TOPICS IN DEVELOPMENT, INCOME INEQUALITY AND HEALTH

TOPICS IN DEVELOPMENT, INCOME INEQUALITY AND HEALTH

By

FREDRICK ASWANI JNR., M.A.

A Thesis

Submitted to the School of Graduate Studies

In Partial Fulfilment of the Requirements

for the Degree

Doctor of Philosophy

McMaster University

DOCTOR OF PHILOSOPHY
(2000)
(Economics)

McMASTER UNIVERSITY
Hamilton, Ontario

TITLE: **Topics in Development, Income Inequality and Health**

AUTHOR: **Fredrick Aswani Jnr., M.A. (Lakehead University)**
B.A. (University of Nairobi)

SUPERVISING COMMITTEE: **David Feeny**
Lonnie Magee

NUMBER OF PAGES: **x, 134**

ABSTRACT

The purpose of this study is to contribute to the existing knowledge on determinants of child health status. Such knowledge may assist in formulating policy in a number of areas especially with regard to child care and development.

The aggregate analysis is based on purchasing power parity adjusted per capita real gross domestic product at 1985 prices (per capita *RGDP*) and its distribution, and their relationship to life expectancy at birth and infant mortality rate. We find quantitatively important and statistically significant relationships between life expectancy and infant mortality rates and per capita *RGDP*. The relationship is well represented by a quadratic spline model that shows a diminishing contribution of income to health status as per capita income grows. We further incorporate selected socio-economic indicators and show that literacy levels have significant impacts on infant mortality and life expectancy. Overall, income distribution did not seem to matter controlling for the non-linear effect of per capita income on health. The potential importance of the non-linear effect is illustrated by simulation exercises that show that redistributing income may improve overall health status especially in the middle income, but has little effect at lower and higher per capita income ranges.

The micro-economic (household) analysis utilizes demographic and health survey data sets for Kenya. Using the logit model we estimate the probability of child survival conditional on certain household, parental and socioeconomic factors. Results show that mother's age and education, presence of other young siblings in the household, presence of other wives, marital status and breast-feeding have quantitatively important and statistically significant impacts on the child's survival. Policy implications are discussed.

ACKNOWLEDGEMENTS

I am greatly indebted to my co-supervisors, Professor David Feeny and Professor Lonnie Magee for their intensive guidance and support. I thank you greatly for the assistance you gave me throughout all stages of this thesis. I also wish to acknowledge the help of Professors Browning and Dooley who also contributed to the work reported in the thesis.

I would also like to thank Professor Mike Veall for his kind support and guidance throughout my PhD programme. I am grateful to the administration and secretarial staff in the Economics Department especially Mrs. Aloma Munian, who was always ready to assist me both on departmental issues and as a friend.

I wish to thank posthumously my brother John Baptiste who always gave me reason to smile even when things looked enervating. You, and my late son B.J., are my heroes.

Special thanks to my wife Rosemary who had to bear with my years of schooling. Thanks for your strength, encouragement and support. Without your being with me, beside me, behind me and around me this task would have been impossible to accomplish. Life's caravan does not turn back. And to my daughters Tina, Vicky and Joanne who endured many evenings of "Quiet! Dad is working" and my son Ngashe, may you all realise your dreams and reach your goals.

To my sisters Ursula, Imelida, Jacinta and Seli thank you all for your support. To my parents Maria and Waswani, "murio muno", I owe you all. This is for you.

TABLE OF CONTENTS

LIST OF TABLES

Chapter 3. Aggregate Analysis

LIST OF TABLES cont…

Chapter 4. Micro Household Analysis

LIST OF FIGURES

CHAPTER 1

INTRODUCTION

1. Introduction:

The overall theme of this study is the determination of child health with particular emphasis on economic factors. Improving our basic understanding of the factors that influence health status is important. Such understandings can help shape policy in a number of areas including public health, health care, education, income distribution, and social services as well as policies to promote overall economic growth and development. After presenting the household model on which the analysis is based, we address determinants of health at the aggregate level. This is followed by an analysis of child health determinants at the microeconomic (household) level. These micro-focused studies complement the more aggregate approaches.

Chapter 1 provides the introduction to the study. Chapter 2 focuses on the model and variable specifications. We provide the microeconomic foundations of the household model that leads to the estimating equations at both the household and aggregate levels. We estimate "stylized" reduced form equations of the household model. Later in the chapter, we discuss in detail the expected influence of various variables identified in the household model on health status at the aggregate level.

Chapter 3 deals with aggregate analysis using macro-economic data. After discussing the data sources, we focus on cross-sectional analysis with particular emphasis on purchasing power parity (PPP), adjusted per capita real gross domestic product (hereinafter referred to as per capita RGDP) at 1985 prices, including its distribution and their relationship to life expectancy and infant mortality using circa 1990 data. We also examine the time series trends in life expectancy and infant mortality rates and compare the structural relationship between life expectancy at birth and per capita RGDP between 1970 and 1990. We extend our analysis to compare cross-sectional results for different periods, the 1970s, 1980s and 1990s.

Chapter 4 focuses on microeconomic (household) determinants of child health. This Chapter uses data from the 1988-89 and 1993 Kenya National Demographic and Health Surveys. The last chapter provides a summary, proposed extensions and conclusions.

CHAPTER 2

MODEL AND VARIABLE SPECIFICATIONS

2.1. Conceptual Model:

The household model emphasises the roles of human capital, household wealth endowment and the allocation of time in the production of household goods and services. Extensions of the household model examine inter-generational considerations, biological and environmental factors, imperfect information and the tastes of households in determining household production and consumption. In the model, market goods and household time are used to produce non-marketed goods. The household derives utility from both marketed and non-marketed goods. Health is one such non-marketed good.

We will develop the household model from its micro-economic foundations and review some issues that should be taken into consideration in aggregation. Aggregation is necessary in order to apply the model to aggregate cross-sectional and time series data as we do in Chapter 3. The model developed below is a modified version of the household model as presented by, among others, Behrman and Deolalikar (1988), Behrman and Wolfe (1987), Rosenzweig & Schultz (1983), Strauss (1986) and Feeny (unpublished).

The household model is based on the assumption that household preferences can be represented by a household utility function. For the purposes of this study the household utility function is represented by:

$$U = U(H,C,S) \tag{2.1}$$

where each argument is a vector:

H - Health status of household members, i.e., male, female and child health status separately.

C - Consumption of final composite goods by household members (includes leisure).

S - Number of surviving children.

We assume that the utility function is well behaved, i.e. quasi-concave and with positive partial derivatives.

In addition to the number of surviving children, household utility is also a function of quality of children, which is in part a function of child care. For example, a family with a large number of poorly educated, low quality children may generate less utility than a family with fewer but better educated, higher quality children. The contribution of quality of children (other than health status) is subsumed under the composite good.

The household maximizes its utility function subject to three sets of constraints: a production function constraint, an expenditure constraint and a time constraint. The production constraint can be represented by the following implicit production function.

$$F(H,C,S,N,E,T_{fi},Tmi,I,MG,P,MC) = 0 \qquad (2.2)$$

This function is assumed to satisfy the usual properties of production functions, i.e. it is quasi-convex, increasing in outputs and decreasing in inputs. The first three arguments represent the outputs as defined in the household utility function (2.1) above while the rest represent inputs where each variable represents a vector over household members.

N - Nutrient intake

E - Education

T_{fi}, T_{mi} - Time devoted by female and male household members respectively towards production of i, where i represents the three outputs H, C and S.

I - Household composition and characteristics. These include:

. sex and age of children
. parents' ages
. marital status
. residence

MG - Market goods

P - Physical environment
e.g. access to safe water and sewage facilities.

MC - Medical care

Implicit specification of the production function has certain advantages over the explicit form. Strauss (1986) notes that this implicit specification allows for both separate and joint production of different outputs.

The household faces a standard budget or expenditure constraint represented by:

$$Y = MG + MC \tag{2.3}$$

which specifies that expenditure on market goods *(MG)* plus that on medical care goods *(MC)* has to sum to the total money income *(Y)*. This is a one period model with no bequests. Total money income is generated as follows:

$$Y = L + V$$

where L and V are labour and non-labour income respectively and

$$L = w_f T_{fw} + w_m T_{mw}. \tag{2.4}$$

w_f and w_m are female and male market wages respectively.

The third constraint is a time constraint,

$$T = T_{fc}+T_{fh}+T_{fs}+T_{fw} = T_{mc}+T_{mh}+T_{ms}+T_{mw} \tag{2.5}$$

T - Total amount of time available per adult

T_{fc} - Time female adults spend on production of composite good. This includes all activities within the household that may or may not indirectly lead to improved health status, e.g. cooking, knitting, listening to music and house cleaning.

T_{fs} - Time female adults spend on care of surviving children

T_{fh} - Time female adults spend on direct production of health status

T_{fw} - Time female adults spend in the labour market

Time components for males are similarly defined.

The full income constraint that is derived from time and expenditure constraints is represented by:

$$FI = w_f T + w_m T + V = p_c C + p_h H + p_s S \tag{2.6}$$

where

FI - full income

p_c - shadow price of composite final goods

p_h - shadow price for health status

p_s - shadow price of surviving children

V - non-labour income

w - wage rate

The full income constraint is captured by a corner solution in which the household spends all of its time in the labour market. Thus full income represents the total value of household time. If the household chooses to use some of its time for other activities (including leisure), it must deem that the value of that time in those other activities is at least as valuable as the opportunity cost of its labour services. This full income constraint states that total expenditure on composite goods, health status and surviving children exhausts total potential income accruing from the labour market and non-wage income.

The household consequently maximizes utility (2.1) subject to the three constraints with fixed prices for market goods, medical care and labour services *(w)*. Shadow prices of health status (p_h), of surviving children (p_s) and composite goods (p_c) are endogenous. If we assume interior solutions we could represent each optimal solution for *C, H* and *S* as an explicit function of both endogenous and predetermined variables. Behrman and Wolfe (1987), however, noted that the choice of the functional form would be arbitrary.

Rosenzweig and Schultz (1983) assumed Cobb-Douglas function formulations but in the end these formulations tended to be too restrictive and made it difficult to identify all of the structural parameters. However, even if we were to assume some functional form for the utility function and production functions, along with identifying restrictions, we would end up with a simultaneous system of equations, which we do not have sufficient data to estimate.

We follow the approach adopted by earlier studies, for example, Behrman and Wolfe (1987) and estimate linear "stylized" reduced equations for health status because this is our primary focus. As Behrman and Wolfe note, these linear relations could be viewed as local approximations to the more complex ones that would result from maximization of the complete system with explicit functional forms and then estimating the resulting simultaneous equation system.

2.2 A "Reduced Form" Equation:

As stated in the previous section we estimate "stylized" reduced form equations with health indicators as the dependent variables. These are derived from maximization of the household utility function (2.1) subject to the constraints (2.2), (2.3), (2.4), (2.5) and (2.6). Based on previous studies we indicate variables that have been found to play a major role in the system. To reiterate what was indicated in the previous section these linear formulations could be viewed as local approximations to the more complex ones. The "reduced form" solution at the micro-economic (household) level is given by:

$$H = H(N,E,P,I,MG,MC,Y) \tag{2.7}$$

where H is a measure of health status, in this case life expectancy at birth and infant mortality rate. The right hand variables are as defined in (2.2) and Y is money income as in (2.3). Each variable is a vector over household members.

