficient for AGM5 is statistically insignificant and negative. Unexpectedly the coefficient for AGM6 (over 40) is again significantly positive. This result may be due to the idiosyncrasies of the few cases in this group. While 64 deaths fall in the over 40 class, only 13 of the more heavily weighted living children do (see Table 3 in the appendix to this chapter).

It is interesting to compare Equation II.1 with Equations II.5 and II.6. The results for the older mothers are about the same, but the coefficients for younger mothers (AGM1 and AGM2) are somewhat reduced in size and significance. Reasonably enough, since Equations II.5 and II.6 omit first births, many fewer mothers

---

1/ The birthweights shown on Charts II.5-II.9 have been derived by adding or subtracting the differential associated with the particular characteristic from the average birthweight for infants in the omitted class, rather than from the constant term of the equation. Using the mean implies that the characteristics of the infants in the omitted class, other than the one of interest in the chart, were those that actually occurred in the population. The constant term implies a comparison with infants all of whose characteristics are those of the omitted classes.

Chart II.5. Relationship Between Birthweight and Age of Mother at Birth. (Based on coefficients for AGM1-AGM6, Equation II.1.)

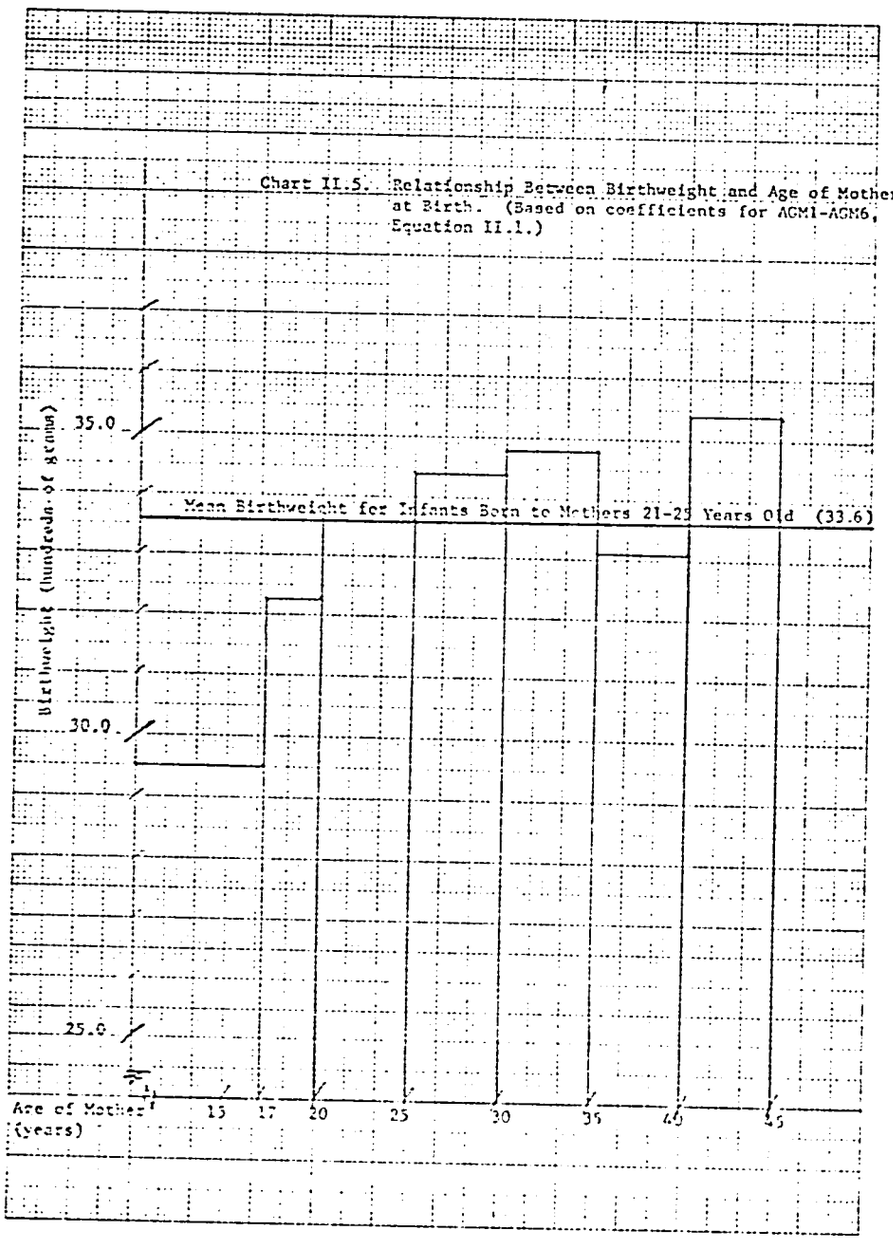


Chart II.6. Relationship Between Birthweight and Number of Pregnancies. (Based on coefficients for TPRG1-TPRG4, Equation II.1.)

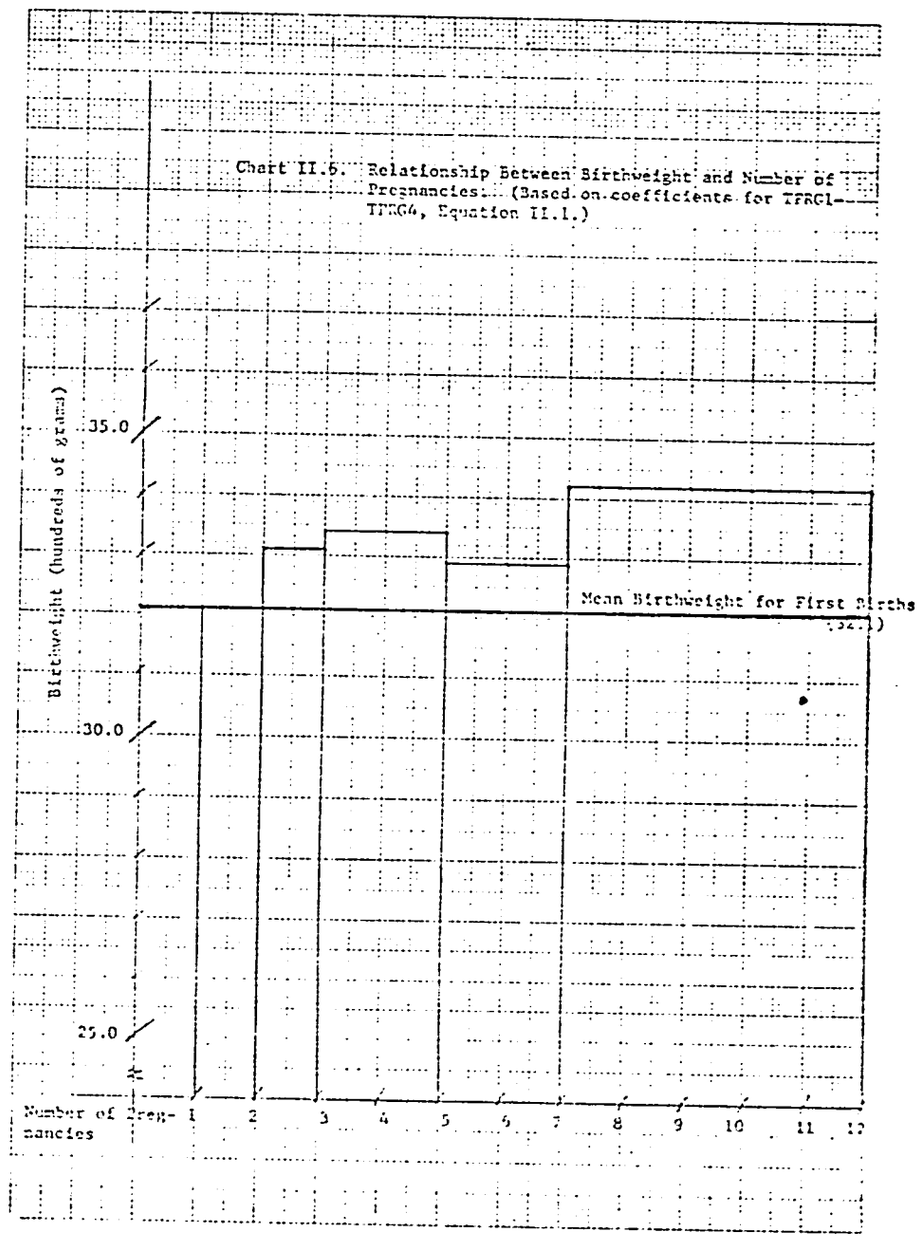


Chart II.7. Relationship Between Birthweight and Length of Pregnancy.(Based on coefficients for PLG1-PLG3, Equation II.1.)

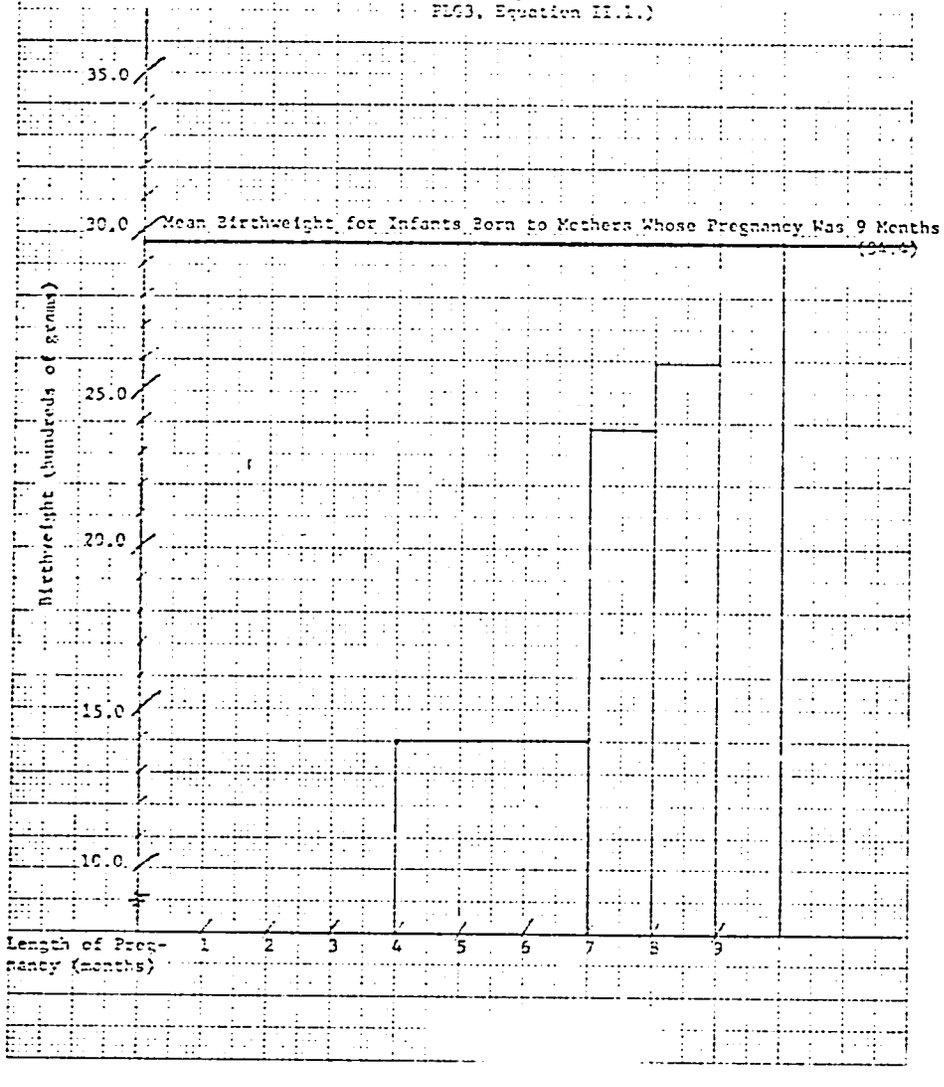
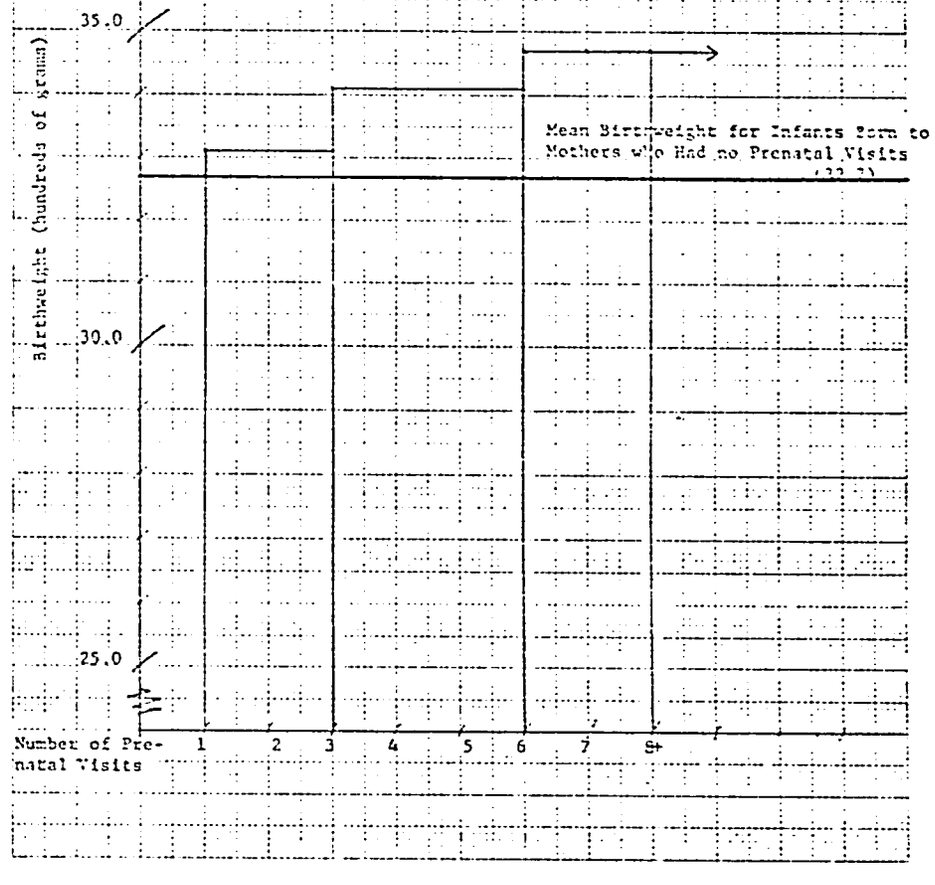


Chart II.8. Relationship Between Birthweight and Number of Prenatal Visits.(Based on coefficients for VST1-VST3, Equation II.1.)



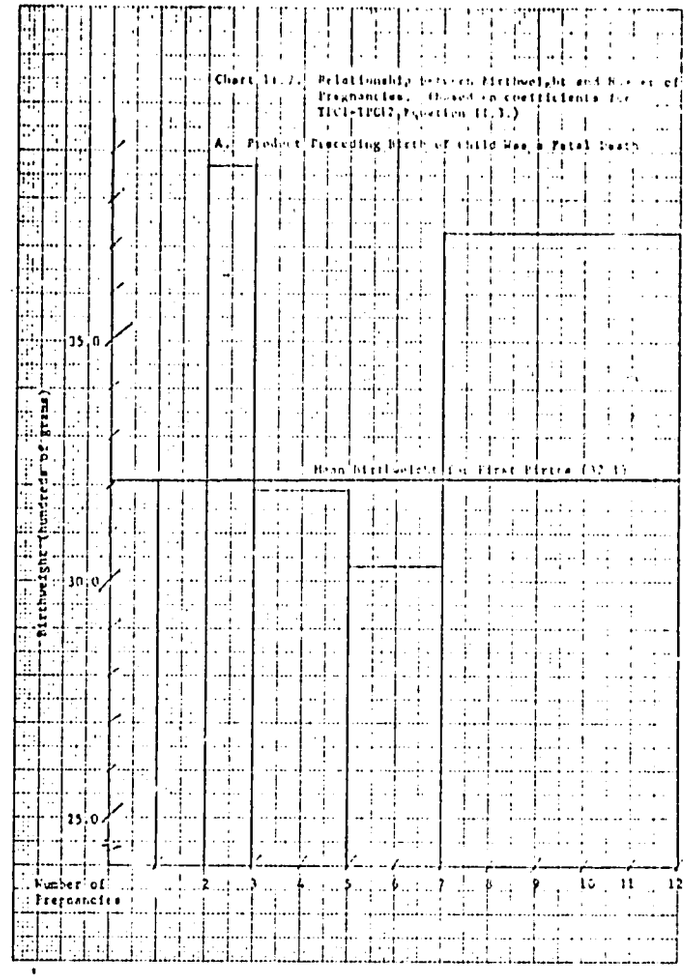


Chart II.7. Relationship between Birthweight and Number of Pregnancies. (Based on coefficients for TIC1-TPC12, Equation II.1.)
A. Product Preceding Birth of Child Was a Fetal Death
Mean Birthweight for First Births (32.1)
Birthweight (hundreds of grams)
35.0
30.0
25.0
Number of Pregnancies
2
3
4
5
6
7
8
9
10
11
12

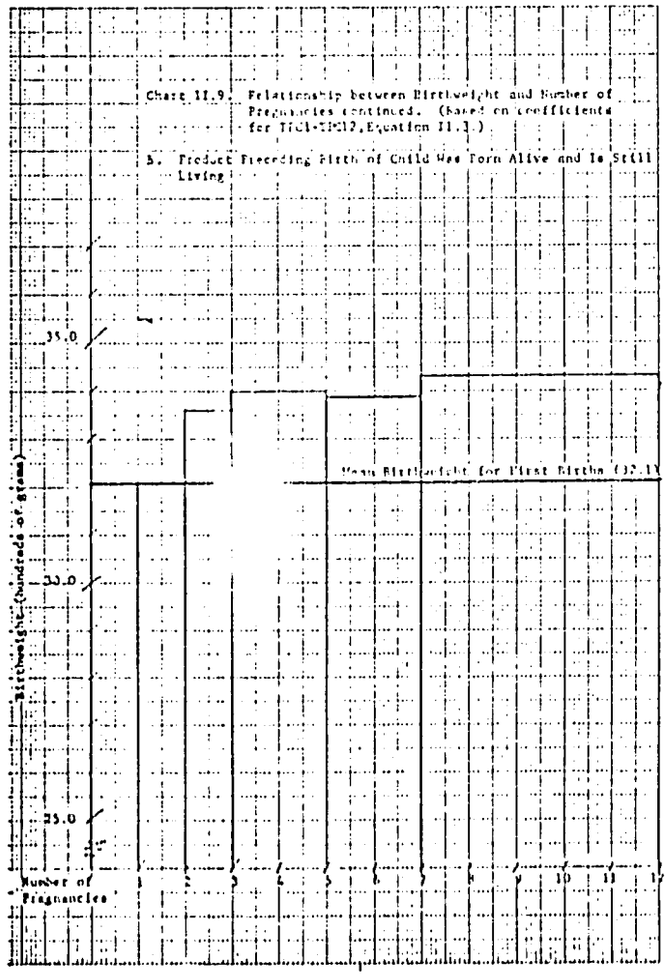


Chart II.9. Relationship between Birthweight and Number of Pregnancies continued. (Based on coefficients for TIC1-TPC12, Equation II.1.)
B. Product Preceding Birth of Child Was Born Alive and Is Still Living
Mean Birthweight for First Births (32.1)
Birthweight (hundreds of grams)
35.0
30.0
25.0
Number of Pregnancies
1
2
3
4
5
6
7
8
9
10
11
12

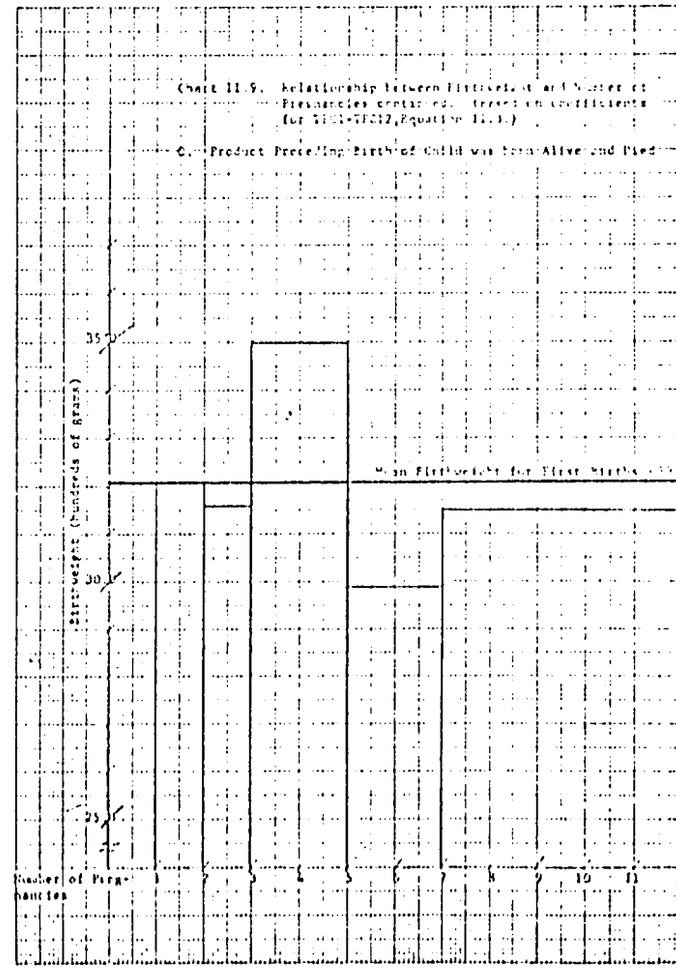


Chart II.9. Relationship between Birthweight and Number of Pregnancies continued. (Based on coefficients for TIC1-TPC12, Equation II.1.)
C. Product Preceding Birth of Child was Born Alive and Died
Mean Birthweight for First Births
Birthweight (hundreds of grams)
35.0
30.0
25.0
Number of Pregnancies
1
2
3
4
5
6
7
8
9
10
11
12

in this set of observations fall in these young age classes, and those that remain tend to fall towards the upper end of the interval. These facts are undoubtedly a large part of the explanation for the differences between Equations II.1, and II.5 and II.6. But the changes also suggest the possibility that there is an interaction between parity and maternal age that makes the effect of the two factors combined greater than the sum of their individual effects.

Taken together these results agree in a general way with the tabulations presented by St. George, et.al. (1970). Their tabulations break the age range into fewer classes than we used, but show birthweights for mothers under 20 to be considerably less than those for older mothers, both in Trinidad and in England and Wales. Birthweights increase steadily with maternal age through the 30's in both countries, and on into the 40's in England, but they decline after 40 in Trinidad. Pethybridge, et.al. (1974) also report average birthweights increasing with maternal age, but their data show a small (although statistically insignificant) decline after the age of 35.

Reproductive History of the Mother (Number of Pregnancies and Outcome of the Previous Pregnancy)

Information about two aspects of each mother's reproductive history was available from the survey -- the total number of her pregnancies and the outcome of each pregnancy (whether the infant died, and its age at death). Equation II.1 includes variables representing the total number of pregnancies. First children are the omitted class, and the four variables divide the rest of the range into two pregnancies (TPRG1), three or four pregnancies (TPRG2), five or six pregnancies (TPRG3), and seven or more pregnancies (TPRG4).

The coefficients for these variables in Equation II.1 show that children of higher parities are, without exception, heavier than first children. Chart II.6 graphs these results. While the differential is uneven -- rising, falling, and then rising again, the differences between the higher parity groups are not significant at the usual confidence levels.[1/] This agrees with the tabular findings of St. George, et.al.(1970) and Pethybridge, et.al. (1974), which show a substantial difference between first and subsequent children, but smaller and irregular differences between children of higher parities.

In the case of first births there is, of course, no previous pregnancy, hence the definition of variables representing outcome of the previous pregnancy presents difficulties when first births are included in the observation set. Equation II.3 offers one way around this problem. A set of variables was included in this equation which represents a cross-classification of the total number of pregnancies with the outcome of the previous pregnancy. First births are again the omitted class. The variables TPC1 through TPC12 denote the cells of the cross-classification -- for example, TPC8 denotes all cases of fifth or sixth pregnancies in which the previous child was born alive and was still living at the time of interview. TPC1, TPC4, TPC7, and TPC10 represent cases in which the previous pregnancy ended in a fetal death. TPC3, TPC6, TPC9, and TPC12 represent cases in which the previous pregnancy ended in a live born child who subsequently died. The remaining variables refer to cases in which the previous child was still living at the time of interview.[2/]

1/ It was noted earlier that the coefficients for binary variables measure differences from the omitted class, and the standard errors and resulting t-scores refer to the statistical significance of that difference. Comparison of any two included classes proceeds like the ordinary test of the difference between two means. The standard error of the difference between the coefficients is approximated by $\sqrt{\sigma_a^2 + \sigma_b^2}$, where σ_a and σ_b denote the standard errors of the coefficients in question (Melichar, 1965). Since $t=\mu/\sigma$ the reader can, of course, derive the needed standard errors by dividing the coefficients by their t-scores.

The results, graphed in Chart II.9, are quite interesting. Part B of the chart shows the results for the four parity classes with previous child born alive and still living. The increase in weight for children beyond the first is statistically significant, as before, but the relationship is considerably smoother than in Equation II.1 (Chart II.6).

The results for cases with death as the outcome of the previous pregnancy are less clearcut. Part A of Chart II.9 shows the differentials by parity when the previous pregnancy ended in a fetal death; Part C of the chart shows the case for death after live birth of the previous child. In both cases the results are mixed, with some parities showing higher birthweights than first children and others showing lower weights.

But when we look again at Part B of Chart II.9 and realize that children at these parities would normally always be significantly heavier than first children, the results are illuminating. While the negative differentials between many of these higher parity infants and first births are not statistically significant at the 95 percent level of confidence, some of the differentials between them and children <u>of the same parity</u> with previous child living are significant. In particular, infants of the third through sixth parities preceded by a fetal death (TPC4 and TPC7) are significantly lighter at birth than children of the same parities preceded by a living child (TPC5 and TPC8). The increase in the R^2 of the equation also suggests that the classification is important.

These results, while interesting, must be taken as highly tentative. As Table 4 in the appendix to the chapter shows, the numbers of cases in some of the classes are quite small, and this is undoubtedly the cause of some of the irregularities in the results.

Equations II.5 and II.6 were based on the set of observations that excluded first births. For these equations separate sets of variables representing number of pregnancies (TPRG2-4) and the outcome of the previous pregnancy (CL1-2) were included. The number of pregnancies shows little relationship to birthweight, in line with the earlier observation that there does not appear to be a significant pattern of differences among higher parity infants. And in its new form the outcome of the previous pregnancy is statistically insignificant, a reasonable result when we recall the irregular pattern that emerged when outcome was cross-classified with parity in Equation II.3.

The survey provides a complete history of all the outcomes of previous pregnancies for each mother. It is already evident that there are problems with using this information -- the fewness of observations in even the simplest classifications that separate previous deaths from living products, and the fact that first children have no preceding products, second children have only one, and so on. But the information is there and other ways of summarizing it might be able to reveal relationships beyond those discussed here.

Complications of Pregnancy (Length of Pregnancy and Multiple Births)

The term "complications of pregnancy" brings to mind conditions like edema, hemorrhage, infectious disease, and Caesarean delivery. The survey tried to collect information about these conditions, but the results were disappointing. For one thing, the information was requested only for deaths, not for living children. But even for the deaths the data actually collected were so poor that it was only possible to say that there had, or had not, been any complications: differentiation by type of complication was impossible.

Because of these deficiencies in the data, no variables for this sort of complication were tested. It would be possible to test whether the presence

of complications (type unknown) affects birthweight by separating out the birthweights for the deaths and computing additional regressions for this set of observations. But given the constraints on programming time, and the poor quality of the information, it did not seem worthwhile to do this.[1/]

Two other characteristics of the pregnancy were available -- its length, and whether the birth was single or multiple. As expected, multiple births are associated with significantly lower birthweights. Equation II.1 shows that multiple births averaged 703 grams (about 1 pound, 9 ounces) less than single births. Equations II.5 and II.6 indicate that the differential is somewhat higher for births beyond the first -- about 860 grams. These results agree closely with the differential between multiple and single births of about 800 grams found by Pethybridge, et.al. (1974).