Note that in both the utility function and the reduced form, there is a need for incorporating household composition, i.e. the age, sex and number of household members. Household composition may influence what it takes to produce health and therefore influence health status and consequently utility. Chapter 4 of the thesis will focus on these micro-economic reduced form solutions.

Due to the potential for income to have nonlinear effects on health (Preston (1975), Lindert (1986) and Wennemo (1993)), we consider income distribution in the aggregate analysis. There is also evidence that those with low incomes tend to have both poor nutrition and health. At the micro level we have:

$$hs_{ij} = h(z_{ij}, g_{ij}, Q_{ij}) \tag{2.8}$$

where:

hs_{ij} - health status for household i in country j

z_{ij} - characteristics of the household

g_{ij} - household income

Q_{ij} - other factors mentioned in (2.7)

If we average over all households, we derive

$$hs_j = f[w(g_j, s), Q_j] \tag{2.9}$$

where:

hs_j - is a general health indicator for country j, notably life expectancy or infant mortality rate.

The term $w(g_j,s)$ shows that the health indicator for country j is not only a function of its average income measure (g_j) but may also be a function of income distribution (s).

Q_j - refers to other factors influencing health at aggregate level.

It may also be noted that in aggregation, random errors that are household specific will tend to average out. A reduced form aggregate equation which is based on the household preference equation (2.1), health production function (2.2), time constraint (2.5), the money income constraint (2.6) and the reduced microeconomic formulation (2.7), which takes aggregation into consideration can be given by

$$H_j = H(Y_j, \theta_j, E_j, P_j, O_j) \tag{2.10}$$

where:

H_j - Health indicator for a given country. In this study as already stated, we use the average life expectancy at birth and infant mortality rate.

Y_j - Income (in this study, measured by per capita *RGDP*). Income distribution could also play a major role in the observed differences in health status. Measures of income distribution are therefore, as noted earlier, incorporated in the analysis.

θ_j - Price vector. This will be proxied by:

. population/doctor ratio
. percentage of urban population

E_j - Average education levels which is a measure for stock of human capital.

P_j - Physical environment

O_j - Public policy, e.g. the tax system, organization of health care, and physician reimbursement mechanisms.

2.3. Selection of Variables for Aggregate Analysis:

The household model provides guidance on the selection of variables. We now need to consider the quantity and quality of various indicators related to the conceptual categories identified in the household model. Below we focus on variables that are analysed in the aggregate time series estimation in Chapter 3. Later in Chapter 4, we also discuss the selection of variables for the microdata estimation. The aim is to provide a justification for their inclusion. The attention to these factors has grown out of the increased interest of differentials in health status across countries that have been associated with cultural, socioeconomic and behavioural factors.

These factors are captured in the basic household model discussed in the previous sections. The World Bank Development Reports (1990 &1993) note the importance of income and medical interventions in the dramatic mortality declines in the Less Developed Countries (LDCs) in the 1950s and 1960s. As indicated in section (2.2), the dependent variables in both the household and aggregate analyses will be health indicators. At the aggregate level, we use infant mortality rates and life expectancy at birth. It should be noted that although the relationship between health indicators and the right hand variables could differ across ages, we elected to use life expectancy at birth and infant mortality rates. We do not investigate life expectancy at later ages mainly due to the lack of data for a large number of countries. Earnings in adulthood are also influenced by the health of the individual. Therefore we risk running into the problem of reverse causation if we investigate the relationship between income and life expectancy in later years.

Life expectancy at birth is normally regarded as a good overall measure of population health status. Infant mortality, on the other hand, is a good indicator of child health and data is available for most years and countries. It should be noted further that our major interest in this study is the determinants of child health and although life expectancy at birth is not directly a measure of child health, Figure 1 shows that its correlation coefficient with infant mortality using 1990 data is very high in magnitude: -0.96.

The relationship between health status and income may be nonlinear as we shall see later. Therefore in an aggregate analysis we ought to take this potential nonlinearity into consideration. The nonlinear effect of income on health implies that there are diminishing returns to income with regard to health status as we move from low to high income countries. It is therefore important to further investigate the effects of income distribution. Apart from analysing the general effect of income distribution we will also investigate the question of whether differences in general health status among countries at same levels of real per capita incomes could be attributed solely to the nonlinear effects of income.

We investigate the effect of daily calorie supply on the general health of the population. This data comes from the Human Development Report 1992 and mainly covers LDCs, (UNDP, 1992).

We may also expect literacy levels to have a positive influence on health status. In the estimation of health status "stylized" reduced form equations, it is especially important to incorporate female literacy levels. In general, mothers are more directly involved in the production of child health than fathers. We have used, however, total illiteracy in the aggregate macro-economic analysis because this data is more readily available for most countries from the World Bank sources. Also, total and female illiteracy are highly correlated (0.99) in 1990 for the set of countries for which both measures were available. Female illiteracy was not reported for most developed countries in the World Bank reports.

The percentage of a country's population that is urban may also play an important role in the county's health status as measured by life expectancy at birth and infant mortality rate. Physicians tend to be attracted to urban areas. Thus we might expect better health for urban children relative to rural ones. Generally, there are better health care, water and sewage facilities in urban areas, but also more need for them due to a higher population density. Urbanization is thus included in both household and aggregate analyses as a proxy for the price of health care. The distance and cost involved in accessing health care in urban areas is generally lower. This will affect utilization and consequently health status. Urbanization complements the population/doctor ratio as proxies for the price of health care.

For lack of adequate data availability, the effect of physical environment on health will be analyzed only at the micro level. At household level we will consider the availability of toilet facilities. At the aggregate level, incorporation of availability of sewage facilities may lead to endogeneity problems. Sewage facilities may be endogenous at the aggregate level because their availability is dependent on taxes, which are a function of per capita national income and the degree of economic development. We do not have any proxy for genetic traits at the household or aggregate levels. This variable is not, therefore, investigated in the study.

In chapter 3 we will focus on cross-sectional analysis of these relationships across countries, in 1970, 1980 and 1990. We also consider trends in health status indicators, i.e. infant mortality rates and life expectancy at birth.

In order to deal with issues of multicollinearity we excluded certain variables whenever they were highly correlated with other explanatory variables. For example, share of agriculture in GDP was highly correlated to urbanization. This led to the exclusion of share of agriculture as an explanatory variable. The exclusion criteria for variables are based on maintaining a reasonable sample size, observed correlation coefficients between explanatory variables plus F-restriction tests. Variables that have data available for a wide selection of countries are given priority.

CHAPTER 3

AGGREGATE ANALYSIS

3.1. Aims and Objectives:

It has generally been observed that despite reduced financial barriers to access medical care, there are still marked differences in health within and across countries. This trend is observed in both developing and developed countries. In United Kingdom, for example, it was observed that despite the effectiveness of the National Health System in removing the financial barriers to health care in the 1980s, the gap between health status of the different social classes was widening (CIAR, 1991).

In a health conference held in 1978 (referred to as the Alma Ata Conference) which examined the issues of health care in developing countries, strategies were mapped out for primary health care with the primary objective of achieving the WHO goal of health for all by the year 2000, (ENHR, 1991). Attention during the intervening years was to focus on research and dissemination of information on the association between health status and the variables identified in the household model above. This research attempts to contribute to this broad WHO objective.

The equations estimated, as discussed in Chapter 2, are "stylized" reduced form equations of an aggregate health status production function model, obtained by aggregating over households based on the reduced form solution of the household model. In this chapter, cross sectional and time series data are used to investigate the interplay of various economic, demographic and socio-economic factors with health. Studies have documented relationships between life expectancy (and other indicators of health status including crude death rates, age specific mortality rates, life years gained and disability adjusted life years) and RGDP, per capita RGDP, other national income accounting measures, inequalities in distribution of income, population doctor ratio, nutrition, level of unemployment and literacy levels among others.

The motivation for the work that is covered in this chapter mainly arises out of the need to estimate the structural relationship between health status indicators and income, and further, to attempt to show what has happened to this structural relationship over time. We attempt to replicate and extend the work done mainly by Preston (1975), Waldmann (1992) and Pritchett and Summers (1996) using more recent data sets. Their contributions are reviewed later, in section 3.3.

3.2. Data Sources:

The income data used in the aggregate analysis is from the Penn World Tables 5.6 (PWT56). We use the series RGDPC for 1970, 1980 and 1990. RGDPC is GDP per capita at 1985 Real Purchasing Power (PPP) adjusted dollars. We refer to it in the study as per capita *RGDP*. This data set represents the most recent one available from the International Comparisons Project and adjusts for price level differences across countries. (It is available online at http://datacentre.chass.utoronto.ca/pwt/index.html). Data on most other variables were supplied on diskette from the World Bank, Washington D.C. This comes with a retrieval program (called STARS) and covers time series and cross-sectional data on social indicators of development (1993 edition) over the period 1965-1991 for 185 countries. The World Bank Development Report (1993) entitled "Investing in Health" was a major complement to this data set. These main sources of data are supplemented by other World Bank Development Reports and the 1992 UNDP Human Development Report. Appendix 1 gives a detailed list of the data sources and variable definitions. The complete RGDPC data set for the three years that are used is given as Appendix 2. Appendix 3 outlines the history and importance of using purchasing power parity income rather than exchange rate adjusted income data.

The World Bank Development Report (1990) notes that data reported for infant mortality rates in various World Bank reports are based on interpolations, extrapolations or at times comparisons with other countries. The consistency in these data imposed by World Bank analysts makes the data useful for comparing health indicator levels across countries at a point in time. Life expectancy on the other hand is derived under the assumptions that a cohort born today goes through the existing age specific death rates. It is heavily influenced, therefore, by the current infant mortality rates. In this study we focus on cross country analysis of health indicators at a point in time.

Since we will be estimating structural shifts later in Chapter 3, for consistency, from the outset we only consider countries that had per capita *RGDP* data for all three years: 1970, 1980 and 1990. There were a total of 107 countries. Out of these 107 countries, two countries, Jordan and Seychelles, were left out because they did not have complete infant mortality and life expectancy data for the period being analysed. In addition, Lesotho was omitted because of suspected errors in their reported infant mortality rate figures between 1973-1981. The reported infant mortality rates for Lesotho for these years were unrealistically high. Lesotho is, for example, reported to have had an infant mortality rate of 130 in 1972, 347 in 1973 and peaking up at 1213 in 1977. Between 1977-1981 the reported infant mortality rate was over 300 and then settled back to a level of 100 in 1982.

3.3. Related Empirical Studies:

Health status in both more developed countries (MDCs) and less developed countries (LDCs) has improved dramatically in the last 40 years. This improvement has been felt more in LDCs. In 1950 the average life expectancy in the LDCs was 40 years; this increased to 60 years by 1990. Over the same period, the number of children dying before the age of 5 declined from about 30 to 10 out of every 100 live births. This is still, however, about 10 times higher than the rate in the MDCs. The improvement in health is attributed to, among other factors, increased levels of income, increases in literacy rates, improvements in general public health and the introduction of particular interventions such as inoculation for smallpox, immunization for measles and polio, and malarial suppression.

Table 3.1 shows improvements in health and associated health indicators for selected demographic regions. Figure 3.2 also provides the disability adjusted life years (DALYs) associated with premature mortality and disability for the same regions. There is wide variation in DALYs per 1000 population associated with premature mortality compared to that associated with disability between various geographic regions. DALYs associated with premature mortality was 442 per 1000 population for Sub-Saharan Africa as compared to only 62 for established market economies (EME). On the other hand DALYs associated with disability was 132 and 60 per 1000 population for Sub-Saharan Africa and EME respectively. The world average was 170 per 1000 population for premature mortality and 89 for disability.