Length of pregnancy is divided into four classes: six months or less (PLG1), seven months (PLG2), eight months (PLG3), and nine months, with nine months serving as the omitted class. The results are highly significant, as expected. Equation II.1 and Chart II.7 show that pregnancies of six months or less produce babies 2,026 grams (about 4 pounds,8 ounces) lighter than full-term babies, on the average. The differentials decline as the length of pregnancy increases: seven month births average 561 grams (1 pound, 4 ounces) lighter than full-term, while eight month births are 351 grams lighter (about 12 ounces).

The coefficients are stable across the specifications that include first births in the observation set (Equations II.1-II.4). But when first births are excluded, the differential between seven and eight months disappears (Equations II.5 and II.6), and the coefficients for both categories show a

---

1/ Since no information about complications was collected for the sample of living children, it is not possible to test for the direct effects of complications on mortality.

difference from full-term babies of about 400 grams. The PAHO staff warned that the length of pregnancy data might be unreliable when obtained from the home interview and it seems probable that this result reflects inaccuracies in the distinction between seven and eight month gestation periods.[1/]

The results from Equations II.1-II.4 are in general agreement with the findings of St. George et.al. (1970).[2/] Their tabulations for Trinidad show that infants of 28 to 30 weeks gestation, about the sixth month, are 2,052 grams lighter than full-term babies. Infants of 32 to 34 weeks gestation, well into the seventh month, are 1,040 grams lighter and those born during weeks 36 to 38, the eighth month, are 342 grams lighter. The differential for seven month births is much larger in their tabulations than in our equations, but the other differentials are surprisingly close.

Spacing

The literature review in the feasibility study noted that both very close and very distant spacing are associated with higher rates of stillbirth and neonatal mortality (Morris and Heady, 1955). This raises the possibility that an infant's birthweight is influenced by the length of time that elapses between a previous birth and its conception. St. George, et.al. (1970) found no general connection between spacing and birthweight but, on the basis of a few observations, they speculated that repeated closely spaced conceptions resulted in lower birthweights.

1/ In the feasibility report (Russell and Burke, 1974, Part Two) we took advantage of the fact that, for the deaths, length of pregnancy was frequently available from both the home interview and medical records. The relatively low correlations between the two sources (.83 for deaths under one month, but .55 for older deaths) suggest that the home interview was an unreliable source for this information.

2/ Exact comparison is not possible since their findings are broken out by two week intervals, and the information given is insufficient to allow us to combine intervals to match the ones used in the regressions.

In order to investigate the effects of spacing we first had to omit first births from the set of observations. The desired measure of spacing is the time between the birth of one child and the conception of the next, but the measure available is the time between the two births. This measure was divided into seven classes: 12 months or less (SPC1), 13-15 months (SPC2), 16-18 months (SPC3), 19-24 months (SPC4), 25-36 months (the omitted class), 37-48 months (SPC5), and over 48 months (SPC6). SPC1 is intended to represent those infants conceived within three months of the previous birth, but it is an imperfect proxy to the extent that the child of record was premature.

Two of the spacing variables show a statistically significant difference from the omitted class in Equations II.5 and II.6. When the interval between births is 12 months or less, birthweights are significantly reduced -- by 152 grams in Equation II.5 and 137 grams in Equation II.6. Birthweight increases thereafter, and is significantly higher (by 147 grams) than the omitted class when the infant follows a previous birth by 19 to 24 months. Intervals longer than three years are not significantly different from those in the omitted class.

Medical Care Before and During Birth

The initial literature review did not uncover any studies that examined the relationship between medical care and birthweight, and the evidence about links between infant mortality and medical care was sparse and contradictory. But on the basis of their Trinidad sample, St. George, et.al. (1970) report that, when the pregnancy was complicated, infants whose mothers received no prenatal care were substantially lighter than infants whose mothers did receive care.

The survey gathered data on the number of visits for prenatal care and this information was used in the regressions. (There is, however, no information about the content of the visits.) Prenatal visits were split into four classes: one or two visits (VST1); three, four, or five visits (VST2); six or more visits (VST3)[1/]; and the omitted class -- those who had no prenatal care.

The results are of considerable interest both because of the paucity of evidence on this point for developing countries and because of their significance for programs aimed at reducing infant mortality. As Equation 1 shows, prenatal care is associated with a significant increase in the infant's birthweight. One or two visits are statistically insignificant, but three to five visits are associated with an increase in weight of 136 grams and the coefficient is significant at the 99 percent level of confidence. Six or more visits are associated with a still larger weight gain -- 197 grams, or about seven ounces. The difference between the coefficients for VST2 and VST3 is not statistically significant.

Equation II.2 uses an alternate representation of the prenatal care data. VSTX groups all mothers who received some prenatal care and compares them with those who did not. The coefficient shows that, on average, prenatal care resulted in a weight increase of 142 grams.

Comparison of Equation II.1 with Equations II.3 through II.6 -- all of which use the same variables to represent prenatal care -- shows that the coefficients are stable, in both value and significance, across specifications. Equation II.4 is perhaps the most interesting because it includes two socio-economic factors, and one might expect prenatal care to be correlated with these factors thus reflecting their influence whenever they are omitted.

1/ The survey did not distinguish those with eight visits from those with more.

But the coefficient values are virtually identical with those in Equation II.1, as are the t-scores.

Two measures of care received during the actual birth are also included in the regressions -- place of birth and attendant at birth. The coefficient for BPLC1 is never statistically significant, indicating that babies born at home do not differ in weight from those born in the hospital. The attendant variables are statistically insignificant except in Equation II.4. This result need not be taken seriously since it is not supported by the other equations, but it is worth noting two possible explanations for the significantly positive coefficients in this equation. The first possibility is that they may represent a tendency for the birthweights of infants not delivered by physicians to be over-reported. The second is that the result may reflect a tendency for recognized high-risk mothers to be attended by a physician.

Maternal Nutrition

The literature review (Russell and Burke, 1974, Part One) repeatedly turned up evidence of the importance of maternal nutrition to infant birthweight, and our subsequent reading has confirmed this relationship. Without exception, authors who have used the height and weight of the mother as proxies for her nutritional status have found that taller, heavier mothers have heavier babies (St. George, et.al., 1970; Kapur, et.al., 1971; Shah and Shah, 1972). In an investigation of a different aspect of maternal nutrition, Harrison and Ibeziako (1973) found that anemic mothers whose anemia went uncorrected during pregnancy had lower weight babies, and that the infant's weight was related to the degree of maternal anemia.

No information about maternal nutrition was collected in the survey. This is a serious problem, both because it precludes direct investigation of one

program possibility, prenatal food supplements, and because it is likely to give rise to statistical problems. Specifically, if an important factor like maternal nutrition is omitted, and if it is correlated with some of the included variables, these latter variables will pick up some of the influence actually due to maternal nutrition. There is no way around the policy problem, but we attempted to solve the statistical problem by experimenting with several socio-economic variables that might be related to the mother's nutrition -- the type of neighborhood, father's occupation, the education of the parents, and the water and toilet facilities available to the household.

Equation II.4 shows the results when type of neighborhood and occupation of the father are included. The coefficients of the variables are reasonable and suggest that the two factors may be serving their intended purpose -- to stand in for the influence of maternal nutrition. For example, infants born into well-to-do, or good, neighborhoods (NBR1) have significantly higher birthweights than those in the omitted class ("medium" neighborhoods). Those born into poor, very poor, or rural neighborhoods have lower birthweights, although the coefficients are not statistically significant at the 95 percent level of confidence.

Similarly, the results for father's occupation are reasonable. The infants of fathers in the professions (OCF2), clerical or high-level sales work (OCF3 and OCF4), assorted crafts (OCF12), and miscellaneous (OCF15) have significantly higher birthweights than those born to farm workers, the omitted class. Only infants whose fathers are laborers (OCF11) or police, firefighters, guards, etc. (OCF14) have significantly lower birthweights. Equation II.6 adds type of neighborhood and father's occupation to the basic specification of Equation II.5, with similar results.

Comparing Equations II.1 and II.4 we see that the added factors have little effect on the coefficients of the original variables. If we are willing to assume that maternal nutrition is reasonably well represented by type of neighborhood and father's occupation, this comparison lends support to the validity of our results. Some specific comparisons were noted earlier, but perhaps most interesting is the fact that the coefficients for the prenatal care variables are virtually identical in the two equations.

APPENDIX TO CHAPTER II
Distributions of Observations with Valid Birthweights by Factors Used in the Regressions

(A) All records with valid birthweights and under 2 years of age at the time of interview or death.

(B) All records with valid birthweights who were under two years of age at the time of interview or death, and were not first children.

1. By Birthweight (in hundreds of grams)

	(A)		(B)	
	Deaths	Sample	Deaths	Sample
Under 14	176	0	130	0
15-19	181	2	129	2
20-24	206	17	148	11
25-29	305	56	220	37
30-34	338	130	254	105
35-39	355	112	276	92
40-44	100	51	82	42
45 and over	40	15	32	14
TOTAL	1,701	383	1,271	303

Note: Birthweight is the dependent variable in this chapter and the values were not grouped in the regressions. They are grouped here only for convenience.

2. By Sex of Child

	(A)		(B)	
	Deaths	Sample	Deaths	Sample
Female	776	189	587	148
Male	925	194	684	155
TOTAL	1,701	383	1,271	303

3. By Age of Mother

	(A) Deaths	(A) Sample	(B) Deaths	(B) Sample
Unspecified	6	1	2	1
10-16 years	82	6	19	3
17-20 years	380	48	192	27
21-25 years	458	144	351	102
26-30 years	347	79	306	66
31-35 years	224	70	205	69
36-40 years	140	22	132	22
41-49 years	64	13	64	13
TOTAL	1,701	383	1,271	303

4. By Total Number of Pregnancies (Parity)

	(A) Deaths	(A) Sample	(B) Deaths	(B) Sample
One	422	80	-	-
Two	329	100	329	100
Three-Four	417	126	417	126
Five-Six	247	46	247	46
Seven or more	278	31	278	31
Unspecified	8	0	0	0
TOTAL	1,701	383	1,271	303

5. By Cross-classification of Total Number of Pregnancies (Parity) with Outcome of Last Pregnancy

Parity/Outcome	(A) Deaths	(A) Sample	(B) Not Applicable
One	422	80	
Two/fetal death	33	5	
Two/still living	250	68	
Two/died after birth	36	3	
Three or four/fetal death	50	14	
Three or four/still living	272	97	
Three or four/died after birth	55	10	
Five or six/fetal death	40	6	
Five or six/still living	164	39	
Five or six/died after birth	29	1	
Seven or more/fetal death	44	4	
Seven or more/still living	194	25	
Seven or more/died after birth	31	2	
Unspecified	81	29	
TOTAL	1,701	383	

6. By Outcome of Previous Pregnancy

	(A) Not Applicable	(B) Deaths	(B) Sample
Still Living		880	229
Fetal death		167	29
Died after live birth		151	16
Unspecified		73	29
TOTAL		1,271	303

7. By Length of Pregnancy

	(A) Deaths	(A) Sample	(B) Deaths	(B) Sample
Six months or less	148	1	97	0
Seven months	207	2	154	2
Eight months	271	39	202	29
Nine months	1,062	337	808	267
Unspecified	13	4	10	5
TOTAL	1,701	383	1,271	303

8. By Multiple or Single Birth

	(A) Deaths	(A) Sample	(B) Deaths	(B) Sample
Single	1,523	368	1,132	290
Multiple	170	8	139	6
Unspecified	8	7	0	7
TOTAL	1,701	383	1,271	303

9. By Interval Since Last Birth

	(A) Not Applicable	(B) Deaths	(B) Sample
12 months or less		164	24
13-15 months		195	31
16-18 months		147	40
19-24 months		243	43
25-36 months		264	60
37-48 months		114	38
more than 48 months		134	38
Unspecified		10	29
TOTAL		1,271	303

APPENDIX TO CHAPTER II, continued

10. By Number of Visits for Prenatal Care

	(A)		(B)	
	Deaths	Sample	Deaths	Sample
None	287	41	213	35
1-2 visits	373	66	282	53
3-5 visits	509	101	394	82
6 or more visits	428	161	308	121
Unspecified	104	14	74	12
TOTAL	1,701	383	1,271	303

11. By Place of Birth

	(A)		(B)	
	Deaths	Sample	Deaths	Sample
Hospital	1,552	345	1,148	270
Home, or other	148	37	123	32
Unspecified	1	1	0	1
TOTAL	1,701	383	1,271	303

12. By Attendant at Birth

	(A)		(B)	
	Deaths	Sample	Deaths	Sample
Physician	478	131	342	91
Midwife	1,082	224	817	186
Other, or no one	107	21	86	21
Unspecified	34	7	26	5
TOTAL	1,701	383	1,271	303

13. By Type of Neighborhood

	(A)		(B)	
	Deaths	Sample	Deaths	Sample
Well-to-do, or good	68	42	42	37
Medium	336	117	234	79
Poor	308	88	238	71
Very poor	116	10	92	10
Rural, or no neighborhood	862	124	656	104
Not specified	11	2	9	2
TOTAL	1,701	383	1,271	303

APPENDIX TO CHAPTER II, continued

14. By Occupation of Father

	(A)		(B)	
	Deaths	Sample	Deaths	Sample
Farm Workers	643	99	505	89
Unspecified	170	29	80	20
Professionals	28	25	18	19
Other Clerical Workers	57	16	41	13
Sales: Working Proprietors, insurance and real estate, manufacturer's agents, etc.	13	18	10	12
Shop Assistants and Salesmen	92	29	70	19
Farmers and Farm Managers	50	4	40	2
Drivers, Road Transport	53	16	38	10
Toolmakers, Machinists, Plumbers, etc.	68	26	50	17
Bricklayers, Plasterers, and Other Construction Workers	149	27	123	19
Millers, Bakers, Brewmasters, and Related Workers	36	16	24	13
Laborers	67	5	55	5
Other Crafts	66	29	55	26
Electricians, Carpenters, Painters, etc.	56	12	42	10
Firefighters, Police, Guards, etc.	52	9	43	7
All Other Occupations	101	23	77	22
TOTAL	1,701	383	1,271	303

CHAPTER III. Application of the Nutritional Classification System for Infants Under One Year

Introduction

Numerous studies have found nutritional status to be an important determinant of infant mortality. This is indicated in the literature review reported in the feasibility study (Russell and Burke, 1974, Part One). For example, a Peruvian study found statistically significant differences in infant mortality between an experimental group which received food supplements and a control group which did not (Baertl, et.al., 1970). In a Guatemalan study there was significantly less mortality, as well as slightly lower morbidity and increased physical growth, in a group which received food supplements, as compared to both a control group and a group which had special medical care and improved sanitation facilities (Scrimshaw, 1970). A regression analysis of data from Costa Rica and Mexico found nutrition to be one of two factors that were consistently significant determinants of infant and child mortality (Sloan, 1970). Many studies cite two mechanisms via which nutrition is expected to influence mortality: poor nutrition reduces resistence to infection, and infectious diseases may precipitate malnutrition (Sloan, 1970; Hillman and Hall, 1964; Scrimshaw, et.al., 1966, 1968; Jelliffe, 1968, 1969, 1971).

Tabulations regarding the underlying and associated causes of infant mortality, reported in Patterns of Mortality in Childhood (Puffer and Serrano, 1973), indicate the importance of nutritional status. Table III.1, which contains these tabulations for the South American sites surveyed by PAHO, shows that nutritional deficiency is an associated cause of death in a high proportion of cases in many of the projects, although it is seldom an underlying cause of death. For example, in rural San Juan, 24.9 deaths per 1,000 live births had nutritional deficiency as an associated cause of death, and 3 deaths per 1,000 live births had nutritional deficiency as an underlying cause of death.

Many factors affect nutritional status. These can be divided into two types -- primary and secondary. The primary factor is an inadequate diet. Secondary factors are poor digestion, malabsorption, inefficient utilization and excessive waste.[1/] Together, these factors determine an individual's nutritional status.

The PAHO survey contains a wealth of information about the primary factor, food intake; it contains no information on the secondary factors. In other words, the survey information pertains to primary malnutrition, or malnutrition due to inadequate food intake; it does not pertain to secondary malnutrition, which is common among severely ill children. As part of the feasibility study

---

1/ These are four bodily processes involved in supplying food for synthesis of tissue or for energy. If there is a problem associated with any one of these processes, adequate quantities of nutrients are not accumulated and, therefore, tissue is not synthesized and there is not sufficient energy to carry out normal activity. The first process is digestion -- the process whereby ingested food is converted in the stomach into material suitable for absorption through the walls of the small intestine. The second process is the absorption of digested food through the walls of the small intestine into the circulatory system. The third process is the assimilation of food into the tissue. And, the last process is the process of discharging waste material from food or in the case of malfunction, the discharge of the residue of food, bacteria and intestinal mucus.

TABLE III.1 Infant Mortality Rate[1/] For Nutritional Deficiency and All Causes of Death, by Underlying or Associated Cause, for South American sites

	Underlying Cause		Associated Cause	
	Nutritional Deficiency	All Causes	Nutritional Deficiency	All Causes
Chaco Province, Argentina				
Resistencia	2.0	76.2	25.5	98.8
Rural Departments	1.3	85.0	26.3	83.3
San Juan Province, Argentina				
San Juan (city)	0.7	50.7	8.8	100.9
Suburban Departments	2.8	87.9	20.8	185.4
Rural Departments	3.0	94.5	24.9	179.8
Bolivia				
La Paz	1.0	73.0	19.3	77.3
Viacha	-	123.5	29.4	96.5
Recife, Brazil	1.5	91.2	34.0	167.2
Ribeirao Preto, Brazil				
Ribeirao Preto (city)	0.2	43.0	12.2	83.0
Franca	1.8	71.5	21.6	122.8
Communities	-	50.8	16.7	85.4
Sao Paulo, Brazil	1.0	65.1	17.2	131.6
Chile				
Santiago	0.9	54.9	11.4	90.5
Comunas	1.2	57.9	18.2	89.1
Colombia				
Cali	2.1	54.6	12.4	81.7
Cartagena	2.2	47.8	13.3	88.7
Medellin	1.2	47.6	14.7	83.6

1/ Number of deaths per 1,000 live births.

Source: Puffer and Serrano, Patterns of Mortality in Childhood, 1973, pp. 78-79.

(Russell and Burke, 1974, Part Two), a system was developed for condensing the food intake information about each infant or child into a few summary measures representing the adequacy of its diet. These measures have been used as independent variables in the analysis of infant and child mortality, reported in Chapters V through VII.

The following sections briefly review the nutritional classification system, describe the results of applying it to the survey data for San Juan, Argentina, and report the results of running correlations between the nutritional measures in an attempt to condense them still further. The final section discusses some of the variables that are likely to affect nutritional status.

Nutritional Classification System

The system developed during the feasibility study uses the survey information available on each infant's food intake -- times per day milk was given, types of weaning foods given, other types of foods given,[1/] and times per week each of these other foods was given. Based on this information, the system classifies each infant's diet as "probably adequate", "probably inadequate," "grossly inadequate" or "unclassifiable" in each of the major nutrients -- proteins, calories, iron, vitamin C and other vitamins and minerals. This involves the application of a set of rules which specify what quantity of each type of food is adequate for infants from various age groups -- 0 to 3 months, 4 to 5 months, 6 to 8 months, and 9 to 11 months. In other words, the system summarizes the diet information to show the adequacy of the infant's diet in each of the major nutrients.

1/ Other types of food included juices, cereals, pulses, roots and tubers, leafy vegetables, eggs, poultry and meat.

The development of the system is described in detail in the feasibility report (Russell and Burke, 1974, Part Two). The system was based on:

(1) the recommended daily dietary allowances of the National Academy of Sciences and the National Research Council,

(2) the nutritional value of foods of each type reported in the survey that are common to Latin America,

(3) the ages at which various types of foods should be introduced into an infant's diet, and

(4) the quantities of food an infant or a child of a specified age is capable of consuming at each feeding.

It was impossible to develop rules for classifying the diets of children over one year of age. This is the case because older children have minimum weekly requirements which convert to much more than seven servings per week for some types of food. Yet the survey reports only whether a child receives various types of foods seven or more times per week.

Application of the System

Table III.2 shows, by age group, the unweighted distributions of the nutritional codes resulting from the application of the classification system to infants less than 12 months old in the San Juan, Argentina project. Of 1,946 infants less than 12 months old, 1,338 were classified into one of three diet groups -- probably adequate, probably inadequate, and grossly inadequate. Thirty-one percent, or 608 of the 1,946 infants less than one year old, were "unclassified" because of missing information. Most of these cases were concentrated in the less than 3 months age group and within this group, they are concentrated in under one month age group. Only about 8 percent of the 602 older infants are missing information, whereas about 48 percent of the

Table III.2 Number and Percent of Infants Classifiable and Unclassifiable under the Nutrition System, by Nutrient Class and Age Group[1/]

Nutrient Class	Live				Dead				Total			
	Classifiable		Unclassifiable		Classifiable		Unclassifiable		Classifiable		Unclassifiable	
	No.	Percent	No.	Percent	No.	Percent	No.	Percent	No.	Percent	No.	Percent
0 to 3 months												
proteins	42	52	39	48	744	59	519	41	786	58	558	42
calories	42	52	39	48	744	59	519	41	786	58	558	42
iron	42	52	39	48	744	59	519	41	786	58	558	42
vitamin C	42	52	39	48	744	59	519	41	786	58	558	42
other[2/]	42	52	39	48	744	59	519	41	786	58	558	42
28 days to 5 months[3/]												
proteins	66	67	33	33	604	88	85	12	670	85	118	15
calories	66	67	33	33	604	88	85	12	670	85	118	15
iron	66	67	33	33	604	88	85	12	670	85	118	15
vitamin C	73	74	26	26	621	90	68	10	694	88	94	12
other[2/]	66	67	33	33	604	80	85	12	670	85	118	15
4 to 5 months												
proteins	31	82	7	18	215	93	17	7	246	91	24	9
calories	31	82	7	18	215	93	17	7	246	91	24	9
iron	31	82	7	18	215	93	17	7	246	91	24	9
vitamin C	38	100	0	0	232	100	0	0	270	100	0	0
other[2/]	31	82	7	18	215	93	17	7	246	91	24	9
6 to 8 months												
proteins	37	95	2	5	145	91	15	9	182	92	17	9
calories	37	95	2	5	145	91	15	9	182	92	17	9
iron	37	95	2	5	145	91	15	9	182	92	17	9
vitamin C	39	100	0	0	160	100	0	0	199	100	0	0
other[2/]	39	100	0	0	160	100	0	0	199	100	0	0
9 to 11 months												
proteins	44	96	2	4	80	92	7	8	124	93	9	7
calories	44	96	2	4	80	92	7	8	124	93	9	7
iron	44	96	2	4	80	92	7	8	124	93	9	7
vitamin C	46	100	0	0	87	100	0	0	133	100	0	0
other[2/]	46	100	0	0	87	100	0	0	133	100	0	0

1/ These tabulations are based on unweighted counts.
2/ Other includes calcium, phosphorous, niacin, riboflavin, thiamine, and vitamin A.
3/ This is not an age group for which separate rules were created under the nutrition system; it is shown because it corresponds to one of the age groups analyzed in the regressions.

the live infants less than 3 months old and 41 percent of the dead infants less than 3 months old are missing information (Table III.2). For many of the infants the reason for the lack of information is clear. Fifty-five percent of the infants who were unclassifiable were in the hospital continuously from birth and later died; nearly all of these were less than 28 days old (Table III.3).

In most cases an infant could not be classified because there was no information about its milk consumption. Examination of the breastfeeding history of unclassifiable infants (Tables III.4 and III.5)shows that most of them were never breastfed and 80 percent of the infants who were never breastfed died within 48 hours after birth. Another 14 percent were in the hospital at the time to begin breastfeeding. Also, 43 percent of the infants who were breastfed but who were nevertheless missing information on milk intake were less than 28 days old. With all this information in mind, no nutritional measures were included in regressions involving infants less than 28 days. Also, this information suggests that attempts to use breastfeeding independently might be misleading because the decision about breastfeeding was not the parents' for many cases, but rather, one determined by whether or not the child was in the hospital.

Now let us turn to considering the classified infants. Examination of Tables III.6 and III.7 indicates that the system does discriminate between infant diets in the crude sense that it does not lump all of the infants into the same diet groups; in each group some infants' diets are classified as probably adequate, while others are classified as probably or grossly inadequate. A larger percentage of dead infants 6 to 11 months old than of live infants 6 to 11 months old have grossly inadequate diets, and vice-versa for probably adequate diets. This is not, however, true for infants less than 6 months old;

TABLE III.3 Distribution of All Unclassifiable Infants by Care of Infant and by Age Group

Care/Age Group	1 to 27 Days	28 Days to 5 Months	6 to 11 Months
	Live		
Hospital	-	-	-
Other	13	33	4
Not Specified	-	-	-
TOTAL	13	33	4
	Dead		
Hospital	335	1	-
Other	111	82	22
Not Specified	5	2	-
TOTAL	451	85	22
	Total		
Hospital	335	1	-
Other	124	115	26
Not Specified	5	2	-
TOTAL	464	118	26

TABLE III.4 Distribution of Unclassifiable Infants by History of Breastfeeding, by Age Group

History of Breastfeeding	1 to 27 Days	28 Days to 5 Months	6 to 11 Months
	Live		
Not Specified	1	-	-
Breastfed	12	33	4
Not Breastfed	-	-	-
TOTAL	13	33	4
	Dead		
Not Specified	2	1	1
Breastfed	91	83	21
Not Breastfed	358	1	-
TOTAL	451	85	22
	Total		
Not Specified	3	1	1
Breastfed	103	116	25
Not Breastfed	358	1	-
TOTAL	464	118	26

TABLE III.5 Distribution of Infants [illegible] Who Were Not Breastfed, by Reason for Not Breastfeeding, by Age Group

Reason for Never Breastfeeding/Age Group	1 to 27 Days	28 Days to 5 Months	6 to 11 Months
		Live	
Child died within 48 hours after birth	-	-	-
Child in hospital at time to begin breastfeeding	-	-	-
Other	-	-	-
Not specified	-	-	-
TOTAL	-	-	-
		Dead	
Child died within 48 hours after birth	294	-	-
Child in hospital at time to begin breastfeeding	50	-	-
Other	13	1	-
Not specified	1	-	-
TOTAL	358	1	-
		Total	
Child died within 48 hours after birth	294	-	-
Child in hospital at time to begin breastfeeding	50	-	-
Other	13	1	-
Not specified	1	-	-
TOTAL	358	1	-

approximately the same share of dead and live infants less than 6 months old have either probably adequate or grossly inadequate diets. This may indicate that the diets of younger infants tend to be more adequate, which is reasonable because most of their nutritional requirements can be met by adequate quantities of milk alone. Or it may indicate that the information on which the system is based is poor, or the system itself is not properly designed.

Also, examination of Tables III.6 and III.7 shows the distinction between probably adequate and probably inadequate appears to be peculiar. In several instances few, or no, infants fall into one of the two categories. For example, virtually none of the diets of infants 6-11 months old were classified as probably inadequate in terms of calories and proteins. For four to five month olds, less than one percent of the diets were classified as probably inadequate in terms of calories; a significant percentage were classified as probably inadequate in terms of protein, but less than one percent were classified as probably adequate in this nutrient. This classification pattern does not exist for vitamins and minerals.

These tabulations tell us several things about the probable nutritional status of infants between 28 days and 11 months of age in the San Juan survey --

(1) About 50 percent of the infants -- about 60 percent of the live infants and 50 percent of the dead infants -- had diets that were probably adequate in protein. About 30 percent had diets grossly inadequate in protein.

(2) About 50 percent of the infants -- about 65 percent of the live infants and about 50 percent of the dead infants -- had diets that were probably adequate in calories. About 35 percent had diets grossly inadequate in calories and may have verged on marasmus, which is a severe calorie and protein deficiency state.