The World Bank Development Report (1993) strongly argues that there exists a positive causal relationship between income and health status. It further asserts that this influence tapers off at higher incomes. This had also been argued earlier in Preston (1975). Substantial empirical work has been done to investigate the relationships between per capita income, income inequality, level of development and health outcomes. (See for example, Ahluwahlia (1976), Ram (1992), Waldmann (1992), Preston (1975), Wennemo (1993), and Wilkinson (1992) among others.) Ahluwahlia's work is summarized in Appendix 4. Preston estimated the relative contribution of economic factors to increases in life expectancy for three different decades, the 1900s, 1930s and 1960s. He used cross sectional data on a number of countries to investigate the relationship between the level of economic development (measured by national income per capita in 1963 dollars) and average life expectancy at birth.

His analysis focused mainly on the following non-linear relationship over the three time periods:

$$exp(x)_t = d/[1+e(a+bY_t)]$$

where: $exp(x)_t$ = life expectancy for country x in year t.

Y_t = per capita income for country x in year

and $e(a+bY_t)$ denotes the antilog of $a+bY_t$.

Preston had sample sizes of 10 countries in 1900, 38 in 1930 and 57 in 1960. For the 30 countries that had data available for both the 1930s and 1960s, he also estimated the following linear multiple regression equation:

$$E=a+b_1X_1+b_2X_2+b_3X_3 \tag{3.1}$$

where:

E - Absolute change in life expectancy between 1930 and 1960.

X_1 - Level of income per head, 1930s

X_2 - Absolute change in income per head between 1930s and 1960s

X_3 - Ratio of income per head in 1960s to 1930s

Preston arrived at three major conclusions. First, the relationship between life expectancy and national income per head had shifted upwards during the twentieth century. Comparing the results for the 1930s and 1960s, he concluded that there had been a shift of 10-12 years between incomes of $100 and $500 and then less thereafter. That is, a country with real per capita income of $100 in 1960s had an average life expectancy of about 10 years more than a country with the same income in 1930s. The relationship between income and life expectancy flattened out at 66.8 years in 1930s compared to 71.5 in 1960s.

The shift, when viewed as a horizontal shift, showed that to attain a particular value of life expectancy between 40 and 60, a country required an income level approximately 2.6 times higher in the 1930s than in 1960s. (Preston (1975, p.236)). Preston also observed that factors other than growth in the country's level of income accounted for 75 - 90 % of the increase in life expectancy between the 1930s and 1960s. In addition, mortality had not become progressively disassociated with standards of living at a moment in time.

Correlation coefficients between life expectancy and log of per capita incomes were 0.885 in the 1930s and 0.880 in the 1960s.

Preston argued that nutrition and literacy were not the major factors explaining the shift between life expectancy and per capita income. When logistic curves were fitted between life expectancy and nutrition and also with literacy there was evidence of structural shift similar to that between life expectancy and real per capita income. Preston argued that diffusion of biomedical knowledge might have influenced the shift. (The information that is available on newspaper circulation in our data could have been used as a proxy for diffusion of biomedical technology but the sample is too small.) More recently, Wheeler (1980), in a simultaneous growth model of income, health and nutrition, found that nutrition mattered more to productivity and consequently health at low per capita incomes while literacy was more important at higher incomes.

Waldmann (1992) using a double log model, found that there was a positive significant association between infant mortality rate and percentage share of income that goes to top income earners even after controlling for the absolute income of the poor (defined as the lowest 20% of the population). Waldmann used data on income shares of the top 5% and imputed absolute per capita income of the lowest fifth from the International Bank for Reconstruction and Development's World Tables (1976 edition) supplemented by data from Summers and Heston (1984). Cross-sectional data on infant mortality rates was also taken from World Tables 1976. This data was supplemented by the UN population and vital statistics report for the year 1974. Estimates of per capita *RGDP* were taken from Summers and Heston (1984). Waldmann used infant mortality and income distribution data for 41 countries. The data were mainly for 1970 but 1960 figures were used where 1970 data were not available. It should be pointed out, however, that data on the distribution of income refer to the year when the information is available for the country and may not coincide exactly to the year for which the data on income and infant mortality were obtained. He estimated log (IMR) as a linear function of the log of the incomes of the low and middle income groups, share of income of top 5% and a dummy, i.e:

log(IMR) =f(log (poor income), log (middle income), rich share, dummy).

Poor and middle income refer to the imputed absolute per capita income of the lowest fifth and middle strata (20^{th} to 95^{th} percentiles) respectively while rich share is the share of income that goes to the top 5% of the population. The dummy variable equals 1 if the year is 1970, 0 if 1960. He also estimated the equation without the income of the middle strata.

Estimation of the above equation for a sample of both developing and developed countries found the coefficient of rich income share to be consistently and significantly positive. The high share of income for the rich might indicate uneven regional development with the poor concentrated in backward areas with limited access to health care. This is especially so for LDCs with majority of the poor living in rural areas and urban slum areas with poor to no public utilities. This worsens with uneven income distribution where the rich influence the allocation of resources. Waldmann attempted to control for the level of female adult literacy. The positive relationship between share of income of the rich and IMR remained significant and both the coefficient and its t-statistic increased.

Waldmann also attempted to control for the effect of social and cultural variables. Dummies were included to take care of continental differences. This did not, however, remove the relationship established above.

In a more recent study, Pritchett and Summers (1996) use OLS and instrumental variables (IV) techniques to focus on isolating the "pure" effect of income from cross country specific and time trend effects on health. They noted that there are three plausible explanations for the positive relation between health and wealth. First, increases in income leads to better health. Second, healthier workers are likely to be more productive and therefore wealthier. This therefore implies reverse causation between income and health. Third, there may be other factors that may lead to improvements in both health and wealth. They mention cultural values and good government as possibilities. Instrumenting for income produced results similar to those obtained when using the income variable itself in the estimating equation. The estimated coefficients after instrumenting for income were similar though slightly larger than the OLS ones. The similarity of the estimates was evidence in favour of a direct causation of income to health.

Pritchett and Summers used five year changes in terms of trade, ratio of investment to GDP and the black market premium for foreign exchange as income instruments. Generally, finding good instruments is difficult. Although the first two instruments may be considered good instruments, i.e. highly correlated to income but uncorrelated to health, the third one, as Pritchett and Summers admit, may not be. A higher black market premium is often associated with a country's political and economic instability that leads to poor performance of the foreign exchange market. This was observed recently in Uganda and

Mozambique. This has a negative effect on the country's balance of payments. This coupled with the political and economic instability may be highly negatively correlated to health status. Therefore black market premium may not satisfy criteria for a good instrument.

Pritchett and Summers estimated that the long run income elasticity of infant and child mortality in LDCs lies between -0.2 and -0.4. This, they argued, implied that over half a million deaths in LDCs in 1990 were attributable to poor economic performance in the 1980s. They concluded that the income-mortality relationship is not a coincidence or due to reverse causation or some sort of incidental association. They argued that there exists a causal relationship running from income to mortality. In the next section we indicate the proposed extensions to the studies reviewed above and also summarize what the rest of the Chapter will focus on.

3.4 Cross-Sectional Analysis Using Circa 1990 Data:

While Preston focused on life expectancy at birth as the dependent variable, Waldmann used infant mortality rate (IMR). Furthermore Preston puts emphasis on income while Waldmann focuses on its distribution. We attempt to reconcile these two by treating both life expectancy and infant mortality rates as dependent variables. In addition we consider both per capita RGDP and its distribution as key independent variables, as well as other potentially important factors including literacy levels, nutrition, urbanization and availability of doctors.

In this study, we first attempt to replicate Preston and Waldmann's results and then do further analyses. Unlike Preston, however, the income data here refers to purchasing power parity (PPP) adjusted per capita RGDP. This is the same kind of data that Waldmann used. In addition we will compare cross-sectional results from different points in time to investigate any changes in the relationships over time.

Waldmann found a positive association between IMR and share of income of the top 5%. Waldmann's analysis was mainly based on linear relationships between the IMR and shares of income going to the various income groups. We will extend this by testing for significance of income distribution variables assuming the true relationship between income and health status to be nonlinear. Once more in a parallel model we also analyze the relationship between life expectancy and income distribution using cross-sectional data.

As discussed in section 3.3, Pritchett and Summers have established that there is a structural relationship running from income to health. They have also established that there is a strong direct causation from income to health and ruled out that this might be incidental. Our focus will be to estimate this relationship over cross sectional data for selected periods. Pritchett and Summers have established through causality tests that the structural relationship between income and health flows from income to health. For readers interested in issues of causation between income and health, see Pritchett and Summers (1996).

This section will cover cross sectional analysis of the aggregate health status production function. We present descriptive statistics pertaining to the data set. We also estimate bivariate relationships between infant mortality rates and life expectancy at birth to per capita *RGDP* and income distribution. Next we analyze multivariate relations between health and selected socio-economic and demographic indicators. Lastly, we examine structural shifts in the relationship between life expectancy at birth and per capita *RGDP*.

As noted before, the real per capita *RGDP* data are based on Penn World Tables 5.6 RGDPC series. This data set is given in Appendix 2.

3.4.1. Descriptive Analysis

Table 3.2 shows summary statistics for the major variables analyzed. The lowest reported life expectancy was 39 for Guinea Bissau followed by Sierra Leone with 42. The highest was Hong Kong with 84 years.

Chad had the lowest per capita *RGDP* of US $399 with a life expectancy at birth of 47 years. The U.S. had the highest per capita *RGDP* of US $18054 with a life expectancy of 74 years, 10 years less than that of Hong Kong for which, however, per capita *RGDP* was lower by US$ 3205. "Positive outlier" countries as shown in Figure 3.3, with higher life expectancies than would be expected on the basis of per capita *RGDP* (per capita *RGDP*, life expectancy at birth) include: Sri Lanka (2096;71), Paraguay (2128;67), Jamaica (2545;73) and Panama (2888;73). "Negative outlier" countries with lower than expected life expectancies include: Gabon (3958;53), Bangladesh (1390;52) and Namibia (2854;57.5). Caldwell (1993) among others indicated that the low health status (especially in Muslim countries) despite high per capita incomes may be attributed to uneven distribution of income and low female literacy levels. Sri Lanka, China and Vietnam, which had per capita incomes of between US$ 330 and $420, had life expectancies in the range of 66 to 71 years. This contrasted sharply with Saudi Arabia, Libya, Oman and Iraq which had life expectancies of between 61 and 64 with per capita incomes over $1500 with Iraq's over $2500.

It was also observed by Caldwell that the former countries spend less on health than the latter. Sri Lanka, for example, spent only 1.7% of its GNP on health care and had a physician/population ratio of 1:5520 compared to Iraq, which spent 4.6% of its GNP on health care and had also a more favourable physician/population ratio of 1:1740. Sri Lanka's policy of food subsidies, high literacy levels and other social services may have tended to reduce effective inequality from the distribution of income (Sen, 1981). A food subsidy increases nutrition among the poor as opposed to direct income grants that could partly end up being spent on non-health enhancing goods. In addition to the comprehensive food policy, Sri Lanka had a higher level of female illiteracy. Caldwell (1993) associated its good performance with high levels of female literacy and political and social will. Caldwell further reiterates that one additional year of mother's schooling is associated with a reduction in child mortality of 6.8%.

Table 3.3 compares the share of income of the lowest 20% of the population and female illiteracy levels for the positive and negative outliers. It is clear that all the negative outliers have female illiteracy levels way above that of the sample average as compared to the positive outliers. This tends to support Wheeler's (1980) views on the relationship between poor health status, high incomes and low literacy levels. There is not enough information on the share of income of the lowest 20% to draw conclusions.

3.4.2. Infant Mortality Rates, Life Expectancy and Income Inequality

As discussed in section 3.3, Waldmann (1992) investigated the relationship between infant mortality rates and income inequality using a log linear relationship. For comparison of results we retain the Waldmann specification and re-estimate the relationship using circa 1990 data. In a parallel model we also estimate the relationship between life expectancy and income distribution. The sample includes 51 countries for which income distribution data are available. Our variables are defined slightly differently from Waldmann. We use absolute average income of the lowest fifth as did Waldmann, but we use percentage income share of the top 10%, whereas Waldmann used the top 5% from World Bank Tables (1976). After 1976, World Development Reports provide the share of income of the top 10% rather than 5%.

We first attempt to replicate Waldmann's work discussed in Section 3.3. The coefficient of the lowest fifth is defined this way so as to capture the association of infant mortality rates with the absolute standard of living of the poor. The share of the top 10% variable acts as an inequality measure. For comparison, Waldmann's results are given in Table 3.4, column numbers 2 to 5. Estimation of the Waldmann type of model with our data gave the results in column number 6. LNYLOW20 was computed by multiplying per capita RGDP by the share of income of lowest fifth, then multiplying the result by five and taking the natural logarithm. We also dropped the log specification and estimated this relationship using both infant mortality rates and life expectancy as dependent variables. The results are given in columns 7 and 8.

The results show that both an increase in average incomes of the lowest fifth and share of income of the top 10% will result in improved health status as measured by lower infant mortality rates and higher life expectancy at birth. These results seem to differ from Waldmann's, who found that an increase in average income of the poor would lead to declining infant mortality rates, but that an increase in percentage income of the "rich" would lead to increases in infant mortality rates. As noted earlier in this section, however, our definition of the "rich" as the top 10% incorporates, in part, what in Waldmann's study is included in the middle income group.

One explanation for the difference in results could therefore be the difference in cut off points in income groups considered. As shown in Table 3.4, Waldmann found that, controlling for other income groups, increases in average income of the middle group would result in declines to infant mortality rates. This negative relation might be offsetting the positive effect of the upper 5% in our estimates. Note especially that when Waldmann left out the income of the middle group (columns numbers 3 and 5), the coefficient on the share of the top 5%, though still positive, became statistically insignificant. Therefore by using the top 10% rather than the top 5%, we might expect the coefficient to become smaller. Note too that the coefficient on the log of the lowest fifth was insignificant when Waldmann included the middle income, but is negative and significant as in our estimates when the middle income category is excluded. Therefore although there is some difference between Waldmann's results and ours, as shown in Table 3.4, which could be due to the differences in income grouping and time periods, we arrive at some similar conclusions. An increase in absolute income of the lowest income group is associated with an improvement in general health status. The last column shows that life expectancy increases as the share of the top 10% and income of the lowest 5th rise.

3.4.3. Generalizability of Estimates that Include Income Distribution

Of the 104 countries for which we have data on per capita *RGDP*, infant mortality rates (IMR) and life expectancy at birth (EXP) for 1990, we have, as mentioned earlier in section 3.4.2, data on income distribution for 51 of them. This results from the fact that, unlike per capita income, which is derived from annual national accounts, income distribution data are compiled from specialized surveys like the living standards measurement studies and household expenditure surveys.

The sample of 51 countries was, however, approximately representative in terms of coverage. On the basis of per capita *RGDP*, 11 were classified by World Bank Development Report 1993 as low income, 20 as middle income and 20 as high income countries. The data on income distribution refers to estimates obtained between 1981 and 1991. (Four countries with income distribution were not included because we did not have the complete per capita *RGDP* data set for them. These countries are Tanzania, Ethiopia, Nepal and Botswana. One country, Lesotho, was omitted for the unrealistically high reported IMR for certain years as mentioned earlier in section 3.2.)

Figure 3.3 shows a plot of life expectancy and per capita RGDP for 1990. The correlation between the two is 0.76. Clearly the relationship between life expectancy and income is non-linear. For both the entire data set and the group for which we have income distribution data, we estimated the relationship using different specifications. In order to capture clearly the relationship, we fitted a quadratic spline allowing for two quadratics in per capita *RGDP*, continuous and with equal slopes at the knot, which is set at the sample mean per capita *RGDP*. Using the data itself to estimate the knot location complicates variance estimation and testing. We followed the standard practice and set the knot ex ante. We chose the sample mean *RGDP* because the relationship between life expectancy (and infant mortality rate) to *RGDP* seems to taper off at this point (See Figure 3.3).

Before analysing the different model specifications we investigate if there are any differences in the structural relationships between life expectancy and income between the entire set of countries and those that have reported data on income distribution. We adopt the quadratic spline formulation with dummy variables and conduct F tests. The following unrestricted equation was estimated in linear form:

$$EXP = f(D, RGDP, RGDP^2, Z^2, D*RGDP, D*RGDP^2, D*Z^2) \quad (3.2)$$

where as Appendix 1 shows:

D - is a dummy variable which equals 1 when a country reports income distribution data, 0 otherwise.

RGDP, $RGDP^2$ refer to per capita *RGDP* and its square.

Z^2 - quadratic spline term. Defined as $(RGDP\text{-}y)^2$ if *RGDP* is greater than *y*, 0 otherwise, where y = mean *RGDP*.

We estimate a similar equation with IMR as the dependent variable. The sample mean *RGDP(y)* for 1990 was US $5082. This was chosen as the knot for the spline. The variable names *D*RGDP*, $D*RGDP^2$ and $D*Z^2$ are interaction variables between the dummy (*D*) and *RGDP*, $RGDP^2$ and Z^2.

The results for the unrestricted equation are presented in columns 2 and 4 of Table 3.5. The results for the restricted equation are given in columns 3 and 5. The dummy intercept term and the interaction dummy variables are individually insignificant at 5% level of significance. The Chow statistic (calculated as shown below) for the joint significance of the dummy interaction variables, including the dummy intercept term, was 2.26 and 2.00 with life expectancy and infant mortality rate as the dependent variable, respectively. The critical 5% and 1% values for the F-statistic with 4 and 97 degrees of freedom in the numerator and denominator respectively are 2.46 and 3.51. For both life expectancy and infant mortality rate equations we do not reject the null hypothesis that the coefficients on all dummy variables are zero at both 5% and 1% levels of significance.

The Chow test statistic, which is a special version of an *F* test, was calculated as:

$$[(e_r'e_r - e'e)/q]/[e'e/(n-k)] \tag{3.3}$$

where:

$e_r'e_r$ - are residual sum of squares from the model with q restrictions

$e'e$ - are residual sum of squares from the unrestricted model with k coefficients and

n - represents number of observations.

We may conclude from these results that countries with reported income distribution data exhibit a similar structural relation between income and life expectancy as countries for which income distribution data are not available. Similar conclusions can be drawn in the case of infant mortality rates.

3.4.4. Comparison of Functional Forms

We have established that the structural relationship between life expectancy at birth and infant mortality rate to per capita *RGDP* is similar for the subset of countries with income distribution data as with the entire data set. Next we will therefore analyze the entire data set without the income distribution variables. Later we will impose the relationship derived from the entire data set to the subset of countries with the income distribution variables. In both cases our principal focus will be the functional form of the relationship between the two health indicators (life expectancy and infant mortality rates), per capita *RGDP* and its distribution. As stated in section 3.3 and at the beginning of this section, this work is along the lines of that done by Preston and Waldmann, among others. As further mentioned in section 2.3 we later extend our analysis to incorporate other socioeconomic and demographic variables.

We estimated a quadratic spline equation over the entire sample. The following was estimated:

$$h_t = a + b_1 RGDP_t + b_2 RGDP^2_t + b_3 Z^2_t \qquad (3.4)$$

where:

h_t - health indicator at time t and $RGDP, RGDP^2$ and Z^2 are as defined in appendix 1. $Z_t = RGDP_t - y_t$ *if* $RGDP_t > y_t$, 0 otherwise. As indicated earlier under (3.2), y_t is mean per capita *RGDP* at time *t*.

In estimating this quadratic spline equation we impose continuity of the function and its first derivatives at the knot. The first restriction is satisfied by observing that the value of the spline function at both sides of the knot evaluated at *y* reduces to:

$$h_t(y) = a + b_1 y + b_2 y^2$$

since the Z term reduces to 0. Continuity of first derivatives is also satisfied since the first derivative of this function becomes b_1+2b_2y on both sides of the knot.

The results from estimating equation (3.4) indicate that the relationship is best approximated by the quadratic spline as compared to linear, linear spline, semi-log and quadratic specifications, as shown in Table 3.6. In all the specifications the dependent variable is life expectancy at birth. In the linear equation we only consider per capita RGDP as the explanatory variable while we add the linear spline term (Z) in the linear spline estimation. For the semi-log we use the natural log of per capita *RGDP* while for the quadratic specifications we estimate an equation similar to (3.4) but without the quadratic spline term (Z^2). With an adjusted R^2 of 0.80 the quadratic spline fits the data best. The semi-log specification seems to perform almost as well as the quadratic spline. Figure 3.4 shows predicted life expectancy from the quadratic spline equation against per capita RGDP for 1990 and it is evident that the quadratic spline fits the data plotted in Figure 3.3 well.

The spline variable coefficient appears exactly equal in magnitude and with opposite sign to the coefficient of $RGDP^2$ and thus the function is nearly linear to the right of the spline. The results show that the relationship is well depicted by two quadratics continuous and smooth at the mean of per capita *RGDP*. Although the double log model estimated by Pritchett and Summers (1996) has the advantage of imposing constant elasticity, thus making the interpretation of coefficients simpler, it may have shortcomings. It assumes that the relationship can be represented by one log function. This might have been suitable for their data because they only considered LDCs. In this study we include MDCs in the sample. By using a quadratic spline we allow the shape of the relationship to differ between low and high income countries.

We re-estimated the equations with infant mortality rate as the dependent variable. The results are reported in Table 3.7. Figure 3.5 shows predicted infant mortality rates from the spline equation. Figure 3.5 and Table 3.7 seem to complement Figure 3.4 and Table 3.6. Among the specifications considered, none performs better than the quadratic spline that has an adjusted R-square of 0.76. We may thus conclude that the quadratic spline model best approximates the relationship between health status and per capita *RGDP*. Note that as in the case where we use life expectancy in Table 3.6, the coefficient of $RGDP^2$ equals that of Z^2 in magnitude and thus the functions are nearly linear to the right of the spline knot.

Wennemo (1993) argues that the reciprocal equation (i.e. regressing infant mortality rates on the inverse of per capita *RGDP*) best represents the relationship between per capita income and infant mortality rates. Our findings show that the reciprocal model ranks after the quadratic spline and semi-log. It is only a better fit compared to the linear, linear spline and quadratic models. See Table 3.7.

From these results we may infer that there is a diminishing contribution to gains in health status as income rises. This leads to a nonlinear relationship between income and heath status indicators. This result is consistent with arguments made by Preston (1975) and Lindert (1986). This nonlinear relationship is well represented by a quadratic spline formulation.