Table III.6 Number of Classifiable Infants by Nutrition Code, Diet Group and Age Group[1/]

Nutrient/Diet Group	Probably Adequate			Probably Inadequate			Grossly Inadequate		
	Live	Dead	Total	Live	Dead	Total	Live	Dead	Total
0 to 3 months									
proteins	35	592	627	4	40	44	3	112	115
calories	26	392	418	9	200	209	7	152	159
iron	34	511	545	5	121	126	3	112	115
vitamin C	34	511	545	5	121	126	3	112	115
other[2/]	34	511	545	5	121	126	3	112	115
28 days to 5 months[3/]									
proteins	32	330	362	24	171	195	10	103	113
calories	46	332	378	8	144	152	12	128	140
iron	31	259	290	5	105	110	30	240	270
vitamin C	38	283	321	9	128	137	26	210	236
other[2/]	53	399	452	11	161	172	2	44	46
4 to 5 months									
proteins	1	1	2	22	140	162	8	74	82
calories	23	143	166	0	3	3	8	69	77
iron	1	1	2	2	3	5	28	211	239
vitamin C	8	25	33	6	26	32	24	181	205
other[2/]	23	141	164	8	59	67	0	15	15
6 to 8 months									
proteins	22	34	56	0	0	0	15	111	126
calories	19	31	50	0	0	0	18	114	132
iron	18	25	43	1	1	2	18	119	137
vitamin C	18	25	43	9	39	48	12	96	108
other[2/]	16	8	24	23	135	158	0	17	17
9 to 11 months									
proteins	31	42	73	0	0	0	13	38	51
calories	29	33	62	0	0	0	15	47	62
iron	27	33	60	1	0	1	16	47	63
vitamin C	22	24	46	11	33	44	13	30	43
other[2/]	23	19	42	23	64	87	0	4	4

1/ These tabulations are based on unweighted counts.
2/ Other includes calcium, phosphorous, niacin, riboflavin, thiamine, and vitamin A.
3/ This is not an age group for which separate rules were created under the nutrition system; it is shown because it corresponds to one of the age groups analyzed in the regressions.

Table III.7 Percent of Classifiable Infants by Nutritional Codes, Diet Group and Age Group[1/]

Nutrient/Diet Group	Probably Adequate			Probably Inadequate			Grossly Inadequate		
	Live	Dead	Total	Live	Dead	Total	Live	Dead	Total
0 to 3 months									
proteins	83	80	80	10	5	6	7	15	15
calories	62	53	53	21	27	27	17	20	20
iron	81	69	69	12	16	16	7	15	15
vitamin C	81	69	69	12	16	16	7	15	15
other[2/]	81	69	69	12	16	16	7	15	15
28 days to 5 months[3/]									
proteins	43	55	54	36	28	29	15	17	17
calories	70	55	56	12	24	23	18	21	21
iron	47	43	43	8	17	16	46	40	40
vitamin C	52	46	46	12	21	20	36	34	34
other[2/]	80	66	68	17	27	26	3	7	7
4 to 5 months									
proteins	3	1	1	71	65	66	26	34	33
calories	74	67	68	0	1	1	26	32	31
iron	3	1	1	7	1	2	90	98	97
vitamin C	21	11	12	16	11	12	63	78	76
other[2/]	74	66	67	26	27	27	0	7	6
6 to 8 months									
proteins	59	23	31	0	0	0	41	77	69
calories	51	21	28	0	0	0	49	79	73
iron	49	17	24	3	1	1	49	82	75
vitamin C	46	16	22	23	24	24	31	60	54
other[2/]	41	5	12	59	84	79	0	11	9
9 to 11 months									
proteins	71	53	59	0	0	0	30	48	41
calories	66	41	50	0	0	0	34	59	50
iron	61	41	48	2	0	1	36	59	51
vitamin C	48	28	35	24	38	33	28	35	32
other[2/]	50	22	32	50	74	65	0	5	3

1/ These tabulations are based on unweighted counts. Also, numbers may not add to 100 because of rounding.
2/ Other includes calcium, phosphorous, niacin, riboflavin, thiamine, and vitamin A.
3/ This is not an age group for which separate rules were created under the nutrition system; it is shown because it corresponds to one of the age groups analyzed in the regressions.

(3) About 65 percent of the infants had diets probably adequate in vitamins and minerals. About 10 percent had diets grossly inadequate in vitamins and minerals.

Correlations Between Nutrition Measures

To explore the possibility that the classifications of one or two of the nutrients could adequately represent the infant's diet, correlations between the nutritional measures were computed. These correlations were computed for two basic age groups on which regressions were run -- 28 days to 5 months and 6 to 11 months. They used variables which represent the three classes of dietary adequacy plus a fourth omitted class, for each of the major nutrients -- proteins, calories, iron, vitamin C and other vitamins and minerals. The three classes are: "probably inadequate", "grossly inadequate", and "unclassifiable"; the omitted class is "probably adequate." So, for example, the protein variables were defined as follows:

PRO1=1,if protein intake is probably inadequate, and PRO1=0, otherwise.

PRO2=1,if protein intake is grossly inadequate, and PRO2=0, otherwise.

PRO3=1, if case is unclassifiable, and PRO3=0, otherwise.

Omitted class: PRO1=0, PRO2=0 and PRO3=0, if protein intake is probably adequate.

Similarly, other variables were defined for the remaining nutrients -- calories (CAL), iron (IRN), vitamin C (VTC), and other vitamins and minerals (OVM).[1/]

Examination of these correlations, which are presented in Tables III.8 and III.9, together with the distributions of the nutrition measures in Tables

---

1/ Appendix B contains definitions of these variables.

III.6 and III.7, indicated that it was desirable to condense the information on the adequacy of food intake still further by omitting and combining some of the measures. Specifically, this examination indicated the following steps should be taken:

28 days to 5 months

Diets for infants 28 days to 5 months can be adequately summarized by the variables for protein and calories, and a selected few of the remaining measures. More specifically,

(1) Omit IRN3, VTC3 and OVM3 (the unclassifiable cases) since PRO3 and CAL3 are highly correlated, sometimes perfectly correlated, with all three variables as well as each other. Each of these variables represents the same, unclassifiable, infants, and one of them can do the job.

(2) Combine IRN1 (probably inadequate cases) with the omitted IRN class (probably adequate cases). There are too few cases classified as "probably inadequate" in terms of iron to justify a separate class.

(3) Combine OVM2 (grossly inadequate cases) with OVM1 (probably inadequate cases) because there are too few cases classified as "grossly inadequate" in terms of other vitamins and minerals to justify two separate classes.

6 to 11 months

Diets for infants 6 to 11 months can be adequately summarized by the protein, calorie, and other vitamin and mineral variables. More specifically,

(1) Omit the following variables since no cases fell in the categories they represent: VTC3 and OVM3 (unclassifiable cases); PRO1 and CAL1 (probably inadequate cases); and OVM2 (grossly inadequate cases).

(2) Omit IRN2 and VTC2 since both PRO2 and CAL2 are highly correlated with IRN2 and VTC2. Omit PRO1 and CAL1 since no cases fall in these classes.

(3) Combine PRO3 with PRO2, and CAL3 with CAL2, because there are too few "unclassifiable" cases to justify separate classes.

TABLE III.8 CORRELATION MATRIX FOR NUTRITIONAL VARIABLES FOR INFANTS 28 DAYS TO 5 MONTHS, SAN JUAN, ARGENTINA 1/,2/

	PRO1	PRO2	PRO3	CAL1	CAL2	CAL3	IRN1	IRN2	IRN3	VTC1	VTC2	VTC3	OVM1	OVM2	OVM3
PRO1 with -	1.00000	-0.19398	-0.38545	-0.17928	-0.05972	-0.38545	0.26552	0.64430	-0.38545	0.21819	0.36156	-0.39658	-0.05138	-0.06624	-0.38545
PRO2 with -		1.00000	-0.23296	-0.10308	0.89318	-0.23296	-0.08348	0.51391	-0.23296	0.10057	0.35158	-0.19738	0.68579	0.45511	-0.23296
PRO3 with -			1.00000	-0.21541	-0.25884	1.00000	-0.16938	-0.45258	1.00000	-0.15539	-0.13047	0.84727	-0.25240	-0.10602	1.00000
CAL1 with -				1.00000	-0.12039	-0.21541	0.17698	-0.20700	-0.21541	0.09995	-0.16963	-0.15251	0.07206	-0.04931	-0.21541
CAL2 with -					1.00000	-0.25884	0.19689	0.42635	-0.25884	0.29308	0.28749	-0.21730	0.50679	0.40961	-0.25884
CAL3 with -						1.00000	-0.16938	-0.45258	1.00000	-0.15539	-0.13047	0.84727	-0.25240	-0.10602	1.00000
IRN1 with -							1.00000	-0.16801	-0.16938	0.61916	-0.15027	-0.14393	0.42745	-0.03589	-0.16938
IRN2 with -								1.00000	-0.45258	0.06929	0.62356	-0.36346	0.29266	0.23426	-0.45258
IRN3 with -									1.00000	-0.15539	-0.13047	0.84727	-0.25240	-0.10602	1.00000
VTC1 with -										1.00000	-0.19899	-0.19059	0.46476	-0.04637	-0.15539
VTC2 with -											1.00000	-0.34711	0.13795	0.25533	-0.13047
VTC3 with -												1.00000	-0.21385	-0.08783	0.84727
OVM1 with -													1.00000	-0.05778	-0.25240
OVM2 with -														1.00000	-0.10602
OVM3 with -															1.00000

1/ See Appendix B for variable definitions.
2/ These correlations were run on weighted data.

TABLE III.9 CORRELATION MATRIX FOR NUTRITIONAL VARIABLES FOR INFANTS 6 TO 11 MONTHS, SAN JUAN, ARGENTINA 1/, 2/

PRO1 with -														
PRO1	PRO2	PRO3	CAL1	CAL2	CAL3	IRN1	IRN2	IRN3	VTC1	VTC2	VTC3	OVM1	OVM2	OVM3
*	*	*	*	*	*	*	*	*	*	*	*	*	*	*
PRO2 with -														
PRO2	PRO3	CAL1	CAL2	CAL3	IRN1	IRN2	IRN3	VTC1	VTC2	VTC3	OVM1	OVM2	OVM3	
1.00000	-0.16000	*	0.88119	-0.16000	-0.10931	0.85949	-0.16000	-0.27002	0.85964	*	0.84326	0.00436	*	
PRO3 with -														
PRO3	CAL1	CAL2	CAL3	IRN1	IRN2	IRN3	VTC1	VTC2	VTC3	OVM1	OVM2	OVM3		
1.000	*	-0.18157	1.00000	-0.03453	-0.18615	1.00000	0.38396	-0.13465	*	-0.12441	0.12291	*		
CAL1 with -														
CAL1	CAL2	CAL3	IRN1	IRN2	IRN3	VTC1	VTC2	VTC3	OVM1	OVM2	OVM3			
*	*	*	*	*	*	*	*	*	*	*	*			
CAL2 with -														
CAL2	CAL3	IRN1	IRN2	IRN3	VTC1	VTC2	VTC3	OVM1	OVM2	OVM3				
1.00000	-0.18157	-0.12465	0.97537	-0.18157	-0.21836	0.75247	*	*	-0.00131	*				
CAL3 with -														
CAL3	IRN1	IRN2	IRN3	VTC1	VTC2	VTC3	OVM1	OVM2	OVM3					
1.00000	-0.03453	-0.18615	1.00000	0.38396	-0.13465	*	-0.02441	0.12291	*					
IRN1 with -														
IRN1	IRN2	IRN3	VTC1	VTC2	VTC3	OVM1	OVM2	OVM3						
1.00000	-0.12718	-0.03453	0.09435	-0.10042	*	-0.16768	-0.00707	*						
IRN2 with -														
IRN2	IRN3	VTC1	VTC2	VTC3	OVM1	OVM2	OVM3							
1.00000	-0.18615	-0.22590	0.73252	*	0.58019	-0.00242	*							
IRN3 with -														
IRN3	VTC1	VTC2	VTC3	OVM1	OVM2	OVM3								
1.00000	0.38396	-0.13465	*	-0.02441	0.12291	*								
VTC1 with -														
VTC1	VTC2	VTC3	OVM1	OVM2	OVM3									
1.00000	-0.36411	*	0.06550	-0.02565	*									
VTC2 with -														
VTC2	VTC3	OVM1	OVM2	OVM3										
1.00000	*	0.53452	0.06566	*										
VTC3 with -														
VTC3	OVM1	OVM2	OVM3											
*	*	*	*											
OVM1 with -														
OVM1	OVM2	OVM3												
1.00000	-0.05070	*												
OVM2 with -														
OVM2	OVM3													
1.00000	*													
OVM3 with -														
OVM3														
*														

* Means that one or both of the variables equalled zero, since no cases fell in the specified class, and, therefore, the correlations were equal to zero.

1/ See Appendix B for variable definitions.

2/ These correlations were run on weighted data.

Determinants of Nutritional Status

From a policy point of view, it is useful to know not only the nutritional status of a population, but the factors that produce that nutritional status. The variables that are likely to affect nutritional status can be separated into those that are expected to have their major influence on food intake, or primary malnutrition, and those that are related to secondary malnutrition or poor digestion, malabsorption, inefficient utilization and excessive losses of nutrients.

Ultimately, of course, primary and secondary malnutrition are synergistic. There is firm agreement that inadequate food intake reduces a child's resistence to infection, and that infectious diseases are associated with poor appetite, diarrhea, and vomiting which reduces food intake as well as adding to nutrient demands for immunity, for repair of damaged tissue, and for supplying the nutrient requirements of infecting agents in some cases. A recent Guatemalan study, among others, has documented this synergism (Scrimshaw, et.al., 1966). The investigators found that 3.4 percent of measles patients under 5 years old in a group receiving food supplements died, compared to 6.8 percent of measles patients under 5 years old in a group that did not receive food supplements. They also found that measles was the precipitating cause of protein malnutrition in 2 of 41 measles patients in the group without food supplements, compared to none of 55 in the group receiving supplements.

Food intake is affected by family income. The poorest fed children are usually found in families with the lowest incomes in both developed and developing countries. In poor families, a larger proportion of food expenditures goes for cereals and, in non-poor families, a larger proportion for milk products and protein foods (Berg, 1973).

Another factor which probably affects nutritional intake is family size. In other words, family size and family income interact because they reduce a family's per capita income (Cook, et.al., 1973). In a regression analysis of data on English children the authors found that both family size and social class were associated with highly significant differences in average daily caloric intake. In an Indian study, Rao and Gopalan (1969) found that while children with 3 or more older siblings constitute 34 percent of the population, they account for 61 percent of all cases of protein-calorie malnutrition in clinics.

Ignorance is purported, in most of the available literature, to be a major cause of inadequate food intake. Jelliffe (1968) reports that, in 1952-54 in Jamaica, one in four of the mothers of children with kwashiorkor would have been able to obtain small additional amounts of milk but continued feeding with corn gruel, sugar and very little condensed milk because they believed this was suitable food. Other examples of harmful customs cited by Jelliffe include reserving animal protein for male elders, fasting, use of purges to treat diarrheal diseases, and prejudice against giving eggs to infants.

Likewise ignorance,coupled with environmental factors such as quality of water supply and accessibility of health services, is cited in the literature as a major cause of secondary malnutrition. For instance, there is lack of knowledge about how to prepare foods; they are often prepared with contaminated water. This is particularly a problem with artificial foods and milk formulas for bottles. Monckeberg (1970) reports, in a study of urban Chile, that 80 percent of the bottles examined had a high bacterial count. Mata and Wyatt (1971) found that children who are breastfed have a lower incidence of diarrhea.

In conclusion, it appears that an econometric investigation of the determinants of infant diets should examine such factors as father's occupation, family size, and mother's and father's education as independent variables. Father's occupation would serve as a proxy for family income, family size would obviously represent itself and parental education might serve as a measure for ignorance.

Likewise it would be desirable to examine factors likely to affect secondary malnutrition -- education of mother, source of water, type of neighborhood, type of breastfeeding, toilet facilities and medical care as proxies for environmental factors. In this case, since food intake is not a measure of secondary malnutrition, it cannot be used as the dependent variable. Some direct measure of nutritional status is needed. Unfortunately, in the PAHO survey, nutritional status at time of interview, a measure of both primary and secondary malnutrition available for living children, was coded using nutritional status at birth in many cases, and so appears to be of little value.

CHAPTER IV. Determinants of Neonatal Mortality

Introduction

An infant faces the highest risk of death in the period immediately following its birth. Chart IV.1 shows, by month of age, the number of infant deaths in San Juan during the two-year survey period. Of the population of 1,900 deaths, 577 occurred during the first week of life and another 329 during the next three weeks. Thus almost half of the deaths during the first year were concentrated in the neonatal period.

This distribution is typical of the other project sites included in the Inter-American Investigation of Mortality in Childhood. In the seven projects shown in Table IV.1, between 40 and 50 percent of the infant deaths occurred during the first four weeks after birth, and a disproportionate number of these occurred during the first week.

Clearly an understanding of neonatal deaths is important to any investigation of the causes of infant mortality. The literature indicates that death during the neonatal period -- and especially the perinatal period, or first week -- depends heavily on the infant's condition at birth. Social and economic factors are believed to play only a minor direct role until the child is older. These considerations suggest that neonatal deaths should be analyzed separately from other infant deaths.

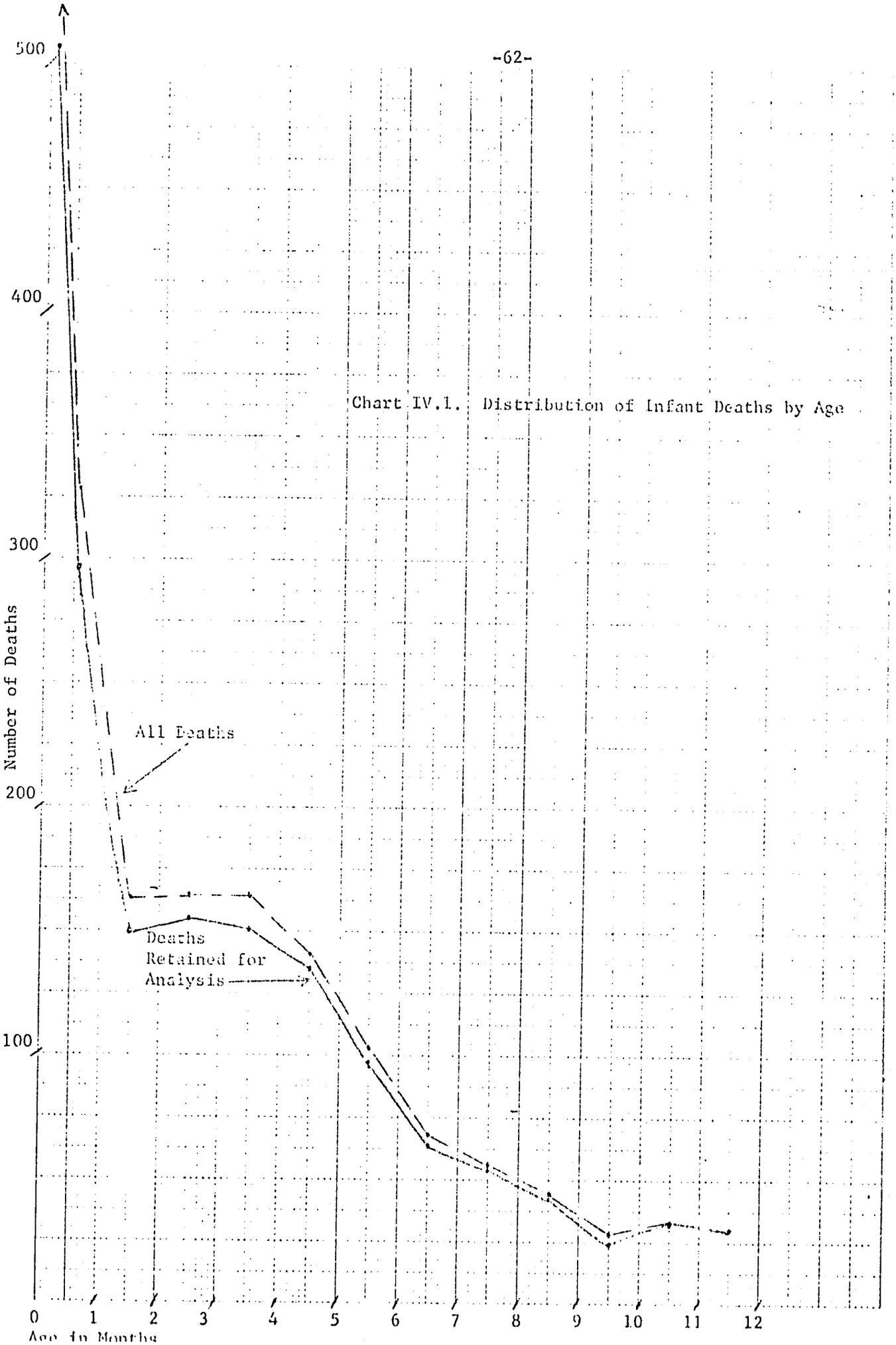


Chart IV.1. Distribution of Infant Deaths by Age

Table IV.1. Infant Mortality, and Number of Infant Deaths by Age, for Selected South American Survey Sites

	Neonatal Deaths		Postneonatal Deaths	Infant Mortality Rate[a]
	Under 7 days	7-27 days		
ARGENTINA: Chaco Province	399	169	842	80.1
ARGENTINA: San Juan Province	577	329	994	81.3
(Retained for analysis in this study)	(509)	(297)	(936)	(—)
BRAZIL: Recife	769	304	1,700	91.2
BRAZIL: Sao Paulo	1,183	775	1,830	65.1
COLOMBIA: Cali	429	107	617	54.6
COLOMBIA: Cartagena	308	93	455	47.8
COLOMBIA: Medellin	315	68	541	47.6

a. Defined as the number of deaths under one year per 1,000 live births.

Source: Puffer and Serrano, 1973, p. 61.

Most of the sets of observations used in this study for the analysis of mortality were defined in the spirit of cohort death rates -- that is, rates that compare deaths in a given age group with the number of living persons in the same age group. But this approach was not feasible for the neonatal deaths because only twenty living children from the same age group were included in the survey. Instead the standard definition of an infant mortality rate was used to suggest the subset to be analyzed. By this definition infant deaths during a given period are compared with all live births during the same period, regardless of whether the children subsequently died.

Employing this analogy we chose the set of all infants under one year, living or dead, for the analysis. The dependent variable, a binary, then took on the value 1 (one) for a neonatal death, and the value 0 (zero) for all other records -- those of deaths between 28 days and one year as well as those of all living children under one year. Thus the comparison group consists of all infants who survived the neonatal period, whether or not they subsequently died. It is made up of 936 infant deaths that occurred after 27 days and 204 living children under one year of age.[1/, 2/]

The literature review in the feasibility report (Russell and Burke, 1974, Part One) surveyed a number of factors that are believed to influence infant mortality. Many authors did not distinguish between the neonatal and post-neonatal periods in their discussions, but such distinctions as were made indicated that the infant's condition at birth, and the events surrounding birth, were of primary importance. These factors include: the prenatal history of the

---

1/ The twenty living children under one month were included in this group on the assumption that most, if not all, of them would survive the neonatal period.

2/ The weights applied to these observations, and to the 806 neonatal deaths, are set out in Appendix C.

child, as summarized in its birthweight and gestation period (length of pregnancy); the sex of the child; and the type of medical care available during the birth. The child's birth order and the spacing between births have also been associated with differences in mortality, although it seems likely that the relationships are primarily due to their effects on birthweight (see Chapter II). Finally, environmental factors, although expected to be much less important than the child's condition at birth, may begin to have some effect during the neonatal period.

A series of regressions were specified to test the importance of these factors and measure their net effects. The factors included in the regressions were, of course, those that are present during the first days of life; no use was made of information which could be relevant only to the older children in the observation set.

The regression results are presented in Equations IV.1 through IV.3. Equation IV.1, the basic equation, contains variables reflecting the characteristics of the child at birth, and the type of medical care received during the delivery. Equation IV.2 varies the basic equation by replacing the variables for parity with variables representing the interval since last birth. Equation IV.3 tests a variety of social and environmental factors in conjunction with the basic specification of Equation IV.1.

Just as for the birthweight equations in Chapter II, the variables shown are in the binary form and each variable represents one of the mutually exclusive classes, or intervals, into which the factor has been divided. For each factor one of the classes has been emitted and is not represented by an explicitly defined variable. Because of this the regression coefficients show

the effect of membership in a specified class, as compared to membership in the omitted class, and the t-scores refer to the statistical significance of these differences.[1/]

Detailed definitions of the classes represented by the variables shown in the equations are given in Appendix B. For the reader's information, tables showing the distributions of the observations across the classes defined for each factor are given in the appendix to this chapter.

The interpretation of the dependent variable in these equations deserves special mention. As noted earlier, the dependent variable takes on the value 1 (one) for neonatal deaths and 0 (zero) for all other infants under one year. When we compute the value of the dependent variable from a particular equation for given characteristics, specified in terms of values of the independent variables, the resulting estimate is interpreted as the <u>probability of death</u> during the neonatal period for children with those characteristics. (A probability is, of course, limited to the range of values between 0 (zero) and 1 (one).) Similarly, the coefficient for a particular variable shows the net contribution of membership in that class to the probability of death. Thus, for example, a coefficient of +0.15 for a particular variable means that children with the characteristic represented by that variable face a probability of death that is .15 higher than the probability that applies to children in the omitted class for the factor.[2/, 3/]

1/ For fuller discussions of the definition and interpretation of binary variables, see Appendix B and Chapter II (the section entitled "The Regressions: Specification and General Discussion").

2/ The correspondence between conventional death rates and the probability of death is simple and direct -- for example, 40 neonatal deaths per 1,000 live births is equivalent to a probability of death of .04.

3/ As before, factor refers to a single causal influence, such as birthweight, while the term variables refers to the binary or binaries used to represent the classes of the factor.

We turn now to examine the regression results. As the coefficients of determination (R^2's) indicate, the factors tested explain a large portion -- about 36 percent -- of the variation in the neonatal survival experience of these infants. And in general, the equations support the notion that the infant's condition at birth is the most important predictor of its survival: birthweight and the length of pregnancy are far and away the most significant elements in the equations.

The following sections discuss the results in detail.

Birthweight

A great many studies attest to the importance of birthweight to an infant's survival (Erhardt, et.al., 1970; Kessner, et.al., 1973; Mata, et.al., 1972; Puffer and Serrano, 1973). On the basis of the evidence presented in these studies, we hypothesized that heavier babies have a higher probability of surviving the neonatal period and we entered birthweight in the regressions. We specified a detailed set of weight classes in order to permit careful investigation of the low and high ends of the range. The nine classes are, in hundreds of grams: 14 or less (BWT1); 15-19 (BWT2); 20-24 (BWT3); 25-29 (BWT4); 30-34 (BWT5); 35-39 (the omitted class); 40-44 (BWT6); 45 and over (BWT7); and unspecified (BWT8).

Birthweight is highly significant, and the coefficients are quite stable, in all of the specifications tested. The coefficients show that infants below 2500 grams (5.5 pounds), the conventional dividing line between low and normal weight, face much higher probabilities of death than do heavier infants. For babies 1400 grams (3 pounds) or less at birth, the probability is very high indeed: the probability of death is .018 for infants 3500 to 3900 grams, the omitted group -- it is .687 higher, or .705 for infants 1400 grams or less (Equation IV.1). The probability is less but still very high for infants between 1500 and 1900 grams -- .423, or .405 higher than for infants in the omitted

EQUATION IV. 1: Basic Equation

Dependent Variable: Neonatal Death = 1

Set of Observations: Records of all infants under 12 months of age.

Name of Factor	Name of Variables	Coefficient	t-score	F for Factor (See Note)
Birthweight	BWT1	0.687	14.17[b]	43.85[b]
	BWT2	0.405	11.86[b]	
	BWT3	0.076	4.36[b]	
	BWT4	0.024	1.79	
	BWT5	-0.004	0.35	
	BWT6	-0.012	0.92	
	BWT7	-0.047	2.31[a]	
	BWT8	-0.022	1.20	
Length of Pregnancy	PLG1	0.240	5.98[b]	45.18[b]
	PLG2	0.355	10.91[b]	
	PLG3	0.005	0.34	
Place of Birth	BPLC1	-0.010	0.69	0.005
Who Attended Birth	ATND1	-0.004	0.46	1.41
	ATND2	0.028	1.42	
Sex of Child	SEX	-0.030	3.85[b]	14.82[b]
Number of Pregnancies (Parity)	TPRG1	-0.004	0.42	4.04[b]
	TPRG2	-0.007	0.71	
	TPRG3	0.001	0.04	
	TPRG4	0.049	3.32[b]	
Constant Term		0.032		

R^2 for the equation: 0.356
F for the equation: 56.17
a. significant at the 95 percent level of confidence.
b. significant at the 99 percent level of confidence.