3.4.5. Multivariate Analysis

Section 3.4.2 dealt mainly with the relationship between health and income distribution along the lines of Waldmann. In section 3.4.3 we focused on the relationship between health and per capita *RGDP* using the approach by Preston among others.

In section 3.4.2 we established that average income of the lowest fifth and share of income of the top 10% are negatively related to infant mortality rate and positively related to life expectancy. We further showed in section 3.4.4 that the relationship between life expectancy and infant mortality rate to per capita *RGDP* can be well represented by a quadratic spline function. In this section we analyze the structural relationship between life expectancy and per capita *RGDP* and its distribution. We add income distribution variables to the quadratic spline in income and estimate the resulting equation over the subset of countries for which we have both income measures. We test for the significance of income distribution, given the non-linear relationship between health and income. The results are reported in Table 3.8.

Note that the results in columns (i) and (iii) are similar to those reported in rows 1 (linear model) and 5 (quadratic spline) in Table 3.6 and consequently lead to similar conclusions about the relationship between income and life expectancy, i.e. the quadratic spline fits the data set best. They differ due to the fact that those in Table 3.6 refer to all countries while the ones referred to in Table 3.8 restricted the data to those countries with income distribution data. The coefficients are, however, comparable. All variables in Table 3.8 are as defined in Appendix 1.

Although the coefficient on SHARETOP10 is positive and just significant, which tends to support Table 3.4 col.8, an F test using (3.3) for the significance of income distribution comparing columns marked (i) and (ii) leads us not to reject the null hypothesis that the income distribution coefficients equal zero. The calculated F-statistic is 2.75 compared to an $F_{0.05}(2,47)$ critical value of 3.2. Column (iii) gives results for the quadratic spline specification over the subset of 51 countries. Column (iv) shows results of the equation in which income distribution is added to the quadratic spline. An F test using the results in columns (iii) and (iv) reaffirms the result that income distribution variables are insignificant, with a calculated F statistic of 0.37 compared to the critical value of 3.2, and both income distribution measures are individually statistically insignificant. Adding the two distribution measures with the spline term does not alter the adjusted R^2.

We extend our analysis and incorporate other variables indicated in the reduced form aggregate health status production function in section 2.2. The results are reported in Table 3.9 for both life expectancy and infant mortality rates. Results under the two columns marked (i) focus on the key variables in the reduced form equation (i.e. income measures, population/doctor ratio and illiteracy level) while those under the columns marked (ii) incorporate in addition percentage urban population (URB) and daily calorie supply (DCS). These results are estimated over the same sample. The constraining variable was DCS. Data on daily calorie supply was taken from the UNDP Human Development Report, 1992. The report did not give DCS data for the more developed countries. We find the two variables, URB and DCS, to be jointly insignificant and indeed incorporating them reduces the adjusted R-square for both IMR and EXP. We therefore dropped URB and DCS from the equation and re-estimated the equation implied by the columns marked (i) with MDCs now included and the results are reported in the two columns marked (iii).

Coefficients of income per capita and its square plus that of the spline term in columns (i)-(iii) for both life expectancy and infant mortality indicate the nonlinear relationship between income and health status discussed in previous sections. The coefficient on POPDOC and that on TILLIT have the expected negative sign for the life expectancy estimates. Countries with high population/doctor ratios and those with high levels of illiteracy have lower life expectancy at birth. This is consistent with Caldwell (1993). The results also seem to point to a positive relationship between population/doctor ratio and infant mortality though the coefficient is statistically insignificant. This could be as a result of costs involved in accessing a physician. Individuals may tend to postpone visits to physicians until their condition gets worse. High levels of illiteracy are unambiguously associated with high infant mortality rates and lower life expectancy at birth. We therefore note that countries with high population/doctor ratios and those with high levels of illiteracy are likely to have lower health status.

In the next section we undertake an exercise to examine the effects of varying income distribution at different levels of income on life expectancy. This, it should be emphasized, is a hypothetical numerical exercise as opposed to the econometric estimation undertaken before. The purpose of this exercise is to simulate the effects on life expectancy of moving from an unequal income distribution to one where each individual receives the nation's per capita *RGDP*.

3.4.6. Demonstrating the Effect of Income Distribution on Life Expectancy

We established in section 3.4.4 that the relationship between life expectancy and per capita *RGDP* is well represented by a quadratic spline. In section 3.4.5 we further established that when we introduce income distribution measures in the quadratic spline equation it had no significant effect and the adjusted R^2 did not change (Table 3.8). Using the results of the quadratic spline for 1990 with life expectancy as the dependent variable from section 3.4.4 as reported in Table 3.6 Row 6, we now investigate the "size" of the effect of changing income distribution on health at different income levels. We investigate this effect by using earlier results that showed that there is non-linearity in the relationship between income and health. This will be done by constructing hypothetical countries containing two income groups. Ideally this exercise should use income/life expectancy relationships estimated with micro data, but with this data being unavailable for most countries we use the aggregate data. We assume that this income/life expectancy relationship estimated from aggregate data also applies to individuals within the hypothetical countries.

We make the following assumptions:

(i) Two income groups: low (L) and high (H)

(ii) Proportion of populations in each income group is p^L and p^H such that $p^L + p^H = 1$. We consider the case where $p^L = 0.7$ and $p^H = 0.3$

(iii) Average incomes in each of the income groups is given by y^L and y^H respectively such that $p^L y^L + p^H y^H = Y$

where *Y* is average income for the country. The average life expectancy is given by

$$p^L exp^L + p^H exp^H = e^U$$

where exp^L and exp^H are predicted average life expectancies from the quadratic spline for the low and high income respectively. The estimated quadratic spline using the 1990 data set in section 3.4.4 and in Table 3.6 is given by:

$$\underset{(30.16)}{EXP = 43.49} + \underset{(10.14)}{1.07RGDP} - \underset{(-7.12)}{0.0099RGDP^2} + \underset{(5.55)}{0.0099Z^2} \quad (3.5)$$

To illustrate the effect of income distribution we take three hypothetical levels of per capita *RGDP* and investigate the effect of varying the income distribution across the two income groups L and H. We consider a hypothetical country in each income group.

We assume that the top 30% in each of the three countries receive 70% of the country's *RGDP*. Countries A, B and C have an average per capita *RGDP* of \$600, \$5000 and \$18000 respectively. The average income of an individual that belongs to each of the two income groups is calculated using a procedure similar to the one used to get YLOW20, described in Appendix 1. The estimated average incomes and life expectancies are given in Table 3.10.

The life expectancy estimates in the columns with sub-title unequal distribution are predicted by the spline equation for each individual assuming the uneven income distribution where top 30% get 70% of income. The values under column exp_U are population weighted average life expectancies for the two income groups. The life expectancy values (exp_E) under the column "equal distribution" represent values when we assume a distribution where each individual receives her country's per capita *RGDP* such that $y^L = y^H = Y$.

The last column represents changes in average life expectancy associated with moving from the unequal distribution to an equal distribution. Income distribution seems to have the greatest impact for the middle income countries. There is a gain of almost six years in Country B. At the low income levels (Country A) there is also a slight increase of 0.2 years by moving towards a more equal income distribution while at the high income levels there does not seem to be any changes in life expectancy resulting from re-distribution of income. These results reflect the concavity of the income/life expectancy relationship established earlier.

In order to further examine the effect of income distribution in the middle income range we restricted our sample to those countries whose per capita *RGDP* falls within the middle 50% of the income continuum, i.e. between the 25^{th} and 75^{th} percentiles. This was because of the above exercise which seems to indicate that income distribution may matter more in the middle income range. These countries are closer to the bend in the quadratic spline relationship shown in Figure 3.3.4, therefore the income distribution might matter most for these countries. As indicated earlier in 3.4.4, Figure 3.3.4 shows that at lower and at very high incomes the relationship between income and life expectancy is almost linear. It is almost vertical at low incomes and flat at high incomes. The per capita *RGDP* range for the "middle" income countries was 1182 (Zimbabwe) to 7478 (Portugal) for a total of 78 countries out of which we have income distribution data for 27 of them. Ideally we should be considering only those countries classified as middle income by World Bank classification. However, we are constrained by income distribution data availability.

We estimated equations similar to those implied by Table 3.8 on the 27 countries. The results are reported in Table 3.11 for both life expectancy and infant mortality rate. Income distribution variables are insignificant for life expectancy and infant mortality rate equations for both linear and quadratic spline equations. Note too that unlike in Tables 3.6, 3.7 and 3.8 when the coefficient of the spline term (Z^2) was almost equal to that of the square income term ($RGDP^2$), for this set of countries it turns out to be insignificant. This is consistent with our earlier assertion that in this range the income-health relationship is concave. We may thus conclude that, although redistributing income, especially in the middle income zone, may influence the health status outcome, our data set does not seem to find any significant effect of income distribution on either life expectancy or infant mortality rate.

In concluding, we observe that at low incomes, income distribution seems to matter only a little as the relationship between income and health is almost vertical. At very high incomes, the relationship is almost flat and again, income distribution does not seem to matter. If income distribution is important anywhere, it should be in the middle income range where the health/income relationship is the most nonlinear. The results in the exercise above seem to indicate that income distribution may matter in this middle income range. However, further investigation using our data set led us to the finding that we cannot reject the null hypothesis that the coefficients of income distribution measures are jointly zero. It should be noted that although the smaller sample size may reduce the power of the tests for the significance of the inequality variables, selecting these particular 27 countries may increase the power as their per capita *RGDP* puts them in the most curved, or least linear, part of the per capita *RGDP*-life expectancy relationship.

3.5. Time Series Aggregate Analysis:

In section 3.4 we focused on the cross sectional results pertaining to the circa 1990 data. We reviewed the descriptive results in Table 3.2 and later through regression estimates established that the relationship of per capita *RGDP* to life expectancy and infant mortality rate is well represented by a quadratic spline model. We also showed that income distribution did not seem to matter if we took nonlinearity into consideration. In this section, after presenting the descriptive statistics for 1970 and 1980, we estimate the relationship over the two data sets for 1970 and 1980. This section will be organized as follows. In subsection 3.5.1 we analyze population weighted trends in life expectancy and infant mortality rates. Subsection 3.5.2 will focus on descriptive statistics for 1970 and 1980 and quadratic spline model estimation with the two data sets. This is followed in subsection 3.5.3 by regression estimates to show how changes in life expectancy and infant mortality rate relate to changes in per capita *RGDP*. It should be reiterated that we use the same set of 104 countries as in the 1990 cross sectional analysis covered in section 3.4.

3.5.1. Trends in Infant Mortality and Life Expectancy

In this section we focus on trends in the two health indicators for the period 1965-91. This is the period that the World Bank data set in this study covers. We consider trends in population weighted and unweighted life expectancies at birth and infant mortality rates for this period. As mentioned above the sample of countries is restricted to the 104 countries analysed in section 3.4. For the unweighted analysis the figures are derived by summing up life expectancies or infant mortalities across all countries and dividing by the number of countries. The weighted results also average across all countries, weighting by the country's population.

Table 3.12 gives the calculated population weighted (*pexp & pimr*) and unweighted (*uexp & uimr*) estimates for life expectancy and infant mortality rates. There is no doubt overall that there has been a general improvement in health status over this period. The number of infants dying before age 1 per thousand live births has declined from about 90 in 1965 to 50 in 1991 while life expectancy at birth increased from about 55 years in 1965 to about 65 in 1991.