NOTE: Each F-value indicates whether the factor as a whole is statistically significant. When a factor is represented by a single variable the F is simply the square of the t-score. When a factor is represented by more than one variable it turns out that, in general, the factor is significant whenever at least one of the variables has a significant t-score. Limitations on computer time made it impossible to calculate these F's for equations other than the basic equation in each chapter but the reader can tell when a factor is significant by examining the t-scores.

EQUATION IV. 2: Basic Equation without Number of Pregnancies and With Interval Since Last Birth

Dependent Variable: Neonatal Death = 1

Set of Observations: Records of all infants under 12 months of age.

Name of Factor	Names of Variables	Coefficient	t-score
Birthweight	BWT1	0.693	14.36[b]
	BWT2	0.415	12.10[b]
	BWT3	0.080	4.57[b]
	BWT4	0.021	1.55
	BWT5	0.001	0.08
	BWT6	0.004	0.29
	BWT7	-0.046	2.20[a]
	BWT8	-0.004	0.21
Length of Pregnancy	PLG1	0.237	5.96[b]
	PLG2	0.367	11.28[b]
	PLG3	0.007	0.55
Place of Birth	BPLC1	0.001	0.07
Who Attended Birth	ATND1	-0.001	0.11
	ATND2	0.029	1.46
Sex of Child	SEX	-0.026	3.38[b]
Interval Since Last Birth	SPCX	-0.0004	0.05
	SPC1	-0.007	0.44
	SPC2	-0.027	1.67
	SPC3	-0.008	0.50
	SPC4	0.008	0.54
	SPC5	-0.023	1.51
	SPC6	-0.052	3.17[b]
Constant Term	-	0.039	-

R^2 for the equation: 0.357
F for the equation: 48.51
a. significant at the 95 percent level of confidence.
b. significant at the 99 percent level of confidence.

EQUATION IV.3. Basic Equation with Environmental Factors

Dependent Variable: Neonatal Death = 1

Set of Observations: Records of all infants under 12 months of age.

Name of Factor	Names of Variables	Coefficient	t-score
Birthweight	BWT1	0.666	13.63[b]
	BWT2	0.398	11.51[b]
	BWT3	0.057	3.03[b]
	BWT4	0.023	1.68
	BWT5	-0.006	0.53
	BWT6	-0.008	0.64
	BWT7	-0.042	2.00[a]
	BWT8	-0.034	1.78
Length of Pregnancy	PLG1	0.250	6.13[b]
	PLG2	0.360	10.85[b]
	PLG3	0.006	0.43
Place of Birth	BPLC1	-0.016	1.02
Who Attended Birth	ATND1	-0.006	0.61
	ATND2	0.028	1.34
Sex of Child	SEX	-0.025	3.09[b]
Number of Pregnancies (Parity)	TPRG1	0.0004	0.03
	TPRG2	0.002	0.18
	TPRG3	0.007	0.42
	TPRG4	0.043	2.32[a]
Type of Neighborhood	NBR1	0.008	0.57
	NBR2	0.002	0.20
	NBR3	0.054	2.08[a]
	NBR4	0.020	2.00[a]
Number of Persons in Household	FAM1	0.007	0.46
	FAM2	-0.005	0.53
	FAM3	0.022	1.60
	FAM4	0.015	0.91
Number of Rooms in Dwelling	RMS1	0.012	0.96
	RMS2	0.012	1.11
	RMS3	0.011	1.08
Marital Status of the Mother	MARM1	0.054	2.71[b]
	MARM2	-0.008	0.73
Constant Term	-	0.010	-

R^2 for the equation: 0.364
F for the equation: 34.24
a. significant at the 95 percent level of confidence.
b. significant at the 99 percent level of confidence.

class. It declines still further for infants between 2000 and 2400 grams -- to about .094 (.018 plus the coefficient for BWT3, .076). The rate of decline quickly falls off after that,and the relationship becomes almost flat, until it finishes up with a final dip -- to virtually certain survival -- for infants 4500 grams (almost 10 pounds) and over.[1/]

Chart IV.2 shows the relationship graphically. As birthweight increases, the steps -- which show the average probability of death over the interval -- decline. The steps are high at first, but the decline is rapid. After 3000 grams (6.5 pounds) the relationship traced by the steps becomes virtually flat, levelling out at very low probabilities of death.

The relationship is shown somewhat differently by the series of connected line segments also drawn on Chart IV.2 This curve is a piecewise linear approximation to the relationship. It is based on a regression using a set of variables so defined that, rather than fitting a step function as the binaries do, they permit the line segment over an interval to take on a slope reflecting the change in the probability of death as birthweight increases.[2/] Four intervals were defined for this variable form -- 1900 grams or less, 2000 to 2900 grams, 3000 to 4000 grams, and over 4000 grams.

1/ The differentials between birthweight groups are constants, given by the coefficients. But the absolute level of the probability depends on the other characteristics of the infant. For example, we could have chosen to compare infants with the characteristics represented by the omitted classes for factors other than birthweight. In this case the probability for children who also fall in the omitted birthweight group is .032, the constant term of the equation (Equation IV.1). The probability for children with all the other omitted group characteristics, but a birthweight of 1400 grams or less, would be .719 (.687 plus .032), and so on.

2/ The definition and interpretation of these variables are described in Appendix B. They are also described in a paper by Watts (1963).

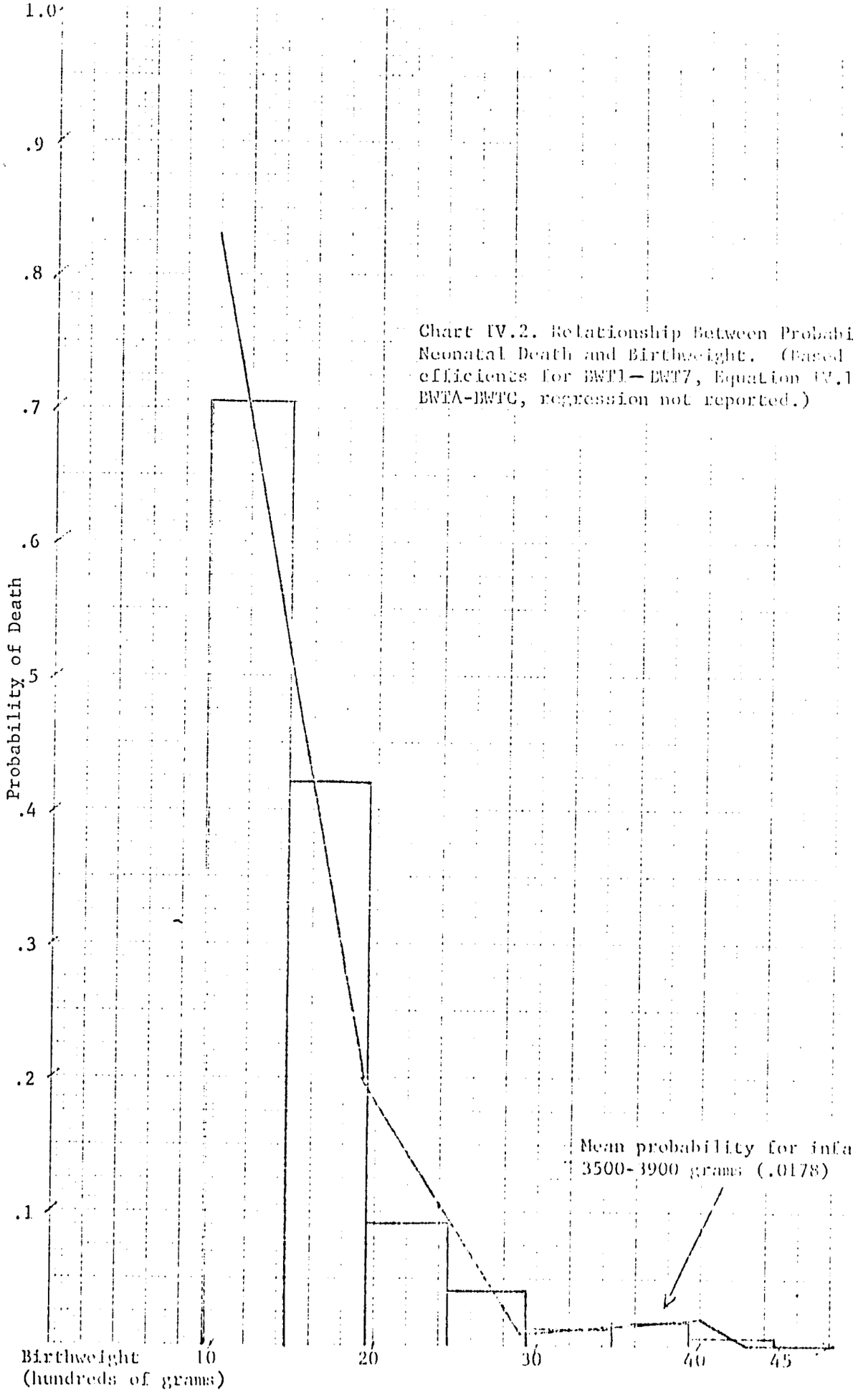


Chart IV.2. Relationship Between Probability of Neonatal Death and Birthweight. (Based on coefficients for BWT1—BWT7, Equation IV.1 and BWTA-BWTC, regression not reported.)

The results are, as they should be, basically the same as for the binaries. But they permit a more precise appreciation of the improvement in an infant's chances of survival as its birthweight increases. For example, over the important interval below 2000 grams, every gain of 100 grams reduces the probability of death by .07. Between 2000 and 2900 grams each additional 100 grams reduces the probability of death by .02.

In summary, the results show that birthweight is extremely important to an infant's survival during the neonatal period. Infants below 2500 grams, and especially below 2000 grams, face much higher risks of death than do heavier infants, and the risk is higher the lower the baby's weight. Above 3000 grams additional weight makes little difference, although it does improve the infant's chances of survival slightly.

Length of Pregnancy

The review of the literature revealed one study (Erhardt, et.al., 1970), which shows that mortality rates are higher for premature than for full-term infants, even when birthweight is held constant. We tested this hypothesis by entering length of pregnancy in the regressions. Four classes were specified: six months or less (PLG1), seven months (PLG2), eight months (PLG3), and nine months -- the omitted class.

The equations show that length of pregnancy is consistently very important as a determinant of survival during the neonatal period. And the fact that it is highly significant in equations which also include birthweight supports

the notion that length of pregnancy has an independent effect on mortality. More specifically, the coefficients show that infants born during the seventh month, or earlier, have significantly higher probabilities of death than full-term infants. There is no statistically significant difference between infants born during the eighth month and those who are full-term.

But the coefficients also suggest a differential between sixth and seventh month babies that is not intuitively reasonable. The coefficient for PLG2, .36, indicates that the added risk of being born during the seventh month is higher than the risk associated with birth during the sixth or earlier months (the coefficient for PLG1 is only .24), when we expect the opposite to be true. This result probably reflects the fact that there is only one living infant, and a handful of infants who survived the neonatal period but died later, in each of these classes.[1/] Thus while the general conclusion -- that prematurity greatly increases an infant's risk of death -- is valid, the very small sample of survivors in the first two classes makes it difficult to trace out the details of the relationship accurately.

Sex of the Child

Although there are some inconsistencies across countries and social classes, female infants generally have lower mortality rates than males, particularly during the neonatal period (Abramowicz and Barnett, 1970; Hammoud, 1965; Morris and Heady, 1955). Since females are lighter at birth (St. George, et.al, 1970; Pethybridge, et.al., 1974; and Chapter II of this report), this differential appears to be independent of the effect of birthweight on mortality.

1/ See Table 2 in the appendix to this chapter.

Equations IV.1-IV.3 indicate that the sex of the child does have an independent effect on its chances for survival. Female infants have a lower probability of death -- about .03 lower -- even after factors like birthweight and length of pregnancy have been taken into account.

Medical Care During Birth (Birthplace and Who Attended Birth)

Studies that have examined the general relationship between medical care and infant mortality have been unable to produce evidence that it plays more than a minor role (e.g., Scrimshaw, 1970). But none of the studies we reviewed examined the effects of differences in medical care during delivery. This is an important area of concern in developing countries, where home deliveries are common and the attendant, if any, is likely to be a midwife rather than a doctor. And from the point of view of program design, it would be useful to have some indication whether efforts should be made to promote one type of care over another. Two measures of medical care during delivery -- the place of birth and the attendant -- were available from the survey. These were tested in the regressions.

A large majority of the infants born in San Juan during the survey period were born in the hospital, and a smaller majority were delivered by midwives (Tables 3 and 4 in the chapter appendix). But the regression results show no difference in the probability of death for the sizeable minority that were born at home -- the coefficient for BPLC1 is statistically insignificant. Similarly, no difference in the probability of death is associated with delivery by a midwife (ATND1) rather than a physician (the omitted class). The reader wil recall that the birthweight analyses failed to reveal any significant difference

in birthweight that could be associated with type of attendant.[1/]

Interestingly, there is a weak association between having an attendant other than a midwife or physician (ATND2) and a higher probability of death during the neonatal period. The differential of about .03 is weakly significant, at the 80 percent level of confidence, in Equations IV.1 and IV.2, but not in Equation IV.3. Needless to say, this result should not be given much weight, but it is noted here simply because it is both reasonable and intriguing.

Parity and Spacing

We initially hypothesized that the effects on mortality of parity and spacing are indirect, and due primarily to their effects on birthweight. The analysis of Chapter II shows that there are important relationships between these factors and birthweight -- first babies are significantly lighter than later ones, and closely spaced pregnancies result in lighter infants than do more widely spaced pregnancies. But some studies have reported higher mortality rates for higher parity infants (Wray, 1971), and the birthweight regressions do not provide an explanation for this. Thus we included parity, and spacing, in the regressions to test for any direct effects they might have on neonatal mortality.

Equation IV.1 shows that parity does have an independent effect on neonatal mortality. While there are no differences in the mortality experience of the lower parities, infants of the seventh or higher parities (TPRG4) have a

---

1/ It seems reasonable to believe that these findings should be generalized beyond San Juan only with the utmost caution, since they probably depend on the type of training midwives and physicians receive, and this may vary between areas. It would be useful to know something about the level of competence of midwives in San Juan.

significantly higher probability of death. Equation IV.3 indicates that this result is stable in the presence of a variety of social factors, including total family size (FAM).

Spacing was also tested for any direct relationship with neonatal mortality.[1/] Equation IV.2 shows that only an interval of more than four years (SPC6) showed any significant association with mortality: infants in this group have a probability of death lower by .052 than infants in the omitted group (25-36 months). We considered the possibility that such a long interval might, in fact, reflect greater age and maturity on the part of the mother, and that this maturity might be the true cause of the mortality differential. But the results for spacing were unchanged in regressions that added age of mother, and the regressions gave no evidence that infants of older mothers experienced lower risks of neonatal death.

Environmental Factors

A number of social and environmental factors were included in the regressions to test the belief that these factors are relatively unimportant during the neonatal period. Equation IV.3 is typical of the results. Residence in a very poor (NBR3) or rural (NBR4) neighborhood is associated with higher mortality, but household size -- which we initially thought might be a measure of the same factor as parity -- and the number of rooms in the dwelling are not significant. The marital status of the mother is significantly associated with neonatal mortality -- infants born to single mothers (MARM1) have a higher probability of death, .054 higher than those born to married women. (MARM2 represents infants born into common-law marriages.)

Regressions not shown here indicated that there is no association between neonatal mortality and type of toilet or water facilities. Only one class of father's occupation was significant and both the class (other clerical workers) and its splendid isolation suggest that little faith should be put in the result.

1/ Because one class of the spacing variable (SPCX=first children) is identical with one class of TPRC (first children are the omitted class), both factors cannot be entered in the same regression. In technical terms, they form a linearly dependent set of variables and the regression cannot be computed because this linear dependence makes it impossible to invert the moments matrix.

The educational levels of the father and the mother were tested in separate regressions, with the same result in each case. Only one class was significant, that with no education, and the coefficient indicated that the probability of death was lower for those infants than for those born to parents with some primary education. The numbers of infants in the no education classes are small, and these results are probably due to the idiosyncrasies of the sample.

Finally, we would have liked to test for any effect of breastfeeding on neonatal mortality. But as Chapter III documents, a great many of the neonatal deaths were never breastfed because they died within 48 hours of birth or because they were in the hospital at the time to begin breastfeeding. Thus it appears that their imminent death precluded breastfeeding, rather than that the failure to breastfeed had anything to do with their deaths. For this reason no attempt was made to include information about breastfeeding in these regressions.

In conclusion, only two of the social/environmental factors -- type of neighborhood and marital status of the mother -- showed any significant relationship with neonatal mortality. Their effects are on a much smaller scale than those of birthweight or length of pregnancy. These results, and those reported in the next chapter, provide new evidence that social and environmental factors are of little importance until later in an infant's life.

APPENDIX TO CHAPTER IV
Distribution of Observations 0 to 11 Months of Age by Factors Used in the Regressions

1. By Birthweight (in hundreds of grams)

	Neonatal Deaths	Deaths 1-11 months	Sample (0-11 months)
Less than 15	166	8	0
15-19	156	23	1
20-24	124	76	11
25-29	103	186	26
30-34	71	230	66
35-39	82	228	49
40-44	43	44	31
Over 45	17	21	8
Not specified	44	120	12
TOTAL	806	936	204

2. By Length of Pregnancy

	Neonatal Deaths	Deaths 1-11 months	Sample (0-11 months)
Six months or less	151	3	1
Seven months	184	34	1
Eight months	135	135	18
Nine months	303	720	183
Not specified	33	44	1
TOTAL	806	936	204

3. By Birthplace

	Neonatal Deaths	Deaths 1-11 months	Sample (0-11 months)
Hospital	702	751	179
Home, or other	104	183	25
Not specified	0	2	0
TOTAL	806	936	204

APPENDIX TO CHAPTER IV, continued

4. By Attendant at Birth

	Neonatal Deaths	Deaths 1-11 months	Sample (0-11 months)
Physician	285	177	74
Midwife	423	592	111
Other, or no one	81	142	14
Not specified	17	25	5
TOTAL	806	936	204

5. By Sex of Child

	Neonatal Deaths	Deaths 1-11 months	Sample (0-11 months)
Female	460	482	101
Male	346	454	103
TOTAL	806	936	204

6. By Total Number of Pregnancies (Parity)

	Neonatal Deaths	Deaths 1-11 months	Sample (0-11 months)
One	227	202	56
Two	167	166	46
Three-Four	175	233	64
Five-Six	101	169	19
Seven or more	136	155	19
Not specified	0	11	0
TOTAL	806	936	204

APPENDIX TO CHAPTER IV, continued

7. By Type of Neighborhood

	Neonatal Deaths	Deaths 1-11 months	Sample (0-11 months)
Well-to-do or good	40	22	20
Medium	166	161	72
Poor	145	156	45
Very poor	47	79	4
Rural or no neighborhood	400	514	62
Not specified	8	4	1
TOTAL	806	936	204

8. By Occupation of the Father

	Neonatal Deaths	Deaths 1-11 months	Sample (0-11 months)
Professionals	15	8	11
Other clerical workers	28	25	10
Sales: insurance and real estate, working proprietors, manufacturers' agents, etc.	7	3	10
Shop assistants and salesmen	38	49	11
Farmers and farm managers	26	24	1
Farm workers	281	399	52
Drivers, road transport	23	30	9
Toolmakers, machinists, plumbers, etc.	35	31	16
Bricklayers, plasterers, and construction workers, n.e.c.	81	73	15
Millers, bakers, brewmasters and related workers	17	15	10
Laborers, n.e.c.	33	34	3
Other crafts	35	32	18
Electricians, carpenters, painters, etc.	33	24	5
Firefighters, police, guards, etc.	32	19	6
Other	44	61	13
Not specified	78	109	14
TOTAL	806	936	204

APPENDIX [illegible] IV, [illegible]

9. By Type of Toilet Facilities

	Neonatal Deaths	Deaths 1-11 months	Sample (0-11 months)
Flush or communal	179	95	71
Other	222	267	64
None	403	570	68
Not specified	2	4	1
TOTAL	806	936	204

10. By Type of Water Facility

	Neonatal Deaths	Deaths 1-11 months	Sample (0-11 months)
Piped inside	196	123	75
Piped outside	90	113	20
Other	518	698	108
Not specified	2	2	1
TOTAL	806	936	204

11. By Interval Since Last Birth (in months)

	Neonatal Deaths	Deaths 1-11 months	Sample (0-11 months)
12 or under	87	86	16
13-15	79	125	19
16-18	61	96	18
19-24	103	149	20
25-36	116	149	26
37-48	60	55	26
over 48	72	63	19
Not specified	228	213	60
TOTAL	806	936	204

APPENDIX TO CHAPTER IV, continued

12. By Marital Status of Mother

	Neonatal Deaths	Deaths 1-11 months	Sample (0-11 months)
Single	114	146	7
Married	527	554	173
Common-law or other	165	232	24
Not specified	0	4	0
TOTAL	806	936	204

13. By Number of Persons in the Household

	Neonatal Deaths	Deaths 1-11 months	Sample (0-11 months)
1 or 2 persons	121	100	0
3 or 4 persons	230	292	79
5 or 6 persons	192	228	68
7 or 8 persons	154	172	35
9 or more persons	109	144	22
Not specified	0	0	0
TOTAL	806	936	204

14. By Number of Rooms in the Dwelling

	Neonatal Deaths	Deaths 1-11 months	Sample (0-11 months)
1 room	121	168	22
2 rooms	190	245	38
3 or 4 rooms	361	402	98
5 or more rooms	126	111	45
Not specified	8	10	1
TOTAL	806	936	204

CHAPTER V. Determinants of Mortality Among Infants 28 Days to 5 Months

Introduction

Infant deaths drop off sharply after the neonatal period. In San Juan 906 of the 1,900 survey deaths occurred during the first four weeks, while only 163 occurred during the following month -- a high number when compared with the experience of developed countries, but much below the toll of the first month. The distribution of deaths by age remains steady over the next couple of months, beginning another decline at four months (Table V.1 and Chart 1 in Chapter IV). After five months the decline is quite rapid.[1/]

As Table V.1 shows, this general pattern is repeated in other projects. Deaths are very high during the first four weeks and decline with the passing months. But the details of the pattern vary somewhat. In some projects (e.g. Cali, Colombia) the decline is less rapid. In others (Sao Paulo, Brazil, for example) it is more rapid.

The decline in the number of deaths with age probably reflects differences in the factors in operation and in the vulnerability of infants of different ages. This suggests that it makes good sense to divide the first

---

1/ In all, deaths during the first six months of life accounted for 86.4 percent of all infant deaths in San Juan, with 47.7 percent during the neonatal period and 38.7 percent in the later months.

TABLE V.1. Number of Infant Deaths by Month of Age for Selected South American Survey Sites

	ARGENTINA: Chaco Province	ARGENTINA: San Juan Number in Survey	ARGENTINA: San Juan Number Retained for Analysis*	BOLIVIA: La Paz	BRAZIL: Sao Paulo	COLOMBIA: Cali
0-27 Days	568	906	806	1049	1958	536
28 Days to 1 Month	111	163	150	277	450	91
2 Months	134	164	156	210	348	61
3 Months	129	164	151	196	273	63
4 Months	96	141	135	144	196	61
5 Months	84	103	97	103	147	65
6 Months	72	68	64	141	99	54
7 Months	54	56	54	115	98	58
8 Months	55	44	42	119	77	48
9 Months	40	28	24	110	58	45
10 Months	39	33	33	113	45	37
11 Months	28	30	30	108	39	34

Source: Puffer and Serrano, 1973, pp. 430-1.

year of life into several age groups for study. The analysis of neonatal deaths has already been discussed. This chapter takes up the analysis of mortality among infants 28 days to 5 months.

The set of observations defined for the analysis included all infants in the survey between 28 days and 5 months at death or interview, a total of 689 deaths and 99 living infants.[1/] This definition corresponds to the concept of cohort death rates -- rates that compare deaths in a certain age group with the living population in the same age group. It differs somewhat from the conventional definition of death rates for infants; conventional rates compare deaths in a particular age group with all live births that occurred during the same period. But since some of the variables used in this analysis depend on information which is specific to the child's age -- for example, whether the infant was vaccinated and its diet -- it seemed most logical to limit comparisons to children in the same fairly narrow range of ages.

The literature review identified a large number of factors associated with differences in infant mortality (Russell and Burke, 1974, Part One). These include the two major determinants of neonatal mortality, birthweight and length of pregnancy, which may still affect the survival of infants in this age class, especially the younger ones. They also include: the medical care given the child and its disease history; the child's diet -- especially the effects of breastfeeding; the sex of the child; and a variety of social and environmental conditions such as the toilet and water facilities available to the family, family size, father's occupation, housing, type of neighborhood and marital status of the mother.

1/ The weights used for these observations are set out in Appendix C.

Information on one or more aspects of each of these factors was available from the survey, and thus it was possible to test them in the regressions. The results are presented in Equations V.1 through V.5. Equation V.1 is the basic equation. Equation V.2 varies from the basic equation only in that it uses a more detailed set of variables to represent the infant's breastfeeding history. Equation V.3 presents the best of the specifications that included measures of diet adequacy. Equations V.4 and V.5 expand the specification of the basic equation by adding a number of social variables.

The dependent variable for this analysis is a binary which takes on the value 1 (one) for deaths and 0 (zero) for living children. As in the neonatal mortality regressions, it is interpreted as a probability of death.[1/] Estimates of the dependent variable calculated from the equations give the probability of death for children with the characteristics specified by the values of the independent variables. Again, as before, all the independent variables are in binary form, and the coefficient for a particular variable gives the difference in the probability of death associated with membership in the class represented by the variable, as compared to membership in the omitted class. For example, the coefficient for VACC gives the differential associated with having been vaccinated -- infants who were not vaccinated are defined as the omitted class for this factor.

Definitions of the variables used to represent each factor are noted, as needed, throughout the text, but the reader will find a complete list of definitions in Appendix B. The distributions of the observations over the classes or intervals represented are given in the chapter appendix. Some of the variables shown in the equations are combinations of classes defined in

---

1/ But the reader should remember that, because of the way we defined the observation set, the probability corresponds to a cohort death rate, not to an infant mortality rate of the standard type.

Appendix B. For example, the first birthweight variable, shown as BWT1, 2, and 3, is the first three birthweight intervals combined, and it includes all infants in the observation set whose birthweight was 2400 grams or less. The combined variables were necessary because this analysis is based on a considerably smaller set of observations than the analyses of Chapters II and IV and, for some factors, there were not enough observations to support the full range of classes defined for the larger observation sets.[1/]

In general, the regressions indicate that a larger and more diverse set of factors becomes important once the infant has survived the neonatal period. At the same time, our ability to generalize across infants and to identify determinants with confidence declines in comparison with the neonatal regressions. For example, the coefficients of determination (R^2's) indicate that less of the variation in the survival experience of this age group has been explained -- they vary from .12 in Equation V.1 to a high of .19 in Equation V.5 And, unlike the results in earlier chapters, the coefficients for individual factors are sensitive to changes in specification or variable form. We will discuss these difficulties in the context of particular factors in the following sections. The last section will return to a consideration of the problem in general and will suggest how it might be dealt with in further research.

Birthweight, Length of Pregnancy, and Sex of the Child

Both low birthweight and prematurity are associated with higher mortality between 28 days and 5 months, although their effects are much less important than

---

1/ The deletion of a class, e.g. NBR1, has the effect of combining it with the omitted class.

EQUATION V.1: Basic Equation

Dependent Variable: Deaths 28 Days to 5 Months = 1

Set of Observations: Records of all infants 28 days to 5 months at time of interview or death.