Table 3.12 further shows that population weighted infant mortality rates are lower than unweighted infant mortality rates except in 1991. On the other hand population weighted life expectancy is lower than unweighted except for 1990. There is however a strange behaviour in population weighted life expectancy especially over the period 1987-1991. Between 1987 and 1990 *wexp* dropped by over 1 year. It then jumped up by almost two and a half years from 63.7 in 1990 to an all time high of 66.3 in 1991.

We examined our data set more carefully and this strange pattern seemed to be due to the reported life expectancy for China. Life expectancy for China covering the period 1979-1991 was reported as follows: 1979-1981;67, 1982-1985;68; 1986-1987;69, 1988;66, 1989;64, 1990;62 and 1991;70. China's population is over 1 billion. We re-estimated life expectancy/infant mortality trends excluding China. The results are reported in Table 3.13. We now observe that the sudden drop and jump in *wexp* disappears. The sudden decline in life expectancy for China between 1988 and 1990 combined with its high population no doubt explains the behaviour of *wexp* in Table 3.12. This decline had a major impact on the numerator thus leading to a "temporary" decrease in the overall *wexp*. In conclusion, we note that the main trend is a rapid reduction in infant mortality rates and an increase in life expectancy.

3.5.2. Cross Sectional Analysis: 1970 and 1980 data

The variables upon which we focus in this section are those that were found to be significantly associated with health indicators as summarized in Table 3.9. Descriptive statistics covering these variables are given in Table 3.14.

Sierra Leone had the lowest reported life expectancy at birth of 34.3 years in 1970 while the highest was Sweden, 74.5 years. In the same year the infant mortality rate was highest in Mali (204 deaths per 1000 live births). The lowest infant mortality was 11 for Sweden. Ten years later in 1980 Guinea Bissau had the lowest life expectancy of 37 years while the highest was Iceland with about 77 years. Japan and Sweden ranked second with life expectancy of 76 years while Sierra Leone was next lowest to Guinea Bissau with life expectancy of 38 years. In the same year, 1980, Mali and Sweden again had the highest (184) and lowest (7) infant mortality rates respectively. Burkina Faso (previously known as Upper Volta) was the poorest country in both years as measured by per capita *RGDP*. It had a per capita *RGDP* of about US $370 in 1970 and 460 in 1980 while the US was the highest income nation in both years with per capita *RGDP* of US $12,963 and $15,295 in 1970 and 1980 respectively.

In terms of adult illiteracy Burkina Faso had the highest illiteracy rate in both 1970 and 1980 (95 % in both) followed by Chad in 1970 (93%) and Mali and Senegal in 1980 (90%). In 1980 Chad's illiteracy rate, although still high, had fallen by 8 percentage points to 85%. The results for Burkina Faso following the World Bank definition of illiteracy show that in both years 95% of the Burkina Faso population over 15 years could not with understanding read and write a short, simple statement on their everyday life (World Bank 1993, p.308). Various developed countries had a 1% illiteracy rate in 1970 and also 1980. However, Finland and Australia had 0% illiteracy rates in 1980. POPDOC, the number of people per doctor, was highest for Burkina Faso (97,120 and 48,150 in 1970 and 1980 respectively). The lowest was Israel in 1970 (410) and Italy in 1980 (340).

We will first estimate the restricted equation for life expectancy and infant mortality rate with only income measures and the quadratic spline term as independent variables over the 1970 and 1980 data sets. We then extend the model to incorporate other variables as shown in Table 3.14. The results given in Table 3.15 under columns marked (i) and (iii) for the restricted equations are similar to those reported in Tables 3.6, 3.7 and 3.8. The relationship of both life expectancy and infant mortality rate with per capita *RGDP* is well represented by the quadratic spline equation. Figures 6 and 7 show predicted life expectancies from the quadratic spline estimated equations plotted against per capita *RGDP* for 1970 and 1980 respectively.

The results for the extended equation incorporating population/doctor ratio (POPDOC) and total adult illiteracy (TILLIT) are given under columns marked (ii) and (iv) in the same Table 3.15. They complement the results for 1990 reported in Table 3.9. Coefficients of per capita *RGDP* are once again significant in both years for life expectancy. For the case of infant mortality rates, the coefficient on per capita *RGDP* becomes insignificant when we incorporate POPDOC and TILLIT for both years. Once again TILLIT seems to have a profound impact on both life expectancy and infant mortality rate. High illiteracy levels are associated with low life expectancy and high infant mortality rates.

3.5.3 Effect of Change in Per Capita *RGDP* on Changes in Health Indicators

Following similar work done by Preston that was discussed in section 3.3, we now focus on how changes in per capita *RGDP* have influenced changes in health. We consider three period intervals 1970-1980, 1980-1990 and 1970-1990. We estimate an equation similar to (3.1) in section 3.3:

$$E = a + b_1X_1 + b_2X_2 + b_3X_3,$$

where E = Absolute change in life expectancy between the two time periods.

X_1 - Level of per capita *RGDP* for the base period

X_2 - Absolute change in per capita *RGDP* between the two periods.

X_3 - Ratio of per capita *RGDP* - end period to base period.

We estimate this equation bearing in mind that with both the ratio of per capita *RGDP* (X_3) and its absolute change (X_2) in the equation, it is hard to interpret their coefficients. However we estimate it this way to compare the results to Preston's. Later we leave out absolute change in per capita *RGDP* and re-estimate the equation.

The results including absolute change in per capita *RGDP* are given in Table 3.16. For comparison Preston's results are also given in the last column. Percentage growth of income, as measured by the ratio of end period per capita *RGDP* to initial per capita *RGDP*, matters for two periods, 1970-1980, and 1970-1990. It has a significant positive influence on changes in life expectancy. Initial level of income is significant for the 1980-90 period and with a negative coefficient for all three periods. Absolute change in income is only significant for the period 1970-1980 with a negative coefficient. In Preston's results, given in the last column of Table 3.16, only initial per capita *RGDP* was significant and also with a negative coefficient. Note that in previous sections when we used the spline model to investigate the effect of income on both life expectancy and infant mortality rates we found that income as measured by per capita *RGDP* was positively associated with current levels of life expectancy. We may thus speculate that though the level of income is a good indicator of the level of life expectancy and infant mortality, faster gains in life expectancy were achieved by countries that had lower per capita *RGDP* in the base period. Anand and Ravallion (1993) suggest that the positive effect of faster economic growth on health is achieved through poverty reduction and increased public spending on health.

We dropped X_2 from the equation and estimated the following equation:

$$\%\ change\ in\ exp = f(level\ of\ RGDP\ (X_1),\ ratio\ of\ RGDP\ (X_3)).$$

As noted earlier in this section, with both X_3 and X_2 in the equation it is difficult to interpret the results in Table 3.16. This is the reason we dropped X_2. The results are reported in Table

3.17 and reinforce those reported in Table 3.16 as well as Preston's, reported in that same table. Lower initial levels of income are associated with higher changes in life expectancy. Our estimates that cover a different time period from Preston's and use PPP adjusted per capita *RGDP* are similar to Preston's. Initial income level is negatively associated with changes in life expectancy. This confirms the general observed trend of faster gains in life expectancy among those countries that had lower per capita *RGDP* in the base period.

We further compared the structural relationship between life expectancy and per capita *RGDP* for the three time periods, 1970, 1980 and 1990. The results show that comparing the structural relationship as per Figures 3.4, 3.6 and 3.7 (drawn to same scale) there is evidence of an upward shift in the relationship between life expectancy and income between 1970, 1980 and 1990. The shift seems more evident at middle incomes. A country with an average per capita *RGDP* of US $4000 had a life expectancy of about 65 years in 1970, 68 in 1980 and almost 70 in 1990 (See Figure 3.8). Note also that the point when the income-life expectancy curve flattens out also seems to have shifted upwards and to the right, due in part to an increase in what could be regarded as the maximum attainable life expectancy during this period.

In conclusion to this chapter we note that there exists a quantitatively important and statistically significant relationship between life expectancy and infant mortality rates and per capita *RGDP*. The relationship is well represented by a quadratic spline model which shows diminishing contribution of income to health status. This does not, however, diminish the importance of income in policy formulation geared at improving health. Income has the greatest impact in the lower income range where the relationship is linear and almost vertical. Increasing income will therefore raise health status for this group. Our simulation exercise showed that redistributing income from rich to poor may improve overall health status in the middle income range. We now move to determinants of child health at the household level. Chapter IV will focus on these micro based estimations.

TABLE 3.1

POPULATION, ECONOMIC INDICATORS AND PROGRESS IN HEALTH BY DEMOGRAPHIC REGION, 1975 - 1990

Region	Population 1990 (millions)	Deaths 1990 (millions)	Income per capita		Child Mortality		Life expectancy at birth (years)	
			Dollars 1990 (millions)	Growth rate, 1975-90 (% per year)	1975	1990	1975	1990
Sub-Saharan Africa	510	7.9	510	-1.0	212	175	48	52
India	850	9.3	360	2.5	195	127	53	58
China	1,134	8.9	370	7.4	85	43	56	69
Other Asia and Islands	683	5.5	1,320	4.6	135	97	56	62
Latin America and the Caribbean	444	3.0	2,190	-0.1	104	60	62	70
Middle Eastern Crescent	503	4.4	1,720	-1.3	174	111	52	61
Formerly socialist economies of Europe (FSE)	346	3.8	2,850	0.5	36	22	70	72
Established market economies (EME)	798	7.1	19,900	2.2	21	11	73	76
Demographically Developing group[a]	4,123	39.1	900	3.0	152	106	56	63
FSE and EME	1,144	10.9	14,690	1.7	25	15	72	75
World	5,267	50.0	4,000	1.2	135	96	60	65

Source:-World Bank (1993), p.2

Note: Child mortality defined as probability of dying between birth and age 5, expressed per 1,000 live births; life expectancy at birth is the average number of years that a person would expect to live at the prevailing age-specific mortality rates.

[a]These countries include, Sub-Saharan Africa, India, China, Other Asia and Islands, Latin America and Caribbean, and Middle Eastern Crescent.

Table 3.2

SUMMARY STATISTICS Circa 1990

VARIABLE	OBS	MEAN	MEDIAN	MINIMUM	MAXIMUM
RGDP90 Real GDP (per capita)	104	5082	2888 (Panama)	399 (Chad)	18054 (USA)
EXP -' life expectancy at birth (e_o)	104	64	66 (Syria, ,Ecuador)	39 (Guinea Bissau)	84 (HongKong)
IMR - infant mortality rate/1000live births	104	54	46 (Cape Verde, Paraguay)	4 (Iceland)	166 (Mali)
SHARETOP10 % share of top 10%	51	31	27 (Yugoslavia)	21 (Hungary, Sweden)	68 (Brazil)
SHARELOW20- % share of lowest 20%	51	6	6 (Jamaica, Israel)	2 (Panama,)	11 (Hungary)
TILLIT - total illiteracy	97	27	19 (Turkey, Nicaragua,Brazil)	1 (Various MDCs)	82 (Burkina Faso)
FILLIT .- (exludes most MDCs) –female illiteracy	69	43	41 (Kenya)	1 (Jamaica)	91 (Burkina Faso)
ACCESS - % popn with access to HC	77	75	80 (Various mid inc. countries)	15 (Mali)	100 (Various MDCs)
WATER - % popn access to safe water	88	71	74 (China)	12 (Centr. Afr. Rep.)	100 (Various MDCs)
URB - % of urban population	106	51	50 (Syria, Zambia)	6 (Burundi)	100 (Singapore)
AGR - % share of agric. in GDP	91	19	16 (Costa Rica, Turkey, Colombia)	0.28 (Singapore)	59.2 (Mozambique)
DCS (excludes MDCs) - Daily calorie supply	67	103	102 (Cote D'Ivoire, Gabon)	70 (Mozambique)	135 (Mexico)
POPDOC Popn./doctor ratio	60	7113	1345 (Egypt,Korea)	210 (Italy)	72990 (Rwanda)
POP - Population (millions)	104	43	10 (Greece)	.26 (Iceland)	1134 (China)
YLOW20 - Absolute average income of lowest 5th	51	2112	1179 (Thailand, Malaysia)	123 (Kenya)	6234 (Japan)

Source:- See appendix 1 for data sources and detailed variable definitions.