Name of Factor	Name of Variables	Coefficient	t-score	F for Factor (see note)
Birthweight	BWT1,2 and 3	0.169	3.63[b]	5.24[b]
	BWT4	0.071	2.19[a]	
	BWT5	0.004	0.16	
	BWT6 and 7	-0.029	0.92	
	BWT8	0.131	2.47[a]	
Length of Pregnancy	PLG1,2 and 3	0.115	3.29[b]	10.83[b]
Whether Child Was Breastfed	BRF1,2,3,4,5 and 6	-0.092	2.59[b]	6.70[b]
Sex of the Child	SEX	-0.036	1.81	3.27
Whether Child Received Any Vaccinations	VACC	-0.087	3.94[b]	15.56[b]
Type of Water Facilities	WTR1	0.047	0.94	0.87
	WTR2	-0.001	0.02	
Type of Toilet Facilities	TLT1	0.040	0.96	3.76
	TLT2	0.106	2.20[a]	
Number of Rooms in Dwelling	RMS1	0.149	4.15[b]	6.08[b]
	RMS2	0.024	0.94	
	RMS3	0.040	1.49	
Constant Term		0.070		

R^2 for the equation: .122
F for the equation: 6.76
a. significant at the 95 percent level of confidence.
b. significant at the 99 percent level of confidence.

NOTE: Each F-value indicates whether the factor as a whole is statistically significant. When a factor is represented by a single variable the F is simply the square of the t-score. When a factor is represented by more than one variable it turns out that, in general, the factor is significant whenever at least one of the variables has a significant t-score. Limitations on computer time made it impossible to calculate these F's for equations other than the basic equation in each chapter but the reader can tell when a factor is significant by examining the t-scores.

EQUATION [illegible]: [illegible]

Dependent Variable: Deaths 28 Days to 5 Months = 1

Set of Observations: Records of all infants 28 days to 5 months at time of interview or death.

Name of Factor	Name of Variable	Coefficient	t-score
Birthweight	BWT1,2 and 3	0.142	2.93[b]
	BWT4	0.067	2.05[a]
	BWT5	-0.017	0.63
	BWT6 and 7	-0.043	1.36
	BWT8	0.125	2.37[a]
Length of Pregnancy	PLG1,2 and 3	0.108	3.13[b]
Whether Child Was Breast-fed, and Age Weaning Started	BRF1	-0.074	1.97[a]
	BRF2	-0.042	0.98
	BRF3,4 and 5	-0.050	1.19
	BRF6	-0.159	4.09[b]
Sex of the Child	SEX	-0.023	1.16
Whether Child Received Any Vaccinations	VACC	-0.086	3.91[b]
Type of Water Facilities	WTR1	0.024	0.47
	WTR2	-0.010	0.23
Type of Toilet Facilities	TLT1	0.086	2.03[a]
	TLT2	0.153	3.13[b]
Number of Rooms in Dwelling	RMS1	0.165	4.63[b]
	RMS2	0.006	0.22
	RMS3	0.038	1.39
Constant Term	-	0.126	-

R^2 for the equation: .147
F for the equation: 6.96
a. significant at the 95 percent level of confidence.
b. significant at the 99 percent level of confidence.

[illegible] 3: [illegible] of Diet Adequacy
[illegible] variables)

[illegible] variable: [illegible] 28 [illegible] 5 months = 1

[illegible] of observations: [illegible] records of [illegible] infants [illegible] days to 5 months at time of interview [illegible] sth.

Name of [illegible]	Name of Variable	Coefficient	t-score
Birthweight	BWT1,2 and 3	0.101	2.10[a]
	BWT4	0.052	1.58
	BWT5	-0.016	0.64
	BWT6 and 7	-0.053	1.65
	BWT8	0.077	1.43
Length of Pregnancy	PLG1,2 and 3	0.141	4.02[b]
Whether the Child Was Breastfed	BRF1,2,3,4,5 and 6	-0.067	1.84
Sex of the Child	SEX	-0.043	2.04[a]
Whether Child Received Any Vaccinations	VACC	-0.071	3.11[b]
Type of Water Facilities	WTR1	0.0005	0.01
	WTR2	-0.019	0.43
Type of Toilet Facilities	TLT1	0.048	1.03
	TLT2	0.105	2.02[a]
Type of Neighborhood	NBR2	0.006	0.22
	NBR3	0.174	2.98[b]
	NBR4	0.052	1.94
Number of Persons in Household	FAM1	0.035	1.14
	FAM2	0.016	0.63
	FAM3	0.128	3.50[b]
	FAM4	0.104	2.33[a]
Adequacy of Diet in Terms of Calories	CAL1	0.069	1.87
	CAL2	-0.007	0.22
	CAL3	-0.064	2.55[a]
Constant Term	-	0.075	-

R^2 for the equation: .157
F for the equation: 6.18
a. significant at the 95 percent level of confidence.
b. significant at the 99 percent level of confidence.

EQUATION V.4: Basic Equation with Social Variables, Including Father's Occupation

Dependent Variable: Deaths 28 Days to 5 Months = 1

Set of Observations: Records of all infants 28 days to 5 months at time of interview or death.

Name of Factor	Name of Variables	Coefficient	t-score
Birthweight	BWT1,2 and 3	0.116	2.26[a]
	BWT4	0.062	1.92
	BWT5	-0.019	0.70
	BWT6 and 7	-0.032	0.99
	BWT8	0.076	1.42
Length of Pregnancy	PLG1,2 and 3	0.115	3.28[b]
Whether the Child Was Breastfed	BRF1,2,3,4,5 and 6	-0.114	3.18[b]
Sex of the Child	SEX	-0.063	2.94[b]
Whether Child Received Any Vaccinations	VACC	-0.069	2.98[b]
Type of Water Facilities	WTR1	0.019	0.36
	WTR2	-0.024	0.54
Type of Toilet Facilities	TLT1	-0.035	0.74
	TLT2	0.040	0.74
Type of Neighborhood	NBR2	0.006	0.20
	NBR3	0.175	2.91[b]
	NBR4	0.092	3.34
Number of Persons in Household	FAM1	0.010	0.34
	FAM2	0.014	0.51
	FAM3	0.109	3.02[b]
	FAM4	0.099	2.25[a]
Number of Rooms in Dwelling	RMS1	0.172	4.61[b]
	RMS2	0.054	2.04[a]
	RMS3	0.014	0.51
Father's Occupation	OCF1	0.0003	0.01
	OCF2,3 and 4	-0.103	2.71[b]
	OCF8,9,10,11,12 and 13	-0.068	2.44[a]
	OCF5,6,7,14 and 15	-0.045	1.45
Constant Term		0.062	

R^2 for the equation: .174
F for the equation: 5.95
a. significant at the 95 percent level of confidence.
b. significant at the 99 percent level of confidence.

EQUATION 7.5: Basic Equation with Social Variables, Including Marital Status of Mother

Dependent Variable: Deaths 28 Days to 5 Months = 1

Set of Observations: Records of all infants 28 days to 5 months at time of interview or death.

Name of Factor	Name of Variables	Coefficient	t-score
Birthweight	BWT1,2 and 3	0.042	0.81
	BWT4	0.060	1.87
	BWT5	-0.017	0.67
	BWT6 and 7	-0.033	1.03
	BWT8	0.082	1.55
Length of Pregnancy	PLG1,2 and 3	0.109	3.12[b]
Whether Child was Breastfed	BRF1,2,3,4,5 and 6	-0.090	2.54[a]
Sex of the Child	SEX	-0.048	2.31[a]
Whether Child Received Any Vaccination	VACC	-0.067	2.96[b]
Type of Water Facilities	WTR1	0.009	0.17
	WTR2	-0.022	0.51
Type of Toilet Facilities	TLT1	-0.014	0.31
	TLT2	0.057	1.09
Type of Neighborhood	NBR2	0.008	0.28
	NBR3	0.126	2.08[a]
	NBR4	0.088	3.28[b]
Number of Persons in Household	FAM1	0.028	0.93
	FAM2	0.014	0.56
	FAM3	0.119	3.37[b]
	FAM4	0.098	2.25[a]
Number of Rooms in Dwelling	RMS1	0.179	4.92[b]
	RMS2	0.038	1.44
	RMS3	0.026	0.95
Marital Status of the Mother	MARM1	0.220	4.14[b]
	MARM2	0.065	1.91
Constant Term		0.046	

R^2 for the equation: .185
F for the equation: 6.91
a. significant at the 95 percent level of confidence.
b. significant at the 99 percent level of confidence.

during the neonatal period. In large part this simply reflects the fact that few very underweight or premature infants survive the first month.[1/]

The coefficients for the birthweight variables in Equations V.1 and V.2 show that infants who weight 2400 grams or less at birth (BWT1,2 and 3), but who nevertheless manage to survive the first month, still face a higher risk of death -- the estimates of the differential range from .14 to .17 in these equations -- than do heavier infants. Infants between 2500 and 2900 grams (BWT4) also face a higher risk of death, but the differential does not always reach significance at the 95 percent level of confidence. The coefficients for the heavier weight classes (BWT5,and BWT6 and 7) do not indicate that additional weight beyond 2000 grams carries any significant advantage in terms of the infant's survival.

The combined variable PLG1,2 and 3 represents all infants who are less than full term at birth -- there were too few to permit a more detailed breakdown. The coefficient, which is stable in value across the different equations, shows that infants born prematurely are consistently at greater risk than full term infants: their probability of death is estimated to be about .12 higher.

Birthweight and prematurity are conditions set at birth, and they have already been shown to have very dramatic effects on the survival experience of infants during the first month. It seems reasonable to believe that the coefficients reported in this chapter reflect their lingering effects on the younger infants in this age group rather than an effect that applies uniformly through the fifth month. But a larger sample, one large enough to permit the age range to be divided more finely, would be required to confirm this hypothesis.

1/ This is clear if we compare the distributions of birthweight ength of pregnancy for neonatal deaths (Chapter IV, appendix) with the r older deaths, or living infants (Chapter IV, appendix, and appendix this chapter).

Finally, the coefficient for sex of the child suggests that female infants continue to have an advantage in survival, other things constant, between the ages of 28 days and 5 months. But this result is quite sensitive to the specification of the rest of the equation. The estimated differential ranges from a low of .022 in Equation V.2 to a high of .063 in Equation V.4 and in two cases the estimate is not significant at the prescribed levels (Equations V.1 and V.2).

Medical Care

As the literature review indicated (Russell and Burke, 1974, Part One), previous studies have had difficulty trying to establish a significant link between medical care and infant mortality. In what was probably the most carefully designed of the studies we reviewed, Scrimshaw (1970) reported that there was a modest decrease in preschool deaths, relative to a control village, in a Guatemalan village that was provided with a series of medical and public health services -- including a clinic, vaccination program, safe water supply, and sanitary privies for each house. But because the services were provided in combination, it was impossible to tell which one, or ones, were responsible for the result.

The survey requested information on several aspects of each child's medical care -- the number of physician or clinic visits, the number of days in the hospital, whether the child was vaccinated, and the types of vaccinations, and we attempted to test these in the regressions. Because of the frequently noted synergism between infectious disease and malnutrition, we also investigated the information about the child's history of diseases. But for all but one of these factors the data were so incomplete that either the test must be considered inconclusive, or no test was possible.

Visits for medical care were distinguished by whether they took place in a clinic or a physician's office and whether the visit was a simple follow-up for a well child, or involved some illness. We dropped the distinction between clinic and physician, but maintained the one between well-child and other visits. Regressions which included numbers of well-child visits showed no difference between infants with such visits (several levels of visits were represented) and those without. Similarly the total number of visits showed no association with the probability of death.

Unfortunately, these results cannot be taken very seriously because of the large number of infants for whom there was no information. For some reason this lack of information was confined to the infant deaths -- 363 of the 689 deaths lacked information on well-child visits and 307 lacked information on other visits, compared to only one of the living children.[1/]

Lack of information was again a serious problem for the child's disease history. We experimented with one crude variable which indicated that the child had suffered from an infectious disease of some sort. But even this much was not known with any certainty for most of the infants, and our failure to find any association between disease history and mortality must be attributed to a simple lack of data.

The one factor for which the information was quite complete was whether the child had received any vaccinations (see Table 5 in the chapter appendix). And the coefficient for VACC in the equations indicates that children who were vaccinated had significantly better chances of survival than those who were not.

1/ We did not test the number of hospital days in the regressions since it seemed highly likely that illness, and imminent death, were the cause of hospitalization, and not the other way around. For the same reason, we did not test visits for other than well-child care separately, but only the total of all visits.

The estimated differential is reasonably stable across specifications, at about .07-.09, and is consistently significant at the 99 percent level of confidence. The types of vaccinations given were frequently left unspecified on the questionnaires and it was thus impossible to test whether any particular kind of vaccination was primarily responsible for this more general result.

Nutrition

Chapter III has described the nutritional classification system that we developed to make use of the diet information collected by the survey. In addition to variables based on this classification, we tested variables representing the breastfeeding history of the infant.

The literature review uncovered very general agreement that breastfeeding reduces infant mortality (Russell and Burke, 1974, Part One). Three reasons are suggested for this: in developing countries breastfeeding is likely to be the most sanitary method of infant feeding; breast milk is considered to be nutritionally superior to the available, or accepted, alternatives; and breast milk contains agents that are believed to increase the child's resistance to infectious disease. The review also indicated that breast milk alone does not provide adequate nutrition after the first few months, but this problem only becomes serious for infants older than those in the age group being considered (Cameron and Hofander, 1971).[1/]

The regression results agree with these earlier studies: breastfed infants were significantly more likely to survive beyond the fifth month than these who were not breastfed. The probability of death is estimated to be about

---

1/ The exception to this statement is iron. An infant's iron stores may be depleted as early as three months and breast milk does not supply iron in sufficient quantities (Burton, 1965).

.09 lower for breastfed infants,[1/] and the differential is stable across specifications.

We attempted to measure the relationship in more detail by defining a set of variables that distinguish between infants both by whether they were breastfed and by the age at which weaning started. (Note that this does not mean that breastfeeding was terminated at that age, but only that other foods were introduced into the infant's diet.[2/]) The results, illustrated in Equation V.2, are not particularly illuminating. Infants whose weaning started at less than one month (BRF1) have a significantly lower probability of death than infants who were not breastfed, and the estimated differential, .07, is close to the coefficient estimated for all breastfed infants in Equation V.1. Infants weaned later (BRF2 and BRF3,4 and 5) also have lower probabilities of death, but the differentials were unstable in the various specifications tested and never reached significance at the 95 percent level. A possible interpretation of these results is that breastfeeding plus early supplementation with other foods is the best nutrition plan, but the results for BRF6 (never weaned or age of weaning unknown) are difficult to incorporate into this interpretation. These infants have even better chances of survival than any of the other three groups.

The variables based on the nutritional classification scheme did not produce useful results for this age group. Understandably, they showed a high intercorrelation with each other and the coefficients were particularly unstable when variables representing several of the nutrient groups were included. They also showed some sensitivity to the inclusion of the breastfeeding variables;

1/ Except in Equation V.3 where the estimate is affected by the collinearity between breastfeeding and the variables representing the adequacy of the child's diet in terms of calories.

2/ We did not attempt to define variables based on the age weaning finished because this information was not specified in a large number of cases.

the child's milk intake was a particularly major item in its diet classification for this age group. But deleting breastfeeding from the regressions did not improve the results. The coefficients were seldom significant, and when they were they were as likely to be perverse as in the expected direction.

Equation V.3 is one of the best of the regressions that included the diet adequacy variables. Diet adequacy in terms of calories was generally more likely to produce the expected signs than was diet adequacy in terms of protein. In Equation V.3 a diet probably inadequate in calories (CAL1) is associated with a higher risk of death, and the coefficient is significant at the 90 percent level; the omitted group is infants with an adequate diet. But a diet classified as grossly inadequate in calories (CAL2) shows no association with mortality. CAL3 represents infants whose diets could not be classified.[1/,2/]

Environmental and Social Conditions

Environmental conditions encompass such things as housing and sanitation, which may contribute to the infant's exposure to disease. Information was available from the survey about three aspects of the child's environment -- the water and toilet facilities available to the family and the number of rooms in the dwelling.

1/ Distributions of the observations over the classes of the nutritional system are given in Chapter III.

2/ There are, of course, many reasons why the nutritional classification might not prove useful. One which has not been mentioned previously is the possibility that not all the diet information collected in the survey was accurately transferred to the tape record. We examined five or six of the questionnaires in the process of doing the tape edit and noticed that one of the records had considerably more diet information than was stored on the tape. Since part of the information was on the tape the infant would have been classified, but incorrectly. This raises the possibility that the diet information for other infants was incompletely transcribed and thus that these infants were misclassified.

The survey reported whether the family had water piped into the dwelling, piped to an outside source, or some other, unspecified, arrangement. Equation V.1 is typical of the results: none of the regressions produced any evidence of an association between the source of water and mortality among infants 28 days to 5 months.

The type of toilet facilities available to the family does show an association with mortality. In Equation V.1 infants whose families have no toilet facilities (TLT2) have significantly higher risks of death than those whose families have access to either a flush (the omitted class) or a non-flush toilet (TLT1). The coefficient for TLT2 indicates that the probability of death is .11 higher for infants in families with no toilet facilities than for those whose families have flush toilets.

But this coefficient is very sensitive to the other variables included in the regressions. Even the more detailed representation of the breastfeeding variables affects the results, although in this case the change increases the size and significance of TLT2 and produces a significant t-score for TLT1 (Equation V.2). But whenever social variables, such as type of neighborhood are included, both coefficients drop in value and their t-scores fall below the prescribed levels of significance (Equations V.4 and V.5). This is due to collinearity between the TLT variables and these social variables, a problem which will be discussed more fully in the last section. At this point suffice it to say that, when such collinearity is present, it is impossible to be sure that the low t-scores really mean that the family's toilet facilities have no influence on the risk of death.

Extremely poor housing conditions also increase an infant's chances of death. Equation V.1 shows that an infant living in a one room dwelling (RMS1) has a much

higher probability of death, about .15 higher, than infants living in a three or four room dwelling.[1/] None of the other RMS variables are associated with significant differentials in the risk of death.

An attempt was made to determine whether the care the child received, as represented by the person primarily responsible for the child, had any effect on its chances of survival. In very few cases was anyone other than the mother named and none of the variables based on this information showed any association with mortality. We were somewhat suspicious of the information, however, especially since the number of working mothers exceeded the number who reported that someone else took primary responsibility for the child. But a variable defined to represent working mothers also showed no association with mortality.

We then tested a variety of more general social variables. Most of these are usually thought of as proxies for the family's standard of living -- family size, type of neighborhood, occupation of the father, and marital status of the mother.[2/] Most of them showed some association with mortality. But, unfortunately, they also exhibited a tendency to be correlated with some of the other variables tested, variables which are capable of having more direct links with the risk of death. As a result, the coefficient estimates and t-scores of these variables proved sensitive to the addition of some of the social variables. As noted earlier, the variables representing toilet facilities were especially sensitive to this problem

Infants born to families in rural (NBR4), or very poor urban (NBR3), neighborhoods faced higher risks of death than those born to families in

---

1/ This variable does not appear to be measuring crowdedness. Persons per room was tested in some of the regressions and was insignificant.

2/ The educational levels of each parent were also tested, separately, with the same results as in the neonatal analysis. Only the very small group of parents with no education showed any difference from the omitted group, those with some primary education, and this difference was in the "wrong" direction.

"medium" urban neighborhoods (Equations V.4 and V.5). Two things should be noted about these results. The first is that like the (potentially) more direct causes of mortality, they are sensitive to the inclusion of other social variables. The other is that there are very few living children in the NBR3 class. This makes the result and especially the size of the coefficient, which is quite large, rather doubtful.

Equations V.4 and V.5 show that infants born into large families -- those with seven or more members in the household (FAM3 and FAM4) - have significantly higher probabilities of death than those born to 3 or 4 person families (the omitted class). The differential is approximately the same for 7 or 8 person families (FAM3) as for larger ones (FAM4) and averages about .10-.12.

Father's occupation shows an intuitively reasonable association with mortality. Infants whose fathers are professionals (OCF2-4) have a considerable advantage over those whose fathers are farm workers -- their probability of death is lower by .10. The infants of craftsmen (OCF8-13) have a somewhat smaller but still significant advantage, while those in the unspecified and miscellaneous classes are not statistically different from the omitted class, farm workers' infants, in their mortality experience (Equation V.4).

Father's occupation and marital status of the mother are highly correlated -- whenever the mother is single, the father's occupation tends to be left unspecified. The results for father's occupation when entered separately have been discussed. Marital status of the mother is also significantly associated with mortality (Equation V.5). Infants of single mothers (MARM1) have much higher probabilities of death, and infants born into common-law marriages (MARM2) have somewhat higher probabilities,[1/] than do infants born to legally

---

1/ The coefficient falls just short of significance at the 95 percent level.

married women. When marital status was included in regressions with father's occupation the coefficient for MARM1 was much higher, about .35 as compared to the .22 in Equation V.5. Even the latter coefficient seems quite high and we suspect that the estimates are biased by the presence of multicollinearity, and possibly by the fewness of the observations in each class as well.[1/]

Two other characteristics of the equations that include marital status merit attention. As Equation V.5 illustrates, whenever marital status was included, the lowest birthweight class (BWT1,2 and 3) was no longer significant. And when both marital status and father's occupation were included, even the vaccination variable was affected by the problems of collinearity. Its coefficient dropped to nearly half its former value and was significant only at the 90 percent level.

The next section takes up the implications of this collinearity and considers what can be done.

Multicollinearity, A Hypothesis, and Directions for Further Research

The social factors used in these equations, and in other chapters of this report, present something of a conceptual problem. They are generally intended as proxies for a suspected causal factor, or factors, rather than as possible direct causes in their own right. For example, type of neighborhood and father's occupation are proxies for the conditions that accompany poverty

---

1/ We are particularly reluctant, however, to stress the fewness of observations in this case. If the distributions by marital status are, in reality, identical for living and dead infants, we could have expected to find 17 living infants with single mothers and 25 born to common-law unions, rather than the 3 and 9, respectively, that were actually surveyed (Table 2, chapter appendix). The numbers actually surveyed may be the product of sampling variability, or they may signal an underlying relationship. Only by analyzing a larger sample can we discover which is the case.

or affluence, while parental education is frequently treated as a proxy for knowledge of child care. But exactly because they are proxies, it is difficult to know whether a statistical association constitutes support for the hypothesized factor, or represents some other, as yet unsuspected, influence.

In this analysis, unlike the 2 of previous chapters, the social factors present a statistical problem as well. They are correlated with some of the factors which can be viewed as possible proximate causes of death -- that is, those that may have a direct link with mortality. This is the reason that the results for these proximate causes are unstable when certain of the social factors are entered in the regressions, a problem that did not arise in the analyses of Chapters II and IV.

The results for the variables representing the toilet facilities available to the family were particularly affected by this problem. When these variables were specified in regressions that included only the factors we believed could be proximate causes of death, their coefficients were of the expected sign and TLT2 (no toilet facilities) was statistically significant. When the specifications were expanded to include some of the social factors, such as type of neighborhood, the toilet variables became insignificant. Yet factors like type of neighborhood are not proximate causes of death; in fact, they stand in for just such factors as toilet facilities, that are expected to be proximate causes.

The significance of the social factors, and the increased R^2's they produced, suggest that they represent not only some of the factors for which measures are available, such as toilet facilities, but others for which no measures are available from the survey. We hypothesize that the proximate causes which the social variables represent are the true causes of mortality and that the results from the equations that exclude social factors are valid.

but we are not able to demonstrate this conclusively because the proximate causes and social factors move together in the observation set.

Basically, the problem is due to insufficient information -- the factors of interest do not show enough independent variation, in this set of observations, to permit their individual effects to be estimated. The econometric label for this condition is multicollinearity. When it is present, insignificant coefficients cannot be taken as evidence that the factor has no effect, only that the collinearity is too severe to permit it to be estimated. Factors with significant coefficients may also be affected. While collinearity does tend to produce insignificant t-scores,[1/] it does not guarantee them, and it also tends to bias the coefficient estimates (Johnston, 1972). Thus we may be misled about the true effect of a significant variable.[2/]

It is not accidental that this problem has appeared for the first time in an analysis based on a relatively small set of observations. In a small sample there is less chance for each of the factors of interest to exhibit enough independent variation. While total sample size is important, it seems likely that the observation set of infants 28 days to 5 months suffers particularly from the small number of living children in the group -- only 99.[3/; 4/]

1/ Because the standard errors of the coefficients become very large as the X'X matrix approaches singularity.

2/ As the discussion indicates collinearity makes it difficult to attribute the correct influence to each variable; but it does not prevent the affected variables from contributing to the explanatory power of the equation, as measured by the R^2.

3/ In the neonatal observation set, not only was the number of living children larger, but the comparison group was expanded by the inclusion of post-neonatal deaths. In the birthweight analysis, the observation set was also considerably larger, and the variation being analyzed was in birthweight rather than mortality so that infant deaths again contributed to a higher density of observations in the "comparison group" -- heavier birthweights.

4/ Part of the difficulty with small sample sizes is illustrated by our earlier discussion of the results for marital status of the mother. There are several reasons to question the validity of these coefficients, but we were reluctant to place much weight on one of them -- the very few living infants in the two classes of interest -- since, had the distributions been identical for the living and dead infants, we could have expected to find many more living infants in these groups. But only a larger sample would permit us to determine for sure whether the actual numbers reflect sampling variability or a valid relationship.

The solution to the problem of multicollinearity is to get more information. And it is clear from the preceding discussion that one way to do this is to expand the sample size. This could be done by combining two or three of the 13 Latin American project sites included in the Investigation. Even combining one other project site, of reasonably good size, to the San Juan survey, should produce a large enough observation set to insure independent variation in each of the factors.

In general, the results of the analysis presented in this chapter are good. It is clearly possible to draw useful and meaningful generalizations from the data. But because of the multicollinearity, some of the results are not as stable and as soundly based as is desirable, and we would like to have more confidence in their statistical basis. We believe that this confidence can be provided by an analysis of several of the project sites combined and that the most profitable direction for further research based on these surveys is to expand the analysis of mortality among infants 28 days to 5 months to a larger set of observations.[1/]

1/ We realize that combining several projects can present problems of data comparability. But we believe that, if care is taken in combining and analyzing several projects, these problems will be outweighed by the benefits.

APPENDIX TO CHAPTER V
Distribution of Observations 28 Days to 5 Months of Age by Factors Used in the Regressions

1. By Birthweight (in hundreds of grams)

	Deaths	Sample
Less than 24	92	4
25-29	147	13
30-34	164	36
35-39	163	27
Over 40	43	16
Not specified	80	3
TOTAL	689	99

2. By Length of Pregnancy

	Deaths	Sample
Less than nine months	126	7
Nine months	538	91
Not specified	25	1
TOTAL	689	99

3. By Whether Infant Was Breastfed and Age Weaning Started

	Deaths	Sample
Breastfed, weaning started at less than 1 month	205	28
Breastfed, weaning started at 1 month	143	14
Breastfed, weaning started at 2 months or more	119	16
Breastfed, never weaned, or age of weaning unknown, or not known whether weaned	109	33
Never breastfed	107	8
Not specified	6	0
TOTAL	689	99

APPENDIX TO CHAPTER V, continued

4. By Sex of Child

	Deaths	Sample
Female	355	43
Male	334	56
TOTAL	689	99

5. By Whether or Not the Child Received Any Vaccinations

	Deaths	Sample
Yes	352	75
No	310	24
Not specified	27	0
TOTAL	689	99

6. By Type of Neighborhood

	Deaths	Sample
Well-to-do, good, or medium	140	36
Poor	112	30
Very Poor	52	3
No neighborhood or rural	383	29
Not specified	2	1
TOTAL	689	99

7. Number of Persons in the Household

	Deaths	Sample
1 or 2	80	0
3 or 4	216	44
5 or 6	154	32
7 or 8	126	17
9 or more	113	6
Not specified	0	0
TOTAL	689	99

APPENDIX TO CHAPTER V, continued

8. By Type of Water Facility

	Deaths	Sample
Piped water inside	98	37
Piped water outside	76	8
Other	514	53
Not specified	1	1
TOTAL	689	99

9. By Type of Toilet Facilities

	Deaths	Sample
Flush or communal	80	33
Other	182	31
None	425	34
Not specified	2	1
TOTAL	689	99

10. By Number of Rooms in the Dwelling

	Deaths	Sample
1 room	137	7
2 rooms	179	20
3 or 4 rooms	283	52
5 or more rooms	82	19
Not specified	8	1
TOTAL	689	99

11. By Marital Status of Mother

	Deaths	Sample
Single	115	3
Married	400	87
Common-law or other	171	9
Not specified	3	0
TOTAL	689	99

APPENDIX TO CHAPTER V, continued

12. By Occupation of the Father

	Deaths	Sample
Farm workers	292	26
Professionals	25	15
Craftsmen	153	32
All other	132	20
Not specified	87	6
TOTAL	689	99

CHAPTER VI. Determinants of Mortality in Children Six Months or Older

Introduction

In the San Juan project, most of the deaths of children under one year -- fully 86 percent -- occurred before the infant reached six months of age. As Table VI.1 shows, there were 1,641 deaths among infants 0-5 months and only 259 among older infants during the two year survey period. Deaths among one-year-olds totalled 153, out of a population of about 21 thousand. There were 55 deaths of two-year-olds, 32 of three-year-olds, and 16 of four-year-olds.