Table 3.3

FEMALE ILLITERACY (FILLIT) AND SHARE OF INCOME OF LOWEST FIFTH (SHARELOW20) FOR OUTLIER COUNTRIES*

COUNTRY	OUTLIER (+/-)	FILLIT	SHARELOW20
Sri Lanka	+	17	5
Paraguay	+	12	Unavailable
Panama	+	12	2
Jamaica	+	1	6
Gabon	-	52	Unavailable
Bangladesh	-	78	Unavailable
Namibia	-	Unavailable	Unavailable
Sample mean		43	6

*See appendix 1 for definition

Table 3.4

EFFECT OF INCOME DISTRIBUTION ON INFANT MORTALITY AND LIFE EXPECTANCY

	log imr						
	WALDMAN[1]				LNIMR	IMR	EXP
	ALL		LDCs				
YLOW20	-	-	-	-	-	-0.016 (-7.56)	0.0043 (8.03)
LNYLOW20	0.07 (0.45)	-0.66 (-6.95)	0.01 (0.04)	-0.50 (-3.54)	-0.96 (-12.93)	-	-
LNYMIDDLE75	-0.80 (-5.57)	-	-0.69 (-3.18)	-	-	-	-
SHARETOP5	2.48 (2.82)	1.09 (1.03)	2.43 (2.32)	1.16 (1.07)	-	-	-
SHARETOP10	-	-	-	-	-0.03 (-3.13)	-1.14 (-2.74)	0.34 (3.19)
D	0.06 (0.44)	0.08 (0.53)	-0.05 (-0.25)	-0.20 (-0.95)	-	-	-
CONSTANT	11.92 (13.2)	9.48 (9.55)	11.41 (7.7)	8.42 (16.58)	10.73 (14.78)	103.08 (6.41)	49.57 (12.00)

Notes:

t-statistics in parentheses

IMR, LIMR and EXP are infant mortality rate, log of infant mortality rate and life expectancy at birth respectively. See Appendix 1 for variable definitions and data sources.
D-Dummy variable which equals 1 if income distribution data is for 1970, 0 if 1960.
[1] Waldman, R.J.(1992) "Income distribution and infant mortality," *Quarterly Journal of Economics*, (107:4), p128,Table II.

TABLE 3. 5

RESULTS FOR CHOW TEST ON INCOME DISTRIBUTION DATA AVAILABILITY, 1990

	Life Expectancy (EXP)		Infant Mortality Rates (IMR)	
	Unrestricted	**Restricted**	**Unrestricted**	**Restricted**
RGDP	0.96 (6.86)	1.07 (10.14)	-3.72 (-6.02)	-4.26 (-9.27)
$RGDP^2$	-0.0087 (-4.31)	-0.01 (-7.12)	0.031 (3.54)	0.04 (6.49)
Z^2	0.0084 (3.11)	0.010 (5.55)	-0.028 (-2.34)	-0.039 (-5.00)
D	2.52 (0.76)	-	-13.03 (-0.89)	-
DRGDP	0.074 (0.33)	-	-0.31 (-0.31)	-
$DRGDP^2$	-0.0011 (-0.37)	-	0.0073 (0.56)	-
DZ^2	0.0012 (0.31)	-	-0.0106 (-0.64)	-
Const.	43.39 (25.87)	43.49 (30.16)	136.11 (18.45)	136.00 (21.55)
RSS	2271	2482	43940	47570

Notes: RGDP- per capita RGDP.
$RGDP^2$- square term for per capita RGDP.
Z^2 - This is the square term for the spline term as defined in appendix 1
D - Dummy variable which equals 1 if income distribution data is available, 0 otherwise
t - statistics are in parenthesis

Table 3.6

LIFE EXPECTANCY AND REAL PER CAPITA RGDP - VARIOUS SPECIFICATIONS, 1990

	CONSTANT	RGDP	$RGDP^2$	LNRGDP	Z	Z^2	R^2 ADJ.
Linear Model	55.81 (55.52)	0.16 (11.76)	-	-	-	-	0.57
Linear Spline	48.66 (44.39)	0.51 (12.81)	-	-	-0.49 (-8.98)	-	0.76
Semi-log	33.88 (21.11)	-	-	8.97 (19.83)	-	-	0.79
Quadratic	49.36 (44.22)	0.53 (11.48)	-0.0023 (-8.13)	-	-	-	0.74
Quadratic Spline	43.49 (30.16)	1.07 (10.14)	-0.0099 (-7.12)	-	-	0.0099 (5.55)	0.80

Notes: t statistics in parentheses.
For variable names, definitions and data sources see Appendix 1.

Table 3.7

INFANT MORTALITY RATES AND PER CAPITA RGDP-VARIOUS SPECIFICATIONS, 1990

	CONS-TANT	RGDP 90	RGDP 90^2	LNRGDP 90	Z	Z^2	1/RGDP	R^2 ADJ.
Linear Model	86.14 (20.28)	-0.64 (-10.79)	-	-	-	-	-	0.53
Linear Spline	115.83 (24.66)	-2.06 (-12.14)	-	-	2.02 (8.70)	-	-	0.73
Semi-log	173.20 (24.85)	-	-	-35.36 (-18.01)	-	-	-	0.76
Quadratic	112.88 (23.63)	-2.14 (-10.89)	0.0096 (7.88)	-	-	-	-	0.70
Quadratic Spline	136.00 (21.55)	-4.26 (-9.27)	0.0039 (6.78)	-	-	-0.04 (-0.039)	-	0.76
Reciprocal	14.04 (4.39)	-	-	-	-	-	685.02 (17.28)	0.74

Notes: t statistics in parantheses

See Appendix 1 for variable names, definitions and data sources

Table 3.8

TESTING SIGNIFICANCE OF INCOME DISTRIBUTION IN THE PRESENCE OF NON-LINEARITY

Dep. Variable-life expectancy n = 51

	(i)	(ii)	(iii)	(iv)
RGDP	0.12 (8.22)	0.10 (2.20)	1.04 (7.37)	0.99 (6.41)
$RGDP^2$	-	-	-0.0098 (-5.65)	-0.0092 (-4.81)
Z^2	-	-	0.0096 (4.66	0.0089 (3.93)
YLOW20	-	0.001 (0.67)	-	0.000021 (0.02)
SHARETOP10	-	0.25 (2.26)	-	0.059 (0.71)
Constant	60.97 (47.52)	52.02 (12.59)	45.91 (20.39)	44.52 (14.04)
RSS	1631	1460	690	679
R^2 Adj.	0.57	0.60	0.81	0.81

Notes: t - statistics in parentheses
Abbreviations are as defined in Table 3.5 and Appendix 1.

Table 3.9

EFFECT OF VARIOUS SOCIOECONOMIC INDICATORS ON LIFE EXPECTANCY AND INFANT MORTALITY RATES.

	LIFE EXPECTANCY (EXP90)			INFANT MORTALITY (IMR90)		
	(i)	(ii)	(iii)	(i)	(ii)	(iii)
RGDP90	0.84 (5.60)	0.93 (4.45)	0.66 (5.86)	-3.19 (-3.50)	-4.17 (-3.39)	-2.62 (-4.84)
$RGDP90^2$	-0.0093 (-4.12)	-0.010 (-3.69)	-0.0063 (-4.72)	0.035 (2.61)	0.046 (2.79)	0.024 (3.83)
Z^2	0.012 (3.13)	0.013 (3.00)	0.0064 (4.04)	-0.046 (-2.02)	-0.060 (-2.32)	-0.238 (-3.13)
POPDOC	-0.0099 (-2.23)	-0.00011 (-2.21)	-0.011 (-2.64)	0.027 (1.03)	0.039 (1.32)	0.032 (1.57)
TILLIT	-0.11 (-3.27)	-0.11 (-3.02)	-0.12 (-3.81)	0.65 (3.13)	0.60 (2.83)	0.66 (4.49)
DCS	-	-0.042 (-0.62)	-	-	0.47 (1.15)	-
URB90	-	-0.013 (-0.31)	-	-	0.096 (0.39)	-
Constant	53.64 (19.94)	57.22 (9.54)	55.50 (22.67)	87.97 (5.45)	50.17 (1.42)	82.00 (7.00)
R^2 Adj.	0.88	0.87	0.90	0.78	0.77	0.87
OBS.	38	38	59	38	38	59
RSS	307	302	509	11057	10549	11675

Notes: t - statistics in parenthesis
See Appendix 1 for detailed variable definitions.

TABLE 3.10

SIMULATION OF EFFECT OF INCOME DISTRIBUTION ON LIFE EXPECTANCY IN THE PRESENCE OF NON-LINEARITY[1]

Country	Per capita RGDP (Y)	Per capita RGDP top 30% (y^H)	Per capita RGDP other 70% (y^L)	Unequal distr.			Equal distr.	Change in exp.
				exp_H	exp_L	exp_U	exp_E	$exp_E - exp_U$
A	600	1,400	257	56.5	46.2	49.3	49.5	0.2
B	5,000	11,667	2,143	76.7	61.8	66.3	72.0	5.7
C	18,000	42,000	7,714	97.9	73.9	81.1	81.1	0.0

[1] The data used for constructing this Table are hypothetical.

TABLE 3.11

DEMONSTRATING THE EFFECT OF INCOME DISTRIBUTION IN THE MIDDLE INCOME RANGES, CIRCA 1990.

n=27

VARIABLE	LIFE EXPECTANCY				INFANT MORTALITY RATE			
	(i)	(ii)	(iii)	(iv)	(i)	(ii)	(iii)	(iv)
RGDP	0.276 (4.40)	0.293 (2.41)	1.48 (4.29)	1.45 (3.82)	-1.164 (-3.87)	-0.980 (-1.63)	-4.877 (-2.71)	-5.14 (-2.59)
$RGDP^2$	-	-	-0.018 (-3.32)	-0.017 (-2.79)	-	-	0.053 (1.83)	0.056 (1.73)
Z^2	-	-	0.063 (0.97)	0.051 (0.68)	-	-	0.122 (0.36)	0.104 (0.26)
YLOW20	-	-0.001 (-0.28)	-	-0.001 (-0.32)	-	-0.006 (-0.33)	-	0.0003 (0.016)
SHARE TOP10	-	0.077 (0.54)	-	-0.011 (-0.09)	-	-0.344 (-0.49)	-	0.178 (0.27)
CONST.	57.94 (27.93)	55.53 (10.83)	41.91 (8.93)	42.96 (6.61)	80.44 (8.08)	91.58 (3.63)	132.35 (5.40)	129.76 (3.82)
R^2ADJ.	0.41	0.40	0.62	0.59	0.35	0.30	0.50	0.46
RSS.	562.57	529.60	336.37	333.91	12965	12832	9179	20732

Variables are as defined in Appendix 2 and t-statistics are in parentheses.