This chapter takes up the analysis of mortality among these older infants and children. Two age groups were analyzed separately -- infants 6 to 11 months and children 1 through 4 years.

The set of observations defined for the analysis of infants 6 to 11 months includes all infants in the survey between 6 and 11 months at death or interview, a total of 247 deaths and 85 living infants (Table VI.1). The set defined for the analysis of 1 through 4-year-olds includes all children 1 through 4 years of age at death or interview, a total of 247 deaths and 674 living children (Table VI.1). As in Chapter V, these definitions correspond to cohort death rates, which compare deaths in a certain age group with the number of living children in the same age group. The weights used for the observations are set out in Appendix C.

TABLE VI.1. San Juan: Population and Number of Survey Observations Retained for Analysis (for two years) by Age Group

	Deaths		Living Children	
	Population	Observations Retained for Analysis	Population	Observations Retained for Analysis
0-5 months	1,641	1,495	20,280	119
6-11 months	259	247		85
1 year	153	145	21,040	203
2 years	55	55	63,140	156
3 years	32	32		159
4 years	16	15		156

Source: Population figures from Puffer and Serrano, 1973, p. 9.

The literature review surveyed a large number of factors that are believed to be associated with the survival of a child 6 months or older (Russell and Burke, 1974, Part One). They are the same factors that are hypothesized to be important to mortality among younger infants, but social and environmental factors are presumed to be primary during this period, and measures of prenatal history such as birthweight and gestation period are presumed to be insignificant. The child's diet, a measure of its nutritional status, is expected to be important. Other factors include: the medical care given the child, including whether or not it received any vaccinations, and its disease history; the sex of the child; and a variety of social and environmental conditions such as the type of water and toilet facilities available to a family, family size, father's occupation, marital status of the mother, number of rooms in the dwelling, and type of neighborhood.

A series of regressions were specified to examine the importance of these factors for each of the two age groups -- 6 to 11 months and 1 to 4 years. The results are presented in Equations VI.1-VI.3. Equation VI.1 is the basic equation for the 6 to 11 month observation set; like the basic equation in Chapter V, it contains only those factors that can be viewed as proximate causes of death. Equation VI.2 adds a number of social variables to the basic specification. Equation VI.3 is a typical equation for the 1 through 4 years age group.

The dependent variable for this analysis is a binary which takes on the value 1 (one) for deaths and 0 (zero) for living children. As in earlier chapters, it is interpreted as the probability of death during the period specified by the observation set -- either 6 to 11 months or 1 to 4 years. Likewise, as before, the independent variables are in binary form, and the coefficient for each variable gives the difference in the probability of death associated

with membership in the class represented by the variable, as compared to membership in the omitted class.[1/] Definitions of the classes represented by the independent variables are given in Appendix B.

In general, the results of the regressions for both observation sets were disappointing. The ability to generalize across children and to identify determinants at high confidence levels does not exist for either set. Even if we relax our requirements for statistical significance, as we have done here, the results provide little useful information. And, of course, the lower confidence levels make this information more tentative. Further, less of the variation in survival experience is explained: the coefficients of determination (R^2's) are .026 and .066 respectively in Equations VI.1 and VI.2 -- the equations for the 6 to 11 month children, and .003 in Equation VI.3 -- the equation for the 1 through 4 year children.

In both cases we suspect that the difficulties are due to the small numbers of deaths involved. The analysis of mortality among infants 28 days to 5 months (Chapter V) has already shown that relatively small sample sizes reduce our ability to draw the appropriate generalizations out of the myriad factors operating in each case.[2/] The number of living children in the observation sets used here are either similar (6 to 11 months), or far greater (1 to 4 years) than in that earlier analysis, but the numbers of deaths are much smaller -- 247 in each case (Table VI.1).

For the 1 to 4-year-olds the problem may ultimately stem from the fact that there are so few deaths in the population. The 247 deaths, essentially

---

1/ For fuller discussions of the definition and interpretation of binary variables, see Appendix B and Chapter II.

2/ Note that this is not per se a problem of having sampled too small a fraction of the population -- the fraction sampled is the same for all age groups in the study, and the population of deaths was surveyed. It is instead a matter of the ability of the resulting observation set to reflect the full range of variation that occurs in each of the factors of interest.

all those that occurred during the 2 year survey period (Table VI.1), are aggregated over such a wide age range that the causes of death may differ considerably. Within each year of age the deaths are so few as to suggest that generalization may be difficult. Thus even if enough projects could be combined to produce a suitable sample size, a useful analysis of this age group may not be possible.

But the results for the 6 to 11-month-olds, while not good enough to offer guides for policy, do suggest that it is possible to generalize about the causes of death for this group. There are a sizeable number of deaths in this narrow age range; expanding the observation set to include several project sites would produce a large enough sample to support a useful analysis.

The following sections will discuss the results in a little more detail. First, the results for infants 6 to 11 months are presented, then the results for children 1 through 4 years.

The Regressions: Infants 6 to 11 Months

On the basis of the literature review we specified regressions for this observation set, all infants 6 to 11 months, which included most of the same factors tested in the regressions for infants 28 days to 5 months. The structure of the regressions specified was also similar. We first tested those factors that seemed likely to be proximate causes of death (Equation VI.1). Then we added a variety of social variables, which serve as proxies for more direct causes of death, to the basic specifications (e.g., Equation VI.2).

The results were faint echoes of those for the 28 days to 5 months age group. In Equation VI.1, for example, the coefficients are generally of the same sign as the counterparts in Chapter V, but not one is significant at even the 80 percent level (although VACC and CAL2 fall just short). In Equation

EQUATION VI.1: Basic Equation for 6 to 11-Month-Olds

Dependent Variable: Deaths 6 to 11 months = 1

Set of Observations: Records of all infants 6 to 11 months of age.

Name of Factor	Name of Variables	Coefficient	t-score	F for Factor (see note)
Sex of Child	SEX	-0.005	0.29	0.084
Type of Water Facilities	WTR1	0.006	0.16	0.10
	WTR2	-0.009	0.21	
Type of Toilet Facilities	TLT1	0.026	0.61	0.60
	TLT2	0.044	0.98	
Whether Child Received Any Vaccinations	VACC	-0.048	1.17	1.37
Whether Child Was Breastfed	BRF1,2,3,4,5 and 6	-0.015	0.39	0.15
Adequacy of Diet in Terms of Calories	CAL2	0.023	1.26	0.84
	CAL3	0.026	0.60	
Constant Term	-	0.043	-	-

R^2 for the equation: 0.026
F for the equation: 0.94
a. significant at the 95 percent level of confidence.
c. significant at the 80 percent level of confidence.

NOTE: Each F-value indicates whether the factor as a whole is statistically significant. When a factor is represented by a single variable the F is simply the square of the t-score. When a factor is represented by more than one variable it turns out that, in general, the factor is significant whenever at least one of the variables has a significant t-score. Limitations on computer time made it impossible to calculate these F's for equations other than the basic equation in each chapter but the reader can tell when a factor is significant by examining the t-scores.

EQUATION VI.2: Basic Equation for 6 to 11-Month-Olds With Social Variables

Dependent Variable: Deaths 6 to 11 Months = 1

Set of Observations: Records of all infants 6 to 11 months of age.

Name of Factor	Name of Variables	Coefficient	t-score
Sex of Child	SEX	-0.019	0.93
Type of Water Facilities	WTR1	0.030	0.63
	WTR2	-0.021	0.41
Type of Toilet Facilities	TLT1	0.019	0.40
	TLT2	0.040	0.83
Whether Child Received Any Vaccinations	VACC	-0.063	1.41[c]
Whether Child Was Breastfed	BRF1,2,3,4,5 and 6	-0.025	0.62
Adequacy of Diet in Terms of Calories	CAL2	0.022	1.03
	CAL3	0.029	0.58
Type of Neighborhood	NBR1	-0.038	1.01
	NBR2 and 3	0.021	0.63
	NBR4	0.016	0.51
Number of Rooms in Dwelling	RMS1	-0.054	1.57[c]
	RMS2	-0.025	0.81
	RMS3	0.002	0.08
Occupation of Father	OCF1	-0.121	2.28[a]
	OCF2,3 and 4	-0.068	1.72[c]
	OCF8,9,10,11,12 and 13	-0.038	1.22
	OCF5,6,7,14 and 15	-0.058	1.73[c]
Number of Persons in Household	FAM1	-0.016	0.57
	FAM2	-0.004	0.14
	FAM3	-0.011	0.32
	FAM4	0.033	0.82
Marital Status of Mother	MARM1	0.165	2.28[a]
	MARM2	0.016	0.55
Constant Term		0.077	

R^2 for the equation: 0.066
F for the equation: 0.86
a. significant at the 95 percent level of confidence.
b. significant at the 99 percent level of confidence.
c. significant at the 80 percent level of confidence.

EQUATION [illegible] Basic Equation [illegible]

Dependent Variable: Deaths 1 through 4 years of Age 1

Set of Observations: Records of all children 1 through 4 years of age.

Name of Factor	Number of Variables	Coefficient	t-[illegible]
Whether or Not Child Was Breastfed and Age Weaning Started	BRF1	0.0004	0.06
	BRF2	0.002	0.20
	BRF3	0.0008	0.09
	BRF4	-0.0002	0.02
	BRF5	0.0005	0.07
	BRF6	-0.002	0.17
Sex of Child	SEX	0.0007	0.19
Whether Child Received Any Vaccinations	VACC	-0.005	0.49
Child's History of Infectious Diseases	DIS1	-0.002	0.22
	DIS2	0.00005	0.01
Person Responsible for Care of the Child	CARE1	-0.001	0.21
Type of Neighborhood	NBR1	-0.0007	0.11
	NBR2	0.002	0.30
	NBR3	0.006	0.41
	NBR4	0.003	0.56
Number of Persons in Household	FAM2	0.001	0.31
	FAM3	0.002	0.38
	FAM4	0.003	0.52
Education of Father	EDF1	-0.003	0.42
	EDF2	0.001	0.19
	EDF3	0.0009	0.12
Type of Water Facilities	WTR1	0.001	0.12
	WTR2	-0.0005	0.06
Type of Toilet Facilities	TLT1	-0.0008	0.11
	TLT2	0.004	0.49
Constant Term		0.006	

R^2 for the equation: 0.003
F for the equation: 0.12
a. significant at the 95 percent level of confidence.
c. significant at the 80 percent level of confidence.

VI.2, the results improve a little with the addition of the social variables. The vaccination variable shows that infants who have been vaccinated have a lower probability of death during this period than those who have not. The value of the coefficient is close to the estimates for the younger group in Chapter V and is significant at the 80 percent level. As before, infants living in one room dwellings (RMS1) have a significantly higher probability of death. The occupation variables indicate that infants of farm workers, the omitted group, continue to face a higher risk of death in this age group. And the marital status of the mother again shows an association with mortality, with infants born to single mothers less likely to survive.

As we noted in our earlier discussion, these results indicate that it is possible to generalize about the causes of death in this age group. But the sample size is too small to allow the generalizations to emerge clearly.[1/] The larger sample that could be generated from several project sites combined would permit a more definitive analysis.

The Regressions: Children 1 through 4 Years

We specified regressions for this observation set -- all children 1 through 4 years -- which included most of the same factors included in the regressions for the 6 to 11 month children. The nutrition measures were not included because the survey data would not allow us to develop such measures.[2/] The measures

---

1/ See the discussion in Chapter V of the problems that arise from inadequate sample size.

2/ See Chapter II for further explanation of why nutritional measures could not be developed.

that were tested are: whether or not child was breastfed, sex of child, whether child received any vaccinations, child's history of infectious diseases, person responsible for care of child, type of neighborhood, education of father, number of persons in household, type of water facilities, and type of toilet facilities. Equation VI.3 is typical of the results. None of the factors in this regression or any of the other regressions specified were significant even at the 80 percent level of confidence. As noted earlier, this undoubtedly reflects the very small number of deaths -- a total of 247 were retained for analysis -- in this age group.

APPENDIX TO CHAPTER VI.
Distribution of Observations 6 to 11 Months of Age by Factors Used in the Regressions

1. By Whether or Not the Child Received Any Vaccinations

	Deaths	Sample
Yes	196	81
No	44	4
Not specified	7	0
TOTAL	247	85

2. By Type of Neighborhood

	Deaths	Sample
Well-to-do, good or medium	43	48
Poor	44	12
Very poor	27	1
Rural or no neighborhood	131	24
Not specified	2	0
TOTAL	247	85

3. By Number of Rooms in the Dwelling

	Deaths	Sample
1 room	31	13
2 rooms	66	14
3 or 4 rooms	119	34
5 or more rooms	29	24
Not specified	2	0
TOTAL	247	85

4. By Marital Status of the Mother

	Deaths	Sample
Single	31	2
Married	154	70
Common-law or other	61	13
Not specified	1	0
TOTAL	247	85

5. By Occupation of the Father

	Deaths	Sample
Professionals	10	14
Farm workers	107	19
Craftsmen	56	27
Others	52	19
Not specified	22	6
TOTAL	247	85

6. By Type of Toilet Facilities

	Deaths	Sample
Flush or communal	15	34
Other	85	25
None	145	26
Not specified	2	0
TOTAL	247	85

7. By Sex of Child	Deaths	Sample
Female	127	43
Male	120	42
TOTAL	247	85

8. By Whether or Not Child Was Breastfed and Age Weaning Started

	Deaths	Sample
Breastfed, weaning started at less than 1 month	39	18
Breastfed, weaning started at 1 month	34	11
Breastfed, weaning started at 2 months or more	121	45
Breastfed, never weaned or age of weaning unknown, or not known whether weaned	31	6
Never breastfed	18	5
Not specified	4	0
TOTAL	247	85

APPENDIX TO CHAPTER VI, continued

9. By Type of Water Facilities

	Deaths	Sample
Piped water inside	25	33
Piped water outside	37	9
Other	184	43
Not specified	1	0
TOTAL	247	85

10. By Number of Persons in Household

	Deaths	Sample
1 or 2	20	0
3 or 4	76	30
5 or 6	74	29
7 or 8	46	15
9 or more	31	11
Not specified	0	0
TOTAL	247	85

11. By Marital Status of Mother

	Deaths	Sample
Single	31	2
Married	154	70
Common-law or other	61	13
Not specified	1	0
TOTAL	247	85

CHAPTER VII. Analysis of a Major Cause of Death: Diarrheal Disease

Introduction

The Inter-American Investigation of Mortality in Childhood paid particular attention to gathering and organizing information on the causes of each death in the survey. Since multiple causes are involved in any death, a hierarchy was established among them. The major underlying cause of each death was first selected -- "the disease or injury which initiated the train of morbid events leading directly to death." Two types of associated causes were recognized. Consequential causes were those conditions "triggered by the underlying cause." Contributory causes included all other conditions which "contributed to the fatal outcome, but which...(were)...not related to the disease or condition directly causing death" (Puffer and Serrano, 1973, p. 74).

The analysis in Chapters IV, V, and VI is based on the assumption that factors like low birthweight, poor nutrition, and poor sanitary and living conditions predispose a child to illness and death, without reference to the type of illness, and that these conditions are truly the underlying causes of death. But it also seems likely that certain of these conditions may be more closely associated with particular illnesses than others and analysis of a single illness, or related group of illnesses, may be able to identify these relationships. Where the cause of death is a major one these more precise results

can help to focus preventive efforts. This chapter takes up the analysis of a major cause of infant mortality: diarrheal disease.

Diarrheal disease was far and away the most important cause of death uncovered by the Investigation. For all the survey sites together, diarrheal disease was the underlying cause of death for 28.6 percent of the deaths of children under five; it was an associated cause in another 11.2 percent of the cases (Puffer and Serrano, 1973, pp. 138-141). The proportions were similar in San Juan -- it was an underlying cause in 25 percent of all deaths under five and an associated cause in another 12 percent (Table VII.1).

The large number of deaths involved makes diarrheal disease a good single cause to analyze.[1/] And in this first attempt at investigating a single cause of death we chose to use the largest possible number of observations, within one of the age groups already defined (and successfully analyzed). Thus we selected the 28 days to 5 months group and used all deaths for which diarrheal disease was <u>either</u> an underlying or an associated cause of death -- a total of 363 deaths (Table VII.1). This choice should be kept in mind, as one refinement of the analysis might be to concentrate on those infants for which it was an underlying cause, omitting those for whom it was only an associated cause.

The full set of observations was then defined, as in Chapters V and VI, to correspond to a cohort death rate. For a specific cause, a cohort rate compares persons who died from that cause with living persons in the same age group. Thus the observation set is composed of the 363 diarrhea deaths and the 99 living children between the ages of 28 days and 5 months.[2/]

1/ Two other categories that might be investigated if this approach were extended are "pneumonia and influenza," and "certain perinatal causes," both of which are underlying or associated causes in a large number of cases (Puffer and Serrano, 1973, p. 378).

2/ See Appendix C for the weights applied to these observations.

TABLE VII.1. Deaths from All Causes and from Diarrheal Disease, by Age[1/]

	All Deaths	Deaths with Diarrheal Disease as: Any Cause[2/]		Underlying Cause	
		Number	Percent	Number	Percent
Under 7 days	509	18	3.5	10	2.0
7-27 days	297	160	53.9	107	36.0
28 days to 5 months	689	363	52.7	253	36.7
6-11 months	247	148	59.9	101	40.9
1-4 years	247	53	21.5	26	10.5
TOTAL	1,989	742	37.3	497	25.0

1/ Tabulation includes only those deaths retained for the regression analysis.
2/ All cases with diarrheal disease as either an underlying or associated cause.

The dependent variable defined for the regressions takes on the value 1 (one) for a death and 0 (zero) for a living child. As before, estimates of the dependent variable derived from an equation are interpreted as the probability of death for children with the characteristics specified. And the coefficients for the independent variables, all of which are binaries, show the differential in the probability of death associated with membership in the class represented by the variable, as compared to membership in the omitted class for the factor.[1/]

The types of factors expected to be related to death from diarrheal disease are of the same sort as those expected to be related to general mortality in this age group. While the literature review (Russell and Burke, 1974, Part One) suggested one or two factors which might show a stronger relationship with diarrheal deaths than with mortality in general -- specifically nutrition and sanitation -- it gave no reason to exclude any of the factors considered in Chapter V. Thus we tested the same specifications here. The factors tested include: birthweight; length of pregnancy; the medical care given the child and its disease history; the child's diet and history of breastfeeding; the sex of the child; and social and environmental conditions, including toilet and water facilities available to the family, family size, father's occupation, housing, type of neighborhood, and marital status of the mother.

1/ Definitions for all the variables shown are listed in Appendix B. The summed variables shown in the equations denote variables which combine the classes indicated by the components of the sum. Distributions of the factors over the classes represented by the variables are given in the chapter appendix.

The striking thing about the results, shown in Equations VII.1-VII.5,[1/] is that they are simply a replay of those for all infants 28 days to 5 months -- with the important difference that the levels of significance and R^2 are lower. Thus the isolation of one cause of death has not succeeded in producing sharper or more specific results. At the same time the analysis provides a very clear example of the effects of sample size -- these equations are based on the same number of living children as were the regressions for all infants 28 days to 5 months, but only about half as many deaths (363 vs. 689).

Because the results are so similar to those in Chapter V, they will be only briefly reviewed in the following section.

The Regression Results

Infants weighing 2400 grams or less at birth face a higher risk of death from diarrheal disease than heavier infants. The coefficient for EWT1,2 and 3 is significant, but a little lower in value than it was in Chapter V; it also shows the same sensitivity to the addition of variables representing the marital status of the mother (Equation VII.5). The remaining birthweight coefficients show, as before, a tendency for the risk of death to decline for heavier infants, but none are significant at the prescribed levels.

Length of pregnancy is reasonably stable in value and always significant at the 90 percent level or better. The coefficient indicates that the probability of death for premature babies is .07-.09 higher than for full-term babies; this differential is below the estimates of Chapter V, which averaged about .12. The sex of the child is significant less often than in Chapter V, but the coefficients indicate that female infants have a lower risk of death.

1/ These are the same specifications shown in Equations V.1-V.5, Chapter V.

EQUATION VII.1: Basic Equation

Dependent Variable: Diarrhea Deaths 28 Days to 5 Months = 1

Set of Observations: Records of all live infants 28 days to 5 months of age, and dead infants 28 days to 5 months of age for whom diarrheal disease was any cause of death.

Name of Factor	Name of Variable	Coefficient	t-score	F for Fac (see note
Birthweight	BWT1,2 and 3	0.133	2.76[b]	2.85[a]
	BWT4	0.048	1.44	
	BWT5	0.010	0.36	
	BWT6 and 7	-0.020	0.63	
	BWT8	0.109	1.98[a]	
Length of Pregnancy	PLG1,2 and 3	0.074	2.04[a]	4.16[a]
Whether Child Was Breastfed	BRF1,2,3,4,5 and 6	-0.080	2.19[a]	4.80[a]
Sex of Child	SEX	-0.034	1.67	2.79
Whether Child Received Any Vaccinations	VACC	-0.060	2.64[b]	6.97[b]
Type of Water Facilities	WTR1	0.023	0.45	0.11
	WTR2	0.007	0.16	
Type of Toilet Facilities	TLT1	0.025	0.60	1.04
	TLT2	0.060	1.22	
Number of Rooms in Dwelling	RMS1	0.120	3.23[b]	-
	RMS2	0.024	0.95	
	RMS3	0.028	1.02	
Constant Term	-	0.035	-	-

R^2 for the equation: 0.105
F for the equation: 3.27
a. significant at the 95 percent level of confidence.
b. significant at the 99 percent level of confidence.

NOTE: Each F-value indicates whether the factor as a whole is statistically significant. When a factor is represented by a single variable the F is simply the square of the t-score. When a factor is represented by more than one variable it turns out that, in general, the factor is significant whenever at least one of the variables has a significant t-score. Limitations on computer time made it impossible to calculate these F's for equations other than the basic equation in each chapter but the reader can tell when a factor is significant by examining the t-scores.

EQUATION VII.2: Basic Equation with Alternative Variables for Breastfeeding History

Dependent Variable: Diarrhea Deaths 28 Days to 5 Months = 1

Set of Observations: Records of all live infants 28 days to 5 months of age, and dead infants 28 days to 5 months of age for whom diarrheal disease was any cause of death.

Name of Factor	Name of Variables	Coefficient	t-score
Birthweight	BWT1,2 and 3	0.110	2.17[a]
	BWT4	0.044	1.31
	BWT5	-0.008	0.29
	BWT6 and 7	-0.032	1.00
	BWT8	0.104	1.89
Length of Pregnancy	PLG1,2 and 3	0.068	1.87
Whether Child Was Breastfed	BRF1	-0.066	1.71
	BRF2	-0.035	0.80
	BRF3,4 and 5	-0.054	1.26
	BRF6	-0.130	3.24[b]
Sex of Child	SEX	-0.025	1.20
Whether Child Received Any Vaccinations	VACC	-0.059	2.59[b]
Type of Water Facilities	WTR1	0.003	0.06
	WTR2	-0.002	0.04
Type of Toilet Facilities	TLT1	0.061	1.41
	TLT2	0.099	1.97[a]
Number of Rooms in Dwelling	RMS1	0.133	3.61[b]
	RMS2	0.011	0.43
	RMS3	0.025	0.90
Constant Term		0.089	

R^2 for the equation: .128
F for the equation: 3.43
a. significant at the 95 percent level of confidence.
b. significant at the 99 percent level of confidence.

EQUATION VIII.3: Basic Equation with Measures of Diet Adequacy (and Two Social Variables)

Dependent Variable: Diarrhea Deaths 28 Days to 5 months = 1

Set of Observations: Records of all live infants 28 days to 5 months of age, and dead infants 28 days to 5 months of age for when diarrheal disease was any cause of death.

Name of Factor	Name of Variables	Coefficient	t-score
Sex of Child	SEX	-0.040	1.83
Birthweight	BWT1,2 and 3	0.088	1.76
	BWT4	0.033	0.95
	BWT5	-0.006	0.22
	BWT6 and 7	-0.038	1.15
	BWT8	0.069	1.21
Length of Pregnancy	PLG1,2 and 3	0.094	2.54[a]
Whether Child Was Breastfed	BRF1,2,3,4,5 and 6	-0.061	1.62
Whether Child Received Any Vaccinations	VACC	-0.048	2.03[a]
Number of Persons in Household	FAM1	0.027	0.86
	FAM2	0.004	0.16
	FAM3	0.090	2.35[a]
	FAM4	0.080	1.68
Type of Neighborhood	NBR2	0.008	0.29
	NBR3	0.131	2.13[a]
	NBR4	0.035	1.26
Type of Water Facilities	WTR1	-0.010	0.19
	WTR2	-0.005	0.11
Type of Toilet Facilities	TLT1	0.031	0.64
	TLT2	0.060	1.11
Adequacy of Diet in Terms of Calories	CAL1	0.047	1.22
	CAL2	-0.009	0.27
	CAL3	-0.046	1.76
Constant Term		0.065	

R^2 for the equation: 0.128
F for the equation: 2.80
a. significant at the 95 percent level of confidence.
b. significant at the 99 percent level of confidence.

EQUATION VII.4: Basic Equation with Social Variables Including Father's Occupation

Dependent Variable: Diarrhea Deaths 28 Days to 5 Months = 1

Set of Observations: Records of all live infants 28 days to 5 months of age, and dead infants 28 days to 5 months of age for whom diarrheal disease was any cause of death.

Name of Factor	Name of Variables	Coefficient	t-score
Sex of Child	SEX	-0.055	2.48[a]
Birthweight	BWT1,2 and 3	0.096	1.75
	BWT4	0.042	1.23
	BWT5	-0.006	0.22
	BWT6 and 7	-0.022	0.67
	BWT8	0.067	1.19
Length of Pregnancy	PLG1,2 and 3	0.073	1.98[a]
Whether Child was Breastfed	BRF1,2,3,4,5 and 6	-0.097	2.59[b]
Whether Child Received Any Vaccinations	VACC	-0.048	2.02[a]
Number of Persons in Household	FAM1	0.009	0.30
	FAM2	0.004	0.14
	FAM3	0.075	1.97[a]
	FAM4	0.076	1.64
Number of Rooms in Dwelling	RMS1	0.135	3.46[b]
	RMS2	0.046	1.65
	RMS3	0.011	0.38
Type of Neighborhood	NBR2	0.006	0.21
	NBR3	0.127	2.01[a]
	NBR4	0.066	2.29[a]
Type of Water Facilities	WTR1	0.003	0.05
	WTR2	-0.007	0.14
Type of Toilet Facilities	TLT1	-0.030	0.62
	TLT2	0.011	0.20
Occupation of Father	OCF1	0.003	0.07
	OCF2,3 and 4	-0.072	1.85
	OCF5,6,7,14 and 15	-0.030	0.92
	OCF8,9,10,11,12 and 13	-0.044	1.52
Constant Term		0.072	

R^2 for the equation: 0.147

F for the equation: 2.78

a. significant at the 95 percent level of confidence.

b. significant at the 99 percent level of confidence.

EQUATION VII.5: Basic Equation with Social Variables, Including Marital Status of Mother

Dependent Variable: Diarrhea Deaths 28 Days to 5 Months = 1

Set of Observations: Records of all live infants 28 days to 5 months of age, and dead infants 28 days to 5 months of age for whom diarrheal disease was any cause of death.

Name of Factor	Name of Variables	Coefficient	t-score
Sex of Child	SEX	-0.044	2.00a
Birthweight	BWT1,2 and 3	0.038	0.67
	BWT4	0.040	1.21
	BWT5	-0.007	0.26
	BWT6 and 7	-0.022	0.68
	BWT8	0.075	1.34
Length of Pregnancy	PLG1,2 and 3	0.067	1.81
Whether Child Was Breastfed	BRF1,2,3,4,5 and 6	-0.077	2.09[a]
Whether Child Received Any Vaccinations	VACC	-0.044	1.88
Number of Persons in Household	FAM1	0.023	0.73
	FAM2	0.004	0.17
	FAM3	0.080	2.15[a]
	FAM4	0.076	1.64
Type of Neighborhood	NBR2	0.009	0.30
	NBR3	0.081	1.27
	NBR4	0.062	2.22[a]
Type of Water Facilities	WTR1	-0.007	0.14
	WTR2	-0.007	0.16
Type of Toilet Facilities	TLT1	-0.016	0.35
	TLT2	0.023	0.42
Number of Rooms in Dwelling	RMS1	0.144	3.77[b]
	RMS2	0.031	1.13
	RMS3	0.019	0.68
Marital Status of Mother	MARM1	0.166	2.83[b]
	MARM2	0.061	1.70
Constant Term		0.042	

R^2 for the equation: 0.158
F for the equation: 3.29
a. significant at the 95 percent level of confidence.
b. significant at the 99 percent level of confidence.

Breastfed infants have a substantially better chance of escaping death from diarrheal disease than infants who are not breastfed, although the estimated differential is somewhat less than that for all deaths in Chapter V. The same sort of pattern emerges here when the breastfeeding variables distinguish among infants by the age at which weaning started: all breastfed infants have an advantage, with those never weaned or age unspecified (BRF6) having the greatest advantage, but three of the four coefficients are not significant at the 95 percent level (Equation VII.2).

The type of toilet facilities are less frequently significant than in Chapter V -- TLT2 (no facilities) is significantly associated with higher mortality only in Equation VII.2 -- and the variables show the same sensitivity to changes in the specification of the rest of the equation. The equations indicate that infants living in one room dwellings face significantly higher risks of death. The type of water facilities available to the family show no association with mortality.

Whether the infant received any vaccinations is consistently significant, as it was in Chapter V. But the estimate of the advantage held by vaccinated infants is lower, about .05-.06.

The social variables also replay the relationships reported in Chapter V. The pattern of differentials for father's occupation is unchanged -- children of professionals have a greater advantage over children of farm workers than do craftsmen's children -- but the coefficients are not significant at the prescribed levels (Equation VII.4). The probability of death is higher for infants in rural, or very poor urban, neighborhoods, but the estimates are unstable in value and significance. Infants in large families are at greater risk, although again the coefficients are smaller and not so strongly significant. The marital status of the mother is associated with mortality; as before, the disadvantage

of infants of single mothers seems suspiciously high (Equation VII.5). It may be one of the results most affected by collinearity in the data set.

The discussion in Chapter V of the difficulties resulting from collinearity apply here even more strongly. The reduced sample size has, in turn, reduced our ability to draw generalizations with any confidence. But the analysis provides no indications that the factors involved in death from diarrheal disease differ from those for general mortality in this age group.

APPENDIX TO CHAPTER VII.
Distribution of Observations 28 Days to 5 Months of Age
by Factors Used in the Regressions

1. By Birthweight (in hundreds of grams)

	Diarrhea Deaths	Sample
Less than 24	57	4
25-29	78	13
30-34	80	36
35-39	79	27
Over 40	18	16
Not specified	51	3
TOTAL	363	99

2. Length of Pregnancy

	Diarrhea Deaths	Sample
Less than nine months	68	7
Nine months	291	91
Not specified	4	1
TOTAL	363	99

3. By Whether Infant Was Breastfed and Age Weaning Started

	Diarrhea Deaths	Sample
Breastfed, weaning started at less than 1 month	109	28
Breastfed, weaning started at 1 month	91	14
Breastfed, weaning started at 2 months or more	123	16
Breastfed, never weaned, or age of weaning unknown, or not known whether weaned	37	33
Never breastfed	0	8
Not specified	3	0
TOTAL	363	99

APPENDIX TO CHAPTER VII, continued

4. By Sex of Child

	Diarrhea Deaths	Sample
Female	201	43
Male	162	56
TOTAL	363	99

5. By Whether or Not the Child Received Any Vaccinations

	Diarrhea Deaths	Sample
Yes	175	75
No	169	24
Not specified	19	0
TOTAL	363	99

6. By Type of Neighborhood

	Diarrhea Deaths	Sample
Well-to-do, good, or medium	67	36
Poor	53	30
Very poor	32	3
No neighborhood or rural	210	29
Not specified	1	1
TOTAL	363	99

7. Number of Persons in Household

	Diarrhea Deaths	Sample
1 or 2	42	0
3 or 4	112	44
5 or 6	75	32
7 or 8	66	17
9 or more	68	6
Not specified	0	0
TOTAL	363	99

APPENDIX TO CHAPTER VII, continued

8. By Type of Water Facility

	Diarrhea Deaths	Sample
Piped water inside	44	37
Piped water outside	33	8
Other	285	53
Not specified	1	1
TOTAL	363	99

9. By Type of Toilet Facility

	Diarrhea Deaths	Sample
Flush or communal	34	33
Other	100	31
None	228	34
Not specified	1	1
TOTAL	363	99

10. By Number of Rooms in Dwelling

	Diarrhea Deaths	Sample
1 room	84	7
2 rooms	104	20
3 or 4 rooms	133	52
5 or more rooms	40	19
Not specified	2	1
TOTAL	363	99

11. By Marital Status of Father

	Diarrhea Deaths	Sample
Single	69	3
Married	188	87
Common-law or other	105	9
Not specified	1	0
TOTAL	363	99

12. By Occupation of the Father

	Diarrhea Deaths	Sample
Farm workers	157	26
Professionals	10	15
Craftsmen	84	32
All Other	59	20
Not specified	53	6
TOTAL	363	99

CHAPTER VIII. Conclusions and Some Implications for Policy

Introduction

Infant mortality rates are high in many of the developing countries. In San Juan during the survey period 81 infants died before their first birthday for every thousand that were born alive. By contrast, there were only 4 or 5 deaths per thousand in the Canadian and United States project sites included in the Investigation (Puffer and Serrano, 1973, p. 61).

Simply for humanitarian reasons it is desirable to reduce this waste of life. But there is also evidence that infant and child mortality may be one of the links in the intricate chain leading to economic development. For example, many students of fertility believe that high infant mortality rates are one cause of the high birth rates in developing countries -- parents have additional children as protection against the risk of losing some of them. This suggests that reductions in mortality might help bring about lower fertility rates. As another example, it is suspected that high child mortality reduces parents' willingness to invest in their children's education and training, and this has repercussions on the amount of human capital embodied in the country's labor force.

Thus programs to combat infant and child mortality may stem from short term humanitarian motives alone, or they may be part of a strategy aimed at economic development over the longer term. In either case, there is a wide

range of possible programs to choose from, ranging from those specific to the problem to others for which reducing mortality may be only a secondary or indirect goal. The program planner, with limited resources at his command, is in need of information to help him decide which programs are likely to be effective and, of these, which will be most effective.

The Conclusions of the Study

One of the purposes of this study has been to lay some of the groundwork needed for these program decisions. The other has been to determine the usefulness of the data collected by the Investigation for this kind of analysis. This chapter will review the conclusions of the study and suggest the most profitable areas for further research. To help focus the discussion the chapter is organized by the types of programs that a program planner might want to consider. The results of the analysis are of interest for the following kinds of programs:

- Nutrition programs for pregnant women;
- Medical care for the mother during pregnancy and delivery;
- Nutrition programs for infants and children;
- Medical care for the child after birth;
- Sanitation and housing;
- Family planning programs.

The results indicate the benefits, in terms of reduced mortality, that can be expected from different programs. They do not provide information about program costs -- for example, the cost of vaccinating infants or of providing prenatal care to pregnant women. The planner will need to obtain this information, much of which is available from other sources, before he can decide which program or programs make the best use of his limited resources.

Most of the results that we discuss apply to infants under six months of age; 86 percent of the deaths during the first year of life occurred in this age group. The study was not able to tell us much about the determinants of mortality in older infants and children because so few deaths were involved. We will return to this point in the section on areas for further research.

Nutrition Programs for Pregnant Women. The link between maternal nutrition and infant mortality has two parts: the mother's nutritional state during pregnancy is an important influence on the birthweight of her baby; and the infant's birthweight is a major determinant of its chances for survival. A program planner interested in introducing a nutrition program for pregnant women, or considering such a program as one possibility in a strategy to reduce infant mortality, needs reasonably precise information about both parts of the relationship. How much difference does a given nutritional status or a specified improvement in that status make to the infant's weight at birth? Then, how much effect do differences in birthweight have on the infant's risk of death?

The study was not able to provide any measures of the relationship between maternal nutrition and birthweight. The survey on which it is based did not collect data on the nutritional status of the mother, her diet, or proxies for her nutrition such as her height and weight. We did test several social variables as proxies for maternal nutrition in the birthweight equations, in an attempt to control for its effect on birthweight (Chapter II). While the results were reasonable and suggested that the proxies were serving their intended purpose, they could not produce the kind of information needed.

The regressions have, however, measured quite precisely the overwhelming importance of birthweight for an infant's survival. The analysis of neonatal deaths in Chapter IV showed that low weight infants face very high risks of

death -- for those 1400 grams (3 pounds) or less, the probability of death in the first month averages about .70.[1/] For infants between 1500 and 1900 grams the probability is still very high, over .40. Those 2000 to 2400 grams, just below the conventional dividing line of 2500 grams (5.5 pounds) that separates low weight from normal weight infants, have about 1 chance in 10 of dying during their first month. For heavier infants the probabilities are much lower -- about .02.[2/]

Low birthweight infants continue to be at risk after the first month. The results reported in Chapter V show that infants less than 2400 grams at birth, who nevertheless manage to survive the first month, face a higher probability of death between 28 days and 5 months than do heavier infants. The estimated differential is, however, rather imprecise, ranging from an increase of .12 to .17.

Medical Care for the Mother During Pregnancy and Delivery. Like the mother's diet, prenatal care acts indirectly on infant mortality through its effect on birthweight. The analysis reported in Chapter II showed that infants whose mothers had some prenatal care were significantly heavier than infants born to mothers who had no care, and greater numbers of visits resulted in greater weight gains. Infants born to women with 3 to 5 visits weighed about 140 grams more than those whose mothers had no care. Infants whose mothers had six or more visits were 200 grams, about 7 ounces, heavier.

The relationship between birthweight and mortality, estimated in Chapter IV, allows us to translate a hypothetical program of prenatal care into lives saved.

1/ The reader should recall that probabilities cannot be greater than one (the event is certain to happen) or less than zero (the event is certain not to happen).

2/ These results are based on the binary variables defined for birthweight. A piecewise linear approximation was also fitted to show how the probability of death changes for small changes in birthweight (see Chart IV.2 in Chapter IV, and the accompanying discussion).

The neonatal analysis represented an estimated population of 43,373 infants.[1/] Of these, 4,770 were born to women who had no prenatal care. Let us suppose that a program had been instituted to give each of these women at least six visits for prenatal care.[2/]

Given the actual birthweights of these 4,770 infants, we can read the probability of death for each one from Chart IV.2.[3/] From these probabilities we can deduce that 163 of the infants could have been expected to die without the benefits of our proposed program. But if the program had provided each woman with the minimum six visits for care, each infant's birthweight would have increased by an average of 200 grams. To estimate the effect of the program, then, we add 200 grams to the recorded weight of each infant and read the new (lower) probability of death from Chart IV.2. The results of these calculations indicate that 41 of the infants could have been saved. Together with the number of visits required (6 for each of 4,770 women for a total of 28,620), the expected savings in lives gives the program planner a first approximation to the resources needed for the program and the results that can be expected.[4/]

1/ See Appendix C.

2/ As noted earlier, the content of the visits for women who had care is not known. This example is based on the assumption that the women in our hypothetical program receive the same (unspecified) services as women who had prenatal care.

3/ Using the piecewise linear representation in the chart rather than the step function.

4/ The analysis did not investigate the determinants of the gestation period (length of pregnancy), in large part because the quality of the data for length of pregnancy appears to be much poorer than the quality of the birthweight data. But it is important to note that prenatal care, maternal nutrition, and other determinants of birthweight may also influence length of pregnancy. Since length of pregnancy has an important independent effect on mortality, any program which influences both variables, as prenatal care may, will have a greater effect on mortality than is indicated simply by its effect on birthweight.

The findings with respect to medical care during delivery are of a negative sort -- they suggest what kinds of programs, or elements of programs, are not likely to reduce infant mortality. They show that delivery by a midwife, rather than by a physician, is not associated with any difference in mortality either directly or indirectly, through birthweight. Similarly, birth at home rather than in a hospital has no effect on birthweight or mortality (Chapters II and IV).[1/]

Nutrition Programs for Infants and Children. This study concurs with others in finding that breastfed infants have a better chance of survival than those who are not breastfed. For infants 28 days to 5 months of age, the probability of death during this period is about .09 lower if the infant has been breastfed (Chapter V).[2/] This differential implies that, for every 11 women who could be persuaded to breastfeed their infants, one additional infant would survive who would not otherwise have survived. It should be noted, however, that the large majority of infants in the San Juan survey -- about 90 percent -- were breastfed for at least some period of time.

We developed a nutritional classification system to summarize the detailed diet information requested for each child, in an attempt to measure the importance of nutrition more generally (Chapter III). The variables based on this system did not produce useful results in the analysis of infants 28 days to 5 months. For infants 6 to 11 months the nutrition variables performed as well as any others, which is to say that none of them showed very strong or stable relationships with mortality.

1/ This does not rule the possibility that the type of care at delivery affects maternal mortality.

2/ Our attempts to determine whether the age at which other foods were introduced made any difference were inconclusive -- the differentials that did appear were not subject to any obvious interpretation (see Chapter V). We could not investigate whether the age when weaning finished affected mortality because the information was lacking for too many infants.

Thus the study provides an estimate of the effect that breastfeeding has on the mortality experience of young infants. The estimate shows that breastfeeding is a significant factor and that programs to promote it could be beneficial.[1/] But we have been unable to contribute to existing knowledge about the relationship between infant mortality and nutrition in general.

Medical Care for the Child. There are many medical services that could be offered to infants with the aim of reducing mortality. The range of services includes check-ups, vaccinations, and hospital and physician care during illnesses. But extensive medical programs can be quite expensive, absorbing limited resources, and it is important to know whether they would be effective.

The analysis has shown that infants who are vaccinated have significantly better chances of survival than those who are not. For those 28 days to 5 months the probability of death was estimated to be lower by .07-.09 if the child was vaccinated (Chapter V). In terms of lives, these estimates imply that one infant could be saved who would otherwise have died, for every 11 to 14 who were vaccinated. While the results for infants 6 to 11 months cannot be given much weight, they suggest that vaccinations continue to reduce mortality in this age group (Chapter VI). Unfortunately, we were unable to identify the type, or types, of vaccination primarily responsible for this result because the survey was seldom able to collect this information.

Lack of information was also a problem for another aspect of medical care -- the effect of visits to a physician or clinic on mortality. Neither the number of well-child visits nor the total number of visits showed any association with mortality, but data were lacking for so many infants that this result must be considered inconclusive (Chapter V).

1/ Since lactating women require more nutrients than non-lact women, the costs of such a program would include not only promotional orts but the need to supplement the diets of poor women.

Sanitation and Housing. The development of safe water supplies and proper sewage treatment systems have been recognized methods of controlling disease for well over a century -- the urban agglomerations that accompanied the Industrial Revolution made the need for such measures especially clear. Lack of adequate sanitation measures is an important public health problem in developing countries.

The regressions indicated that the type of toilet facilities available to the family might have an effect on mortality among infants 28 days to 5 months. The differential associated with having no toilet facilities as compared with having a flush toilet is large, about .10, and often statistically significant. A smaller differential is associated with having a non-flush toilet, but this was seldom significant. Unfortunately, both variables were quite sensitive to the inclusion of social variables such as type of neighborhood and father's occupation, and it is thus impossible to be sure of the size or importance of these relationships (Chapter V). The source of the family's water supply -- piped inside, piped outside, or other -- did not show any association with infant mortality.

Extremely poor housing conditions are associated with higher mortality. An infant living in a one room dwelling had a probability of death higher by about .16 than an infant living in a three or four room house. This result did not seem to be reflecting the crowdedness of the dwelling since persons per room was tested and found insignificant.

Family Planning Programs. The analysis identified a number of relationships which suggest that family planning programs could, by focusing on high risk groups, help to reduce infant mortality. Many of these relationships are already well known and the present study simply adds to the accumulating evidence about them.

For example, very young mothers tend to have lower weight infants (Chapter II). Mothers under 17 have infants that average 406 grams, nearly a pound, less than those born to women 21 to 25. Mothers 17 to 20 years of age are still at a significant disadvantage, although the differential is much less -- about 135 grams, or just under 5 ounces. The very great importance of birthweight to the infant's survival has already been noted. This suggests that one of the aims of family planning programs might be to persuade very young women, particularly those under 17, of the value of postponing pregnancy in the interests of producing a healthy, living child.

The results for parity (number of pregnancies) do not offer any useful guides for family planning programs. The birthweight differential of first children is unavoidable if a woman plans to have children at all, although it offers another reason for postponing the first pregnancy until other conditions are favorable (Chapter II). However, mortality is higher when there are seven or more people (children and adults) in the household (Chapter V), suggesting that limitation of family size would help reduce infant mortality by reducing household size.

The spacing of pregnancies after the first is important. Infants conceived within three or four months of a previous birth are significantly lighter when born than those conceived after longer intervals. The results further suggest that the optimal interval, in terms of the effect on the infant's birthweight, is 19 to 24 months (Chapter II). Thus reasonably long intervals between births might be promoted through family planning programs, again on the grounds that they reduce mortality.

Finally, the marital status of the mother is also associated with mortality -- infants born to single mothers have a significantly higher risk of death -- and this suggests another obvious focus for family planning efforts.

Some of these relationships are frequently used as indicators of high risk cases once a woman is pregnant. From this point of view, they can help to guide the provision of prenatal care. To recapitulate then, the results confirm that the infants of women under 17, single women, and women pregnant within three or four months of a previous birth or pregnant for the first time, should be considered to be at greater than average risk.

Recommendations for Further Research

This study has shown that the data collected in San Juan, Argentina by the Inter-American Investigation of Mortality in Childhood, conducted by the Pan American Health Organization, are capable of supporting a useful regression analysis. The analyses of birthweight and neonatal mortality (Chapters II and IV) produced a number of clear, statistically robust, generalizations which were all the more convincing because of their agreement with earlier work. And the regression estimates provided rather precise estimates of these relationships, estimates which could be helpful to program planners.

The analysis of mortality among infants 28 days to 5 months, reported in Chapter V, similarly demonstrated that useful generalizations can be drawn from these data. But the estimates were less reliable than those of the earlier chapters because there was not enough independent variation in each of the explanatory factors. As we noted in Chapter V, this problem of collinearity stems in large part from the fact that the observation set for this group was considerably smaller than those used for the birthweight and neonatal mortality regressions; in a small sample there is less chance for each factor to vary independently of the others.

The importance of sample size was confirmed by the analysis of deaths from diarrheal disease in Chapter VII. This observation set coincided with the

one for all infants 28 days to 5 months, except that the number of infant deaths was cut nearly in half since only deaths from diarrheal disease were under study. The result was that the analysis identified all the same determinants as for the larger group, but at lower levels of confidence, and less of the variation in mortality was explained.

The analyses of mortality among older infants and children failed to produce useful results, largely because of the small numbers of deaths in these age groups. In the case of the 6 to 11-month-olds the regressions indicated that generalization was possible, but that the observation set was too small. For the 1 to 4-year-olds the number of deaths was so few, and the inability to generalize from this data set so complete, that it is not clear that a useful analysis of this age group is possible.

San Juan is only one of thirteen Latin American project sites included in the Investigation. In light of the results of this study, we believe that further econometric work based on these surveys could most profitably begin by using them to enlarge the set of observations for infants between 28 days and 5 months. The results of the analysis in Chapter V are promising. The sizeable number of deaths in this age group suggests that it is a likely target for programs to reduce infant mortality. By combining the data from one or two additional project sites to that for San Juan it should be possible to produce results that are more precise and more statistically robust, and therefore more reliable guides for policy, than those in Chapter V. We would rate this the most valuable addition to the present study.

We would recommend two other extensions of the present study although, in our opinion, they should be given lower priority. The first is simply the replication of the analyses of birthweight and neonatal mortality, using data from one or more additional project sites. The value of such replication

is that it would serve as a check on the generality of the results and the accuracy and stability of the estimated relationships. Neonatal deaths are such a large part of infant mortality that they must be a major target of attempts to reduce it, and it is obviously desirable that decisions about these programs be founded on the best possible information.

It would probably also be useful to generate a larger set of observations for the analysis of mortality among 6 to 11-month-olds. Deaths in this age group are still numerous, suggesting that it would be helpful to program planners to uncover the generalities that apply to this group. If three or four project sites were combined, the observation set should be large enough to produce these generalities.

Finally, we will just mention several areas in need of research that probably cannot be based on these surveys. The most obvious is the need for more precise information about the effects of maternal nutrition on birth-weight; the surveys cannot help here since no data on the mother's nutrition were collected. Length of pregnancy is so important to the infant's survival that it would be helpful to know more about its determinants; but the length of pregnancy information in the surveys does not seem to be of high enough quality to support such an analysis. And last, more work needs to be done to measure the relationship between infant diet and mortality.

APPENDICES

Appendix A. Selection of the Observations

The San Juan survey originally included 2,156 deaths and 891 living children. But the information collected for some of these records was so incomplete that they could not contribute tothe analysis, and these records were deleted from the file. Table A.1 shows, by age, the number of records originally included in the survey and the number retained for analysis.

The major reason for deleting records was lack of a home interview. This interview was only one of the sources of information about the infant deaths, but it collected most of the information of interest to the study; 161 deaths were dropped from the file because no interview information was available. The problem did not arise for the living children because the interview was the sole source of information about this group, and thus there were no living children without interviews in the original file.

As the second step of the selection process, the missing or invalid items were printed out for each of the remaining records. We then read through the print-out to identify records with too little information to be of use to the study. Very few records had to be deleted as a result of this edit -- only 6 of the deaths and 13 of the living sample.

Table A.1

Number of Records in Survey and Number Retained for Analysis, by Age

Age	Deaths		Sample	
	Number in survey	Number Retained	Number in survey	Number Retained
Under one year, total	1,900	1,742	206	204
0-27 days	906	806	21	20
28 days-1 month	163	150	23	23
two months	164	156	22	22
three months	164	151	16	16
four months	141	135	21	21
five months	103	97	17	17
six months	68	64	14	14
seven months	56	54	14	14
eight months	44	42	12	11
nine months	28	24	13	13
ten months	33	33	11	11
eleven months	30	30	22	22
One year	153	145	204	203
Two years	55	55	158	156
Three years	32	32	162	159
Four years	16	15	161	156
ALL AGES, TOTAL	2,156	1,989	891	878

Appendix B. Definition of Variables Used in the Regressions

The survey information recorded on LMHO's storage tape was described in detail in the feasibility report (see Appendix to Chapter III in Russell and Burke, 1974, Part One). This information had to be re-coded in variable forms appropriate for regression analysis. This appendix first describes the types of variables defined, and then lists in detail the full set of variables.

Much of the information available from the survey was qualitative in form -- for example, the character of the neighborhood (well-to-do, good, poor, very poor) has no natural numeric representation. One or more binary variables were defined to represent this type of information. A binary variable takes on the value 1 (one) for each observation that falls in the designated class (e.g., poor neighborhoods), and 0 (zero) for all other observations. If there are n possible classes, only n-1 binaries are defined. The nth class is uniquely defined by entries of zero for each of the n-1 binaries representing the other classes. The choice of classes to represent depended primarily on the number of classes available and the density of observations in each.

For factors that do have a natural numeric representation there are a number of ways to define variables for regression, and the advantage of such a large data base is that it permits experimentation with several different forms. Each of these factors was represented by three types of variables:

(1) the ordinary linear form, e.g., age of mother in years
(2) a binary, or set of binaries
(3) a series of variables that permit a piecewise linear approximation to the underlying functional form.

The variables for piecewise linear approximation are defined as follows. Let x be the normal numeric form of the variable, and a, b, and c be three "turning points" at which it is suspected that the slope of the relationship may change. Then three new variables are defined:

$x_a = 0$ ([illegible]), or [illegible], whichever is larger
$x_b = 0$ ([illegible]), or [illegible], whichever is larger
$x_c = 0$ ([illegible]), or [illegible], whichever is larger.

Note that a,b, and c are i[illegible] to the [illegible] of x, and divide it into four segments. When the original linear form of the variable, x, and the three new variables x_a, x_b, and x_c, are included in a regression, their coefficients -- call them B, B_a, B_b, and B_c -- can be combined to represent the slopes of the four segments. Moving from lower to higher values of x along the horizontal axis, these slopes are, respectively, B, $B + B_a$, $B + B_a + B_b$, and $B + B_a + B_b + B_c$. This functional form is continuous in the sense that the estimated segments meet at the turning points. It has the advantage, over the binary form, that each segment can take on any slope dictated by the data.

It was not possible, of course, to define the values of variables for which an observation was missing information. This problem was handled in two ways. In the few cases where information was missing for a large number of records, a binary class was created to represent "no information". When only a few records were missing information, these records were assigned the mean value of the variable(s) for the observations with information. As discussed in the feasibility report, since the regression technique forces the estimated equation to pass through the point of means, the created items carry no weight in the products and cross-products that determine the regression coefficients. This procedure is based on the assumption that there is nothing systematic about the missing items, an assumption which does not seem restrictive when only a few records are involved.

It should be noted that the variables described in the following pages are the com[illegible]e set defined for all records in the survey. Not all of the age groups [illegible]ated out for analysis had a sufficient range of observations to warrant using the full set of variables. For these age groups variables

were combined or dropped, as necessary. For example, virtually all cases with length of pregnancy 6 months or less were deaths under one month of age. In analyses based on records for older children, this variable was combined with the ones representing longer pregnancies.

The following pages list the set of variables defined for all records. Variables preceded by an asterisk were also defined on the separate tape file created for records with valid birth weight information.

List of Variables

The definitions of binary variables specify which code, or codes, are to be denoted by = 1 (one). By implication all other codes are given a 0 (zero) entry for that variable.

Item	Variables defined
* DEATH OR SAMPLE	CTL = 1, if death
* RECORD ID NUMBER	
* AGE OF CHILD AT DEATH OR INTERVIEW	years, months, and days
* NEONATAL DEATH	NEO = 1, if infant died under 28 days
* SEX OF CHILD	SEX = 1, if female
* TYPE OF NEIGHBORHOOD	NBR1 = 1, if well-to-do, or good (medium = omitted class) NBR2 = 1, if poor NBR3 = 1, if very poor NBR4 = 1, if rural, or no neighborhood
NUMBER OF ROOMS IN THE DWELLING	RMS = number of rooms RMS1 = 1, if 1 room RMS2 = 1, if 2 rooms (3 or 4 rooms = omitted class) RMS3 = 1, if more than 4 rooms
NUMBER OF PERSONS IN THE HOUSEHOLD	FAM1 = 1, if 1 or 2 (3 or 4 = omitted class) FAM2 = 1, if 5 or 6 FAM3 = 1, if 7 or 8 FAM4 = 1, if 9 or more
NUMBER OF PERSONS PER ROOM (FAM ÷ RMS)	PER = number of persons per room PER1 = 1, if PERM ≤ 1.0 (1.0 < PERM ≤ 2.0 = omitted class) PER2 = 1, if 2.0 < PERM ≤ 3.0 PER3 = 1, if 3.0 < PERM ≤ 4.0 PER4 = 1, if 4.0 < PERM
* SOURCE OF WATER FOR THE HOUSEHOLD	(piped water inside = omitted class) WTR1 = 1, if piped water outside WTR2 = 1, if other
* TOILET FACILITIES	(flush, or communal = omitted class) TLT1 = 1, if other TLT2 = 1, if none

Item	Variables Defined
* EDUCATION OF FATHER	EDF1 = 1, if none (primary - omitted class) EDF2 = 1, if secondary EDF3 = 1, if technical, or university
* OCCUPATION OF FATHER	OCF1 = 1, if OCCF is blank, --, or -2 (unspecified) OCF2 = 1, if OCCF is -3, -4, or code beginning with 0 (0+ through 09) (professionals) OCF3 = 1, if OCCF is 29 (other clerical workers) OCF4 = 1, if OCCF is 30, 31, or 32 (sales: working proprietors, insurance and real estate, manufacturers' agents, etc.) OCF5 = 1, if OCCF is 33 (shop assistants and salesmen) OCF6 = 1, if OCCF is 40 (farmers and farm managers) (41 = omitted class: farm workers, n.e.c.) OCF7 = 1, if OCCF is 64 (drivers, road transport) OCF8 = 1, if OCCF is 75 (toolmakers, machinists, plumbers, etc.) OCF9 = 1, if OCCF is 79 (bricklayers, plasterers, and construction workers, n.e.c.) OCF10 = 1, if OCCF is 82 (millers, bakers, brewmasters, and related workers) OCF11 = 1, if OCCF is 89 (laborers, n.e.c.) OCF12 = 1, if OCCF is 70-74, 80, 81, or 83-88 (other crafts) OCF13 = 1, if OCCF is 76-78 (electricians, carpenters, painters, etc.) OCF14 = 1, if OCCF is 90 (firefighters, police, guards, etc.) OCF15 = 1, if OCCF is any other code (except 41)
* LENGTH OF PREGNANCY	PLG = length in months PLG1 = 1, if 6 months or less PLG2 = 1, if 7 months PLG3 = 1, if 8 months (9 = omitted class)

Item	Variables Defined
* SINGLE OR MULTIPLE BIRTH	MULT1 = 1, if multiple birth
* COMPLICATIONS DURING PREGNANCY	COMP = 1, if complications

Note: This variable is defined only for the child deaths. A "4" code has been entered for living children.

BIRTH WEIGHT	*BWT = weight in grams BWT1 = 1, if BWT ≤ 014 BWT2 = 1, if 015 ≤ BWT ≤ 019 BWT3 = 1, if 020 ≤ BWT ≤ 024 BWT4 = 1, if 025 ≤ BWT ≤ 029 BWT5 = 1, if 030 ≤ BWT ≤ 034 (035 ≤ BWT ≤ 039 is omitted class) BWT6 = 1, if 040 ≤ BWT ≤ 044 BWT7 = 1, if 045 ≤ BWT BWT8 = 1, if birth weight not specified BWTA = max (0, BWT minus 019) BWTB = max (0, BWT minus 029) BWTC = max (0, BWT minus 040)

Note: Combine lowest BWT classes (BWT1 - BWT4) when BWT is used in runs for children over 1 month

* BIRTH PLACE	BPLC1 = 1, if born at home, <u>or other</u> = 0, if born in hospital
* ATTENDANT AT BIRTH	ATND1 = 1, if midwife ATND2 = 1, if other, or <u>no one</u> (physician = omitted class)
* NUMBER OF VISITS FOR PRENATAL CARE	VST = number of visits VSTX = 1, if <u>any</u> prenatal care (0 = omitted class) VST1 = 1, if 1 or 2 visits VST2 = 1, if 3, 4, or 5 visits VST3 = 1, if 6, 7 or 8+ visits VSTA = Max (0, VST minus 2) VSTB = Max (0, VST minus 5)
WHETHER CHILD RECEIVED ANY VACCINATIONS	VACC = 1, if yes (either "1" or "2") (no or "7" = omitted class)
NUMBER OF WELL-CHILD VISITS TO CLINIC OR PHYSICIAN	WEL1 = 1, if blank, "-", or 8 (no visits = omitted class) WEL2 = 1, if 1 or 2 visits WEL3 1, if 3 or more visits

Note: Use this variable only for runs on children over 28 days. Most of the deaths under 28 days were in the hospital from birth to death

Item	Variables created
NUMBER OF ALL VISITS TO CLINIC OR PHYSICIAN	ALL1 = 1, if blank, "-", or 0 (no visits = omitted class) ALL2 = 1, if 1 or 2 visits ALL3 = 1, if 3 or 4 visits ALL4 = 1, if 5 or more visits
PERSON RESPONSIBLE FOR CARE OF CHILD	(mother = omitted class) CARE1 = 1, if codes 2 through 8 CARE2 = 1, if hospital

Note: Use this variable only with children over 28 days.

*EDUCATION OF MOTHER	EDM1 = 1, if none EDM2 = 1, if primary and EDYRM = 0, 1, or 2 (primary, EDYRM 3 or 4 = omitted class) EDM3 = 1, if primary and EDYRM = 5, 6, 7, or 8 EDM4 = 1, if secondary EDM5 = 1, if technical, <u>or</u> university
CHILD'S HISTORY OF INFECTIOUS DISEASES	("--" or 88 = omitted class) DIS1 = 1, if had <u>any</u> infectious disease under 1 year of age DIS2 = 1, if had <u>any</u> infectious disease at 1 year of age or older
CLASSIFICATION SYSTEM DEVELOPED TO SUMMARIZE DETAILED INFORMATION ON CHILD'S DIET (SEE RUSSELL AND BURKE, 1974, PART TWO, FOR DETAILS)	(protein probably adequate = omitted class) PRO1 = 1, if protein probably inadequate PRO2 = 1, if protein grossly inadequate PRO3 = 1, if unclassifiable PRO4 = 9, if child 12 months or older (calories probably adequate = omitted class) CAL1 = 1, if calories probably inadequate CAL2 = 1, if calories grossly inadequate CAL3 = 1, if unclassifiable CAL4 = 9, if child 12 months or older (iron probably adequate = omitted class) IRN1 = 1, if iron probably inadequate IRN2 = 1, if iron grossly inadequate IRN3 = 1, if unclassifiable IRN4 = 9, if child 12 months or older

Item	Variables created
	(vitamin C probably adequate - omitted class) VTC1 = 1, if vitamin C probably inadequate VTC2 = 1, if vitamin C grossly inadequate VTC3 = 1, if unclassifiable VTC4 = 9, if child 12 months or older
	(other vitamins and minerals probably adequate - omitted class) OVM1 = 1, if other V and M probably inadequate OVM2 = 1, if other V and M grossly inadequate OVM3 = 1, if unclassifiable OVM4 = 9, if child 12 months or older
WHETHER CHILD BREASTFED AND AGE WEANING STARTED	(no = omitted class) BRF1 = 1, if yes, weaning started less than 1 month BRF2 = 1, if yes, weaning started 1 mo BRF3 = 1, if yes, weaning started 2 mos BRF4 = 1, if yes, weaning started 3 mos BRF5 = 1, if yes, weaning started 4 mos or more BRF6 = 1, if yes, never weaned, or age of weaning unknown, <u>or</u> not known whether weaned BRFX = 1, if yes, weaning started 3 months or less BRFY = 1, if yes, weaning started more than 3 months

Note: (1) BRFX, BRFY, and BRF6 form a complete set of binaries; (2) these variables are very sensitive to the age group being investigated and the tabulations should be checked before specifying age-stratified regressions--e.g., most of those under 1 month who were never breast fed were in the hospital form birth to death (3) BRF6 is primarily those never weaned for children 3 months or less, and primarily those with age unknown for children 1 year or older.

OUTCOME OF PREVIOUS PREGNANCY	CL1 = 1, if previous pregnancy ended in fetal death (previous child born alive, still living = omitted class) CL2 = 1, if previous child born alive, subsequently died

Item	Variables created
*TOTAL NUMBER OF PREGNANCIES (equivalent to parity of child)	TPRG = no change (1 = omitted class) TPRG1 = 1, if 2 pregnancies TPRG2 = 1, if 3 or 4 pregnancies TPRG3 = 1, if 5 or 6 pregnancies TPRG4 = 1, if 7 or more pregnancies TPRGA = Max (0, TPRG minus 1) TPRGB = Max (0, TPRG minus 4) TPRGC = Max (0, TPRG minus 6) OR TPRGD = Max (0, TPRG minus 2)
*TOTAL NUMBER OF PREGNANCIES BY OUTCOME OF PREVIOUS PREGNANCY (1 = fetal death; 2 = previous child living; 3 or 4 = infant death)	(1 = omitted class) TPC1 = 1, if TPREG = 2, CLAST = 1 TPC2 = 1, if TPREG = 2, CLAST = 2 TPC3 = 1, if TPREG = 2, CLAST = 3 or 4 TPC4 = 1, if TPREG = 3 or 4, CLAST = 1 TPC5 = 1, if TPREG = 3 or 4, CLAST = 2 TPC6 = 1, if TPREG = 3 or 4, CLAST = 3 or 4 TPC7 = 1, if TPREG = 5 or 6, CLAST = 1 TPC8 = 1, if TPREG = 5 or 6, CLAST = 2 TPC9 = 1, if TPREG = 5 or 6, CLAST = 3 or 4 TPC10 = 1, if TPREG = 7 or more, CLAST = 1 TPC11 = 1, if TPREG = 7 or more, CLAST = 2 TPC12 = 1, if TPREG = 7 or more, CLAST = 3 or 4 TPC21 = 1, if TPREG = 2, CLAST = 1, 3, or 4 TPC22 = 1, if TPREG = 2, CLAST = 2 TPC23 = 1, if TPREG = 3 or 4, CLAST = 1, 3, or 4 TPC24 = 1, if TPREG = 3 or 4, CLAST = 2 TPC25 = 1, if TPREG = 5 or 6, CLAST = 1, 3, or 4 TPC26 = 1, if TPREG = 5 or 6, CLAST 2 TPC27 = 1, if TPREG = 7 or more, CLAST = 1, 3, or 4 TPC28 = 1, if TPREG = 7 or more, CLAST = 2
*INTERVAL BETWEEN PREVIOUS BIRTH AND BIRTH OF INDEX CASE, in months	SPC = no change SPCX = 1, if TPREG is 1 (i.e., first child) SPC1 = 1, if 12 months or less SPC2 = 1, if 13-15 months SPC3 = 1, if 16-18 months SPC4 = 1, if 19-24 months (25-36 = omitted class) SPC5 = 1, if 37-48 months SPC6 = 1, if more than 48 months

Item	Variables created
	SPCA = Max (0, SPC minus 12) SPCB = Max (0, SPC minus 15) SPCC = Max (0, SPC minus 18) SPCD = Max (0, SPC minus 24) SPCE = Max (0, SPC minus 36) SPCF = Max (0, SPC minus 48)

Note: Define SPC and SPCA-SPCF for children of the second or higher birth order only.

OCCUPATION OF MOTHER	(blank or "--" = omitted class) OCCM1 = 1, if OCCM = 41 (farm workers n.e.c.) OCCM2 = 1, if OCCM = 91 (housekeepers, cooks, maids) OCCM3 = 1, for all other occupational codes (0+, 0X, 00-99)

Note: For age-stratified regressions, OCCM1-OCCM3 should be combined.

AGE OF FATHER AT INTERVIEW	AGF = no change AGF1 = 1, if AGFI under 24 (24-30 = omitted class) AGF2 = 1, if AGFI 31-40 AGF3 = 1, if AGFI over 40
MARITAL STATUS OF MOTHER	MARM1 = 1, if single (married = omitted class) MARM2 = 1, if common-law <u>or</u> other
WHETHER BOTH PARENTS IN HOUSEHOLD	(both parents in household = omitted group) PHH1 = 1, if mother only, <u>or</u> father only <u>or</u> neither (blank)
CAUSES OF DEATH	Maintain all cause variables as is. In addition, create two new cause variables: CAUS1 = 1, if measles is any cause of death CAUS2 = 1, if diarrhea is any cause of death
*AGE OF MOTHER AT BIRTH	AGM = no change AGM1 = 1, if under 17 AGM2 = 1, if 17-20 (21-25 = omitted class) AGM3 = 1, if 26-30 AGM4 = 1, if 31-35 AGM5 = 1, if 36-40 AGM6 = 1, if over 40

Item	Variables Created
	AGMA = Max (0, AGM minus 16) AGMB = Max (0, AGM minus 20) AGMC = Max (0, AGM minus 25) AGMD = Max (0, AGM minus 30) AGME = Max (0, AGM minus 35) AGMF = MAx (0, AGM minus 40)

Note: For age-stratified regressions, combine AGM1 and AGM2, AGM5 and AGM6, AGMA and AGMB, and AGME and AGMF.

NUTRITIONAL STATUS AT INTERVIEW	(normal = omitted class) NS1 = 1, if malnutrition, degree I NS2 = 1, if malnutrition, degree II or malnutrition, degree III, or not specified

Note: Use only for living children. Virtually all living children under one year are classified as normal.

Appendix C. Weighting the Sample Observations

The San Juan survey is based on a stratified sample. It includes all the deaths of children under 5 that occurred during the two year period August 1968 - July 1970. The population of living children under 5 was sampled by selecting one out of every 48.3348 households for interview. Thus if the survey data are to represent events in the proportions in which they occur in the population, the observations must be weighted. The need for weighting is clear if we recall that the dependent variable in the mortality regressions is interpreted as the probability of death, and that this probability, and the changes in it associated with various factors, will be greatly overstated if the survey observations are combined without weighting.[1/]

The simplest weights that will serve the purpose are defined as

$$w_i = p_i/n_i \qquad (1)$$

where the weight is the ratio of the population in a certain group (e.g., deaths of children under 5) to the number of sample observations in that group. Thus if the survey includes all deaths and 1 out of every 50 living children, the appropriate weights are 1 and 50, respectively. If the sample is representative within strata, and the weights are applied to the elements of the regression moments matrix, the weighted elements will equal (in the expectation) the elements that would have resulted from a survey of the entire population

---

1/ Sampling variability, and the fact that different numbers of observations were dropped from each age group, mean that the survey observations represent slightly different proportions of the population, by age, within the groups of deaths and living children. The weights also adjust for these factors.

But the regression package available through PAHO, like many regression packages, was not designed to handle weighted data If it is fed weighted data, it produces test statistics based on the population moments (created by weighting) and the sample number of observations (N). One solution to the problem is to adjust the weights so that they sum to N over the observations. The weights then produce a moments matrix whose elements are proportional to the population moments, but are reduced in size to match the numbers of survey observations. This requirement produces a slightly more complicated formula for the weights:

$$\hat{w}_i = Np_i/Pn_i \qquad (2)$$

where $N = \Sigma n_i$ is the total number of observations in the sample, and $P = \Sigma p_i$ is the size of the population. All of the weights used in the study were calculated from formula (2).

Because the weights depend on N, new weights had to be defined for each subset of records selected for analysis In the following sections the "subset tapes" are described, the weights are shown, and any special assumptions are noted. Two subset tapes were created from the set of records with valid birth weight information and six were created from the set of all records. The basic information from which the weights were calculated -- the numbers of children in each age group for the sample and the population -- are shown in Table C.1.

I. BIRTHWEIGHT FILES

A. Records of Children under Two Years of Age.

The appropriate population for studying the determinants of birthweight is the complete population of birthweights, the variable of interest, for some

Table C.1. Child Population (for two years) and Numbers of Survey Observations, by Age

Age	Deaths			Sample		
	Population (1)	All Survey Observations (2)	Survey Observations with Valid Birth Weights (3)	Population* (4)	All Survey Observations (5)	Survey Observations with Valid Birth Weights (6)
Under one year, total	1,900	1,742	1,578	20,280	204	192
0-27 days	906	806	762	NA	20	18
28 days-one month	163	150	128	NA	23	23
two months	164	156	145	NA	22	21
three months	164	151	135	NA	16	15
four months	141	135	114	NA	21	20
five months	103	97	87	NA	17	17
six months	68	64	54	NA	14	13
seven months	56	54	46	NA	14	12
eight months	44	42	36	NA	11	10
nine months	28	24	15	NA	13	12
ten months	33	33	29	NA	11	11
eleven months	30	30	27	NA	22	20
One year	153	145	123	21,040	203	191
Two Years	55	55	48	63,140	156	141
Three Years	32	32	27		159	138
Four Years	16	15	13		156	140
ALL AGES, TOTAL	2,156	1,989	1,789	104,460	878	802

* From Puffer and Serrano, Patterns of Mortality in Childhood, p. 9.

specified time period. This is essentially equivalent to the population of live births for the period.[2/]

The San Juan survey provides a retrospective count of births during the period. The set of all records includes: all children who were born during the preceding five year period and are still living; and all children born during the last two years of the period who died. It omits the children who were born, and died, during the first three years of the period -- a sizeable group since death rates are high during the first year after birth. Thus it does not accurately represent the population of birthweights during the period.

But if we limit our consideration to records of children under two years of age, we do have a virtually complete roster of the population of birthweights during a two year period. This group includes all children born during the period and still alive[3/] and, because the survey recorded all deaths for two years, it also includes all children who were born and died during the two years. This group comes quite close to matching the desired population and the "birthweight tape" was created using only these records.

Weights were created for two age groups -- under 1 year, and 1 year -- using the four n_i for these groups from columns (3) and (6) in Table C.1 and the four p_i from columns (1) and (4). No distinction was made between records in the same age group. To illustrate the procedure, the numbers from which the weights were calculated are shown below.

Deaths	p_i	n_i
under 1 year	1,900	1,578
1 year	153	123

2/ The weights of stillborns are omitted by this definition. But if, as here, interest centers on the part birthweight plays in the infant's subsequent survival, the birthweights of live born infants may, in fact, be the preferred population.

3/ This statement is, strictly speaking, not quite accurate because the survey of living children was carried out in phases during the two years, and children under two who were sampled in the early months were not necessarily born within the period. But the population numbers used to calculate the weights are properly defined, and it seems reasonable to believe that the survey children are representative of children born within the period.

Sample	p_i	n_i
under 1 year	20,280	192
1 year	21,040	191
Total	P=43,373	N=2,084

Applying formula (2), the weights are:

	Deaths	Sample
under 1 year	.05785286	5.07510433
1 year	.05976743	5.29286261

B. Records of Children under Two Years of Age and of Parity Two or Higher.

Some of the factors believed to affect birthweight -- specifically the interval since the last birth, and the outcome of the previous pregnancy -- apply only to children after the first, and cannot be entered in the regressions when first births are part of the observation set. To permit investigation of these factors, a second birthweight tape was created that excluded first births from the set of observations defined in I.A. The weights for this new set are given below:

	Deaths	Sample
under 1 year	.05208667	4.88147727
1 year	.05443291	5.05390030

II. INFANT MORTALITY FILES

A basic infant mortality file contained all the variables defined for the regressions, for all infants -- deaths and sample -- 0 to 11 months of age. Four "subset tapes" were created from this file. Three of the four were defined in the spirit of "cohort" death rates, that is, rates that compare deaths in a certain age group with the number of living persons in the same age group. (Since no information was available on the distribution by age of the living population under one year, it was assumed to be distributed evenly by month for the purpose of calculating the weights.) The fourth

"subset tape", described first in this section, was defined along the lines of the standard infant mortality rate, and requires a few words of explanation.

A. All Infants 0-11 Months of Age

The purpose of this subset was to investigate the determinants of neonatal mortality. This purpose is not apparent from the defined subset nor from the weights, both of which are appropriate for other types of analysis as well. It is made clear by the dependent variable used for the analysis, which was defined as one for deaths 0-27 days of age, and as zero for all other deaths and all living children.

It was not feasible to apply the notion of cohort death rates to neonatal deaths because only twenty living children from the same age group were included in the survey. Instead the standard definition of infant mortality rates was used to suggest the subset to be analyzed. By this definition infant deaths during a given period are compared with all live births, regardless of whether the children survived or died after live birth. The match between the concept and the survey data is not exact since the survey of living children was undertaken in phases and children 0-11 months during the early part of the survey period were not necessarily born during the period. But it seems quite reasonable to treat them as representative of children who were, and the population numbers used to calculate the weights are conceptually correct.[4/] In this spirit the dependent variable was defined as described and the weights are given below:

4/ Note, however, that the living children 0-11 months were weighted to resemble the population of living children of that age, not the number of live births. Live births include children who later died and the inclusion of the deaths over 27 days brings the "comparison group" (i.e., the group for which the dependent variable is zero) closer to this concept.

	Deaths	Sample
0-27 days	.09862215	
28 days - 5 months	.09359430	7.47605159
6-11 months	.09199921	10.46647218

B. Infants 28 Days to 5 Months

	Deaths	Sample
28 days-5 months	.09151982	7.32265497

C. Infants 6-11 Months

	Deaths	Sample
6-11 months	.03347721	3.80860151

D. Diarrhea Deaths and All Living Children 28 Days to 5 Months

Death rates due to a specific cause are generally calculated by comparing the deaths from that cause with the living children in the same age group. Accordingly, the subset drawn to investigate diarrhea deaths omitted all deaths from other causes. In order to calculate the weights it was necessary to assume that the number of diarrhea deaths in the survey bore the same relation to the population of diarrhea deaths, as the number of survey deaths from all causes bore to the population of deaths from all causes.

	Deaths (Diarrhea)	Sample
28 days - 5 months	.05697490	4.45775869

III. CHILD MORTALITY FILE

This tape was made up of all variables defined for the regression analysis, for children one to four years of age. This group was analyzed as a whole and the set of weights was defined in the spirit of cohort death rates.

	Deaths	Sample
1 year	.01150947	1.13052892
2 - 4 years	.01101461	1.46222986

APPENDIX D. Bibliography

Abramowicz, M. and H. Parnett, "Sex Ratio of Infant Mortality," American Journal of Diseases of Children, vol. 119 (April, 1970).

Aykroyd, W., "Nutrition and Mortality in Infancy and Early Childhood: Past and Present Relationships," American Journal of Clinical Nutrition, vol. 24 (April, 1971).

Baertl, J.E., E. Morales, G. Verastegui, and G. Graham, "Diet Supplementation for Entire Communities. Growth and Mortality of Infants and Children," American Journal of Clinical Nutrition, vol. 23 (June, 1970).

Berg, A. and R.J. Muscat, The Nutrition Factor, Washington, D.C.: The Brookings Institution, 1973.

Burgess, A. and R.F.A. Dean (eds.), Malnutrition and Food Habits, London: Tavistock Publications, 1962.

Burton, B., The Heinz Book of Nutrition, New York: McGraw-Hill, Inc., 1965.

Cameron, M. and Y. Hoffander, "Manual on Feeding Infants and Young Children," Protein Advisory Group of the United Nations System, 1971.

Cook, J., D.G. Altman, D.M.C. Moore, S.G. Topp and W.W. Holland, "A Survey of the Nutritional Status of School Children: Relation Between Nutrient Intake and Socio-Economic Factors," British Journal of Preventive and Social Medicine, vol. 27 (1973).

Erhardt, C., H. Abramson, J. Pakter and F. Nelson, "An Epidemiological Approach to Infant Mortality," Archives of Environmental Health, vol. 20 (June, 1970).

Hammouda, E., "Studies in Fetal and Infant Mortality: A Methodological Approach to the Definition of Perinatal Mortality," American Journal of Public Health, vol. 55 (July, 1965).

Harrison, K.A., and P.A. Ibeziako, "Maternal Anemia and Fetal Birthweight," Journal of Obstetrics and Gynaecology of the British Commonwealth, vol. 80 (September, 1973).

Hillman, R. and J. Hall, "Nutrition in Pregnancy" in Modern Nutrition in Health and Disease, Wohl and Goodhart (eds.), Philadelphia: Lea and Febiger, 1964.

Jelliffe, D.B., "Infant Nutrition in the Subtropics and Tropics." Geneva: World Health Organization Monograph Series, No. 29, 1968.

APPENDIX D, continued

Jelliffe, D.B., Child Nutrition in Developing Countries: A Handbook for Fieldworkers, Washington, D.C.: Agency for International Development, 1969.

Johnston, J., Econometric Methods, second edition, New York: McGraw-Hill, Inc., 1972.

Kapur, S., G. Kumar, K.C. Mammen, and G. Jesudian, "Height and Weight of South Indian Women of Child Bearing Age and Their Effect on Birth Weight and Length of the Baby," Indian Journal of Medical Research, vol. 59, no. 9 (September, 1971).

Kessner, D.M., J. Singer, C.E. Kalk, and E.R. Schlesinger, Infant Death: An Analysis by Maternal Risk and Health Care, Washington, D.C.: Institute of Medicine, National Academy of Sciences, 1973.

Mata, L., J. Urrutia and A. Lechtig, "Infection and Nutrition of Children of a Low Socio-Economic Rural Community," American Journal of Clinical Nutrition, vol. 24 (February, 1971).

Mata, L. and R. Wyatt, "Host Resistence to Infection," American Journal of Clinical Nutrition, vol. 24 (August, 1971).

Mata, L., J. Urrutia, C. Albertazzi, O. Pellecer and E. Arellano, "Influence of Recurrent Infections on Nutrition and Growth of Children in Guatemala," American Journal of Clinical Nutrition, vol. 25 (November, 1972).

Melichar, E., "Least-Squares Analysis of Economic Survey Data," American Statistical Association, Proceedings of the Business and Economic Statistics Section, 1965.

Monckeberg, F., "Factors Conditioning Malnutrition in Latin America, With Special Reference to Chile. Advice for a Volunteer Action," Bibliotheca Nutritrio et Dieta, no. 4, (1970).

Morris, J. and J. Heady, "Social and Biological Factors of Infant Mortality: I. Objects and Methods," The Lancet (February 12, 1955).

National Academy of Sciences, "Maternal Nutrition and the Course of Pregnancy: Summary Report," (Committee on Maternal Nutrition, Food and Health Board, National Research Council), Washington, D.C. 1970.

Pethybridge, R.J., J.R. Ashford, and J.G. Fryer, "Some Features of the Distribution of Birthweight of Human Infants," British Journal of Preventive and Social Medicine, vol. 28 (February, 1974).

APPENDIX D, continued

Puffer, R.R., and C.V. Serrano, Patterns of Mortality in Childhood, Washington, D.C.: Pan American Health Organization, 1973.

Russell, L.B., and C.S. Burke, Determinants of Infant and Child Mortality: Report of the Feasibility Study, prepared for the Agency for International Development under Contract AID/CM/otr-C-73-212, Work Order No. 1, Washington, D.C.: National Planning Association, Part One (February, 1974) and Part Two (March 1974).

Scrimshaw, N. and J. Salomon, H. Bruch and J. Gordon, "Studies of Diarrheal Disease in Central America, VIII: Measles, Diarrhea and Nutritional Deficiency in Rural Guatemala," American Journal of Tropical Medicine and Hygiene, vol. 15, no. 4 (1966).

Scrimshaw, N., C. Taylor and J. Gordon, "Interactions of Nutrition and Infection," World Health Organization Geneva Monograph Series no. 57, 1968.

Scrimshaw, N., "Synergism of Malnutrition and Infection. Evidence from Field Studies in Guatemala," Journal of American Medical Association, vol. 212 (June 8, 1970).

Shah, K.P., and P.M. Shah, "Relationship of Weight During Pregnancy and Low Birth Weight," Indian Pediatrics, IX:9 (September, 1972).

Sloan, F., "Survival of Progeny in Developing Countries: An Analysis of Evidence from Costa Rica, Mexico, East Pakistan and Puerto Rico," Santa Monica, California: Rand Corporation (AID/CSD-2533), 1971.

St. George, J., E.H. St. John, and D. Josa, "Factors Influencing Birth Weight in Normal Pregnancy," The Journal of Tropical Pediatrics (September, 1970).

Watts, H., "An Introduction to the Theory and Practice of Binary Variables (or All About Dummies)," mimeo, April 15, 1963.

Wray, J., "Population Pressure on Families: Family Size and Child Spacing," Reports on Population/Family Planning, no. 9 (August, 1971).