Table 3.12

Trends in infant mortality rate and life expectancy at birth 1965-1991
China included

YEAR	uexp	uimr	wexp	wimr
1965	55.7	96.7	56.3	93.4
1966	56.1	94.6	57.1	90.8
1967	56.5	92.6	57.9	88.7
1968	56.8	90.7	58.2	86.6
1969	57.2	88.8	58.7	84.5
1970	57.5	87.0	59.1	82.3
1971	57.9	84.9	59.5	80.3
1972	58.2	82.8	60.0	77.8
1973	58.6	80.9	60.4	75.7
1974	59.0	78.9	60.8	73.4
1975	59.4	77.0	61.2	71.8
1976	59.7	74.9	61.5	69.9
1977	60.1	72.9	61.9	68.0
1978	60.5	71.1	62.3	67.0
1979	60.8	69.4	62.6	65.1
1980	61.2	67.8	63.0	64.6
1981	61.5	66.1	63.3	62.9
1982	61.9	64.5	63.7	61.1
1983	62.2	63.1	63.9	60.5
1984	62.5	61.5	64.2	58.6
1985	62.9	60.0	64.4	57.2
1986	63.2	58.5	64.7	56.2
1987	63.6	57.0	65.0	55.2
1988	63.8	55.9	64.6	54.5
1989	64.0	54.8	64.2	53.8
1990	64.1	53.7	63.7	53.1
1991	64.6	52.2	66.3	52.3

uexp = Unweighted life expectancy at birth
uimr = Unweighted infant mortality rate
wexp = Population weighted life expectancy at birth
wimr = Population weighted infant mortality rate

Table 3.13

Trends in infant mortality rate and life expectancy at birth 1965-1991, Excluding China.

Year	uexp	uimr	wexp	wimr
1965	55.7	96.7	56.7	94.6
1966	56.1	94.7	57.0	93.1
1967	56.4	92.7	57.3	91.4
1968	56.8	90.8	57.5	90.0
1969	57.1	88.9	57.8	88.5
1970	57.5	87.1	58.1	87.0
1971	57.8	85.1	58.5	85.4
1972	58.2	83.0	58.8	83.9
1973	58.6	81.1	59.1	82.7
1974	58.9	79.1	59.5	81.5
1975	59.3	77.3	59.9	80.4
1976	59.7	75.2	60.2	79.2
1977	60.1	73.2	60.6	78.0
1978	60.4	71.4	60.9	76.2
1979	60.8	69.7	61.3	74.3
1980	61.1	68.0	61.5	72.4
1981	61.5	66.3	61.9	70.6
1982	61.8	64.7	62.2	68.7
1983	62.1	63.3	62.5	67.2
1984	62.5	61.8	62.8	65.7
1985	62.8	60.2	63.1	64.1
1986	63.2	58.7	63.4	62.6
1987	63.5	57.1	63.7	61.0
1988	63.8	56.0	64.0	60.1
1989	64.0	55.0	64.2	59.1
1990	64.2	53.9	64.3	58.2
1991	64.5	52.3	64.9	57.1

uexp = Unweighted life expectancy at birth
uimr = Unweighted infant mortality rate
wexp = Population weighted life expectancy at birth
wimr = Population weighted infant mortality rate

Table 3.14

SUMMARY STATISTICS 1970 & 1980

Variable Name	Mean		Minimum		Maximum	
	1970	1980	1970	1980	1970	1980
EXP Life Expectancy at Birth	57.5	61.2	34.3 (Sierra Leone)	37 (Guinea Bissau)	74.5 (Sweden)	76.6 (Iceland)
IMR Infant Mortality Rate (per 1000 live births)	87.1	66.9	11 (Sweden)	7 (Sweden)	204 (Mali)	184 (Mali)
RGDP- Per Capita RGDP	3,425	4,436	374 (Burkina Faso)	457 (Burkina Faso)	12,963 (USA)	15,295 (USA)
TILLIT. Total Illiteracy (% of pop. over 15 years)	34.03	33.8	1 (Various MDCs)	0.0 (Finland, Australia)	95 (Burkina Faso)	95 (Burkina Faso)
POPDOC Population/Doctor Ratio	10,692	7,565	410 (Israel)	340 (Italy)	97,120 (Burkina Faso)	48,510 (Burkina Faso)

See appendix 1 for data sources.

TABLE 3.15

EFFECT OF SOCIOECONOMIC INDICATORS ON LIFE EXPECTANCY AT BIRTH AND INFANT MORTALITY RATE, 1970 AND 1980.

VARIABLE	LIFE EXPECTANCY (EXP)				INFANT MORTALITY RATE (IMR)			
	1970		1980		1970	1980		
	(i)	(ii)	(iii)	(iv)	(i)	(ii)	(iii)	(iv)
RGDP	1.39 (6.53)	0.68 (3.130)	1.12 (8.20)	0.45 (3.03)	-5.98 (-5.50)	-2.14 (1.70)	-4.87 (-7.29)	-1.41 (-1.60)
$RGDP^2$	-0.017 (-4.25)	-0.0091 (-2.48)	-0.011 (-5.36)	-0.0043 (-2.20)	0.071 (3.52)	0.026 (1.25)	0.05 (4.81)	0.012 (1.08)
Z^2	0.015 (3.19)	0.0089 (2.16)	0.010 (3.91)	0.0039 (1.72)	-0.063 (-2.65)	-0.025 (-1.06)	-0.046 (-3.54)	-0.010 (-0.78)
POPDOC	-	0.0035 (0.75)	-	-0.0099 (-1.93)	-	-0.015 (-0.60)	-	0.032 (1.07)
TILLIT	-	-0.26 (-10.40)	-	-0.19 (-7.76)	-	1.29 (9.00)	-	1.02 (6.92)
Constant	35.50 (15.82)	57.45 (17.91)	39.61 (23.01)	59.80 (21.40)	183.64 (16.09)	75.00 (4.07)	161.04 (18.94)	58.68 (3.49)
R^2Adj.	0.72	0.89	0.79	0.93	0.67	0.85	0.74	0.87

See appendix 1 for data sources and variable definitions.

TABLE 3.16

EFFECT OF CHANGES IN RGDP ON CHANGES IN LIFE EXPECTANCY FOR SELECTED PERIODS

Dependent variable - Change in life expectancy

VARIABLE	1970 - 1980	1980 - 1990	1970 - 1990	PRESTON'S RESULTS 1930 - 1960
RGDP	-0.0038 (-0.494)	-0.0161 (-2.077)	-0.023 (-1.340)	-0.01883 (-1.902)
Absolute change in RGDP (X2)	-0.077 (-3.017)	0.022 (0.692)	-0.019 (-0.402)	0.00128 (0.218)
Ratio of RGDP (X3)	2.920 (3.771)	-0.827 (-0.548)	1.860 (1.929)	0.38756 (0.229)

t - statistics in parantheses

TABLE 3.17

EFFECT OF CHANGES IN RGDP ON CHANGES IN LIFE EXPECTANCY

VARIABLE	1970 - 1980	1980 - 1990	1970 - 1990
RGDP	-0.021 (-4.563)	-0.013 (-2.110)	-0.029 (-4.304)
Ratio of RGDP (X3)	1.080 (2.178)	-0.0059 (-0.006)	1.552 (2.705)

t - statistics in parentheses.

Figure 3.1

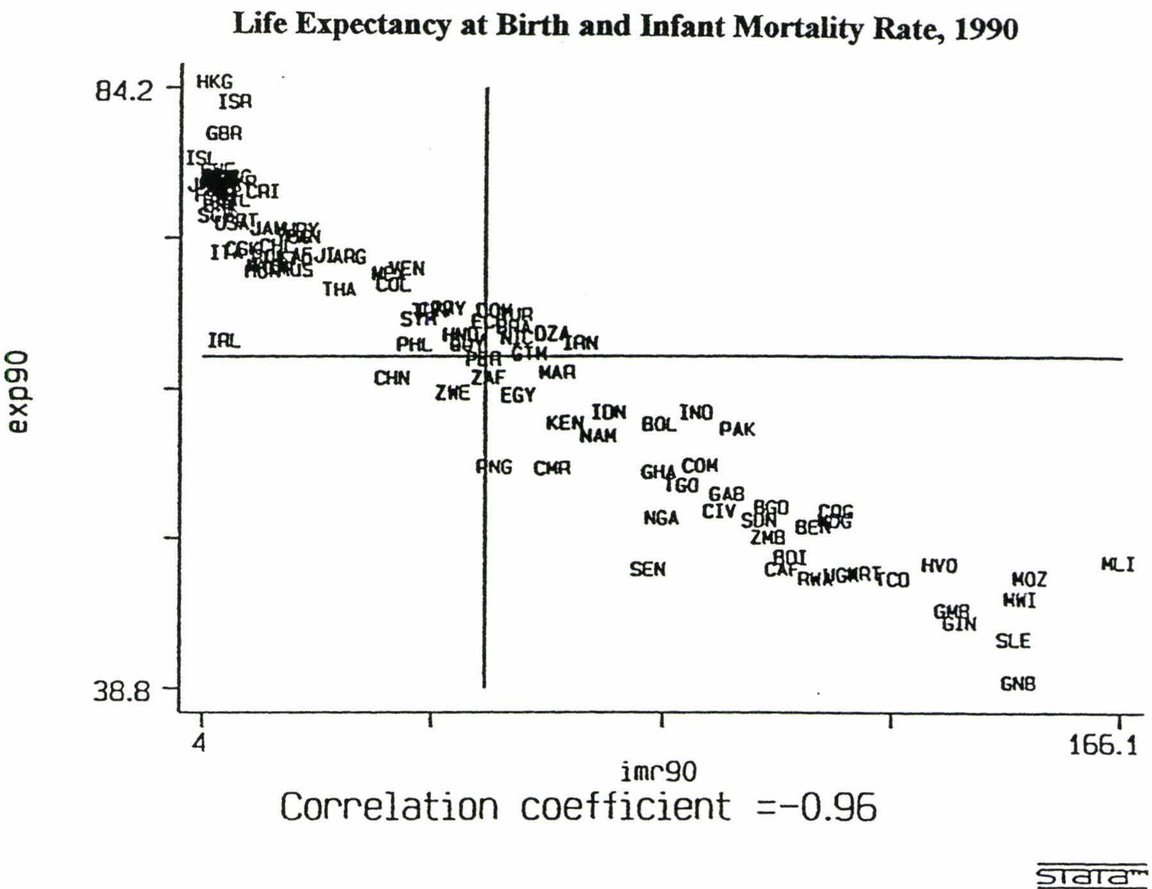

Life Expectancy at Birth and Infant Mortality Rate, 1990
84.2
38.8
exp90
4
166.1
imr90
Correlation coefficient =-0.96
HKG
ISR
GBR
CRI
THA
IRL
PHL
CHN
ZWE
EGY
MAR
IRN
KEN
IDN
NAM
BOL
IND
PAK
PNG
CMR
COM
TGO
GAB
NGA
CIV
ZMB
SEN
HVO
MOZ
MWI
SLE
GNB
MLI

Figure 3.2 **Burden of Disease Attributable to Premature Mortality and Disability by Demographic Region,1990.**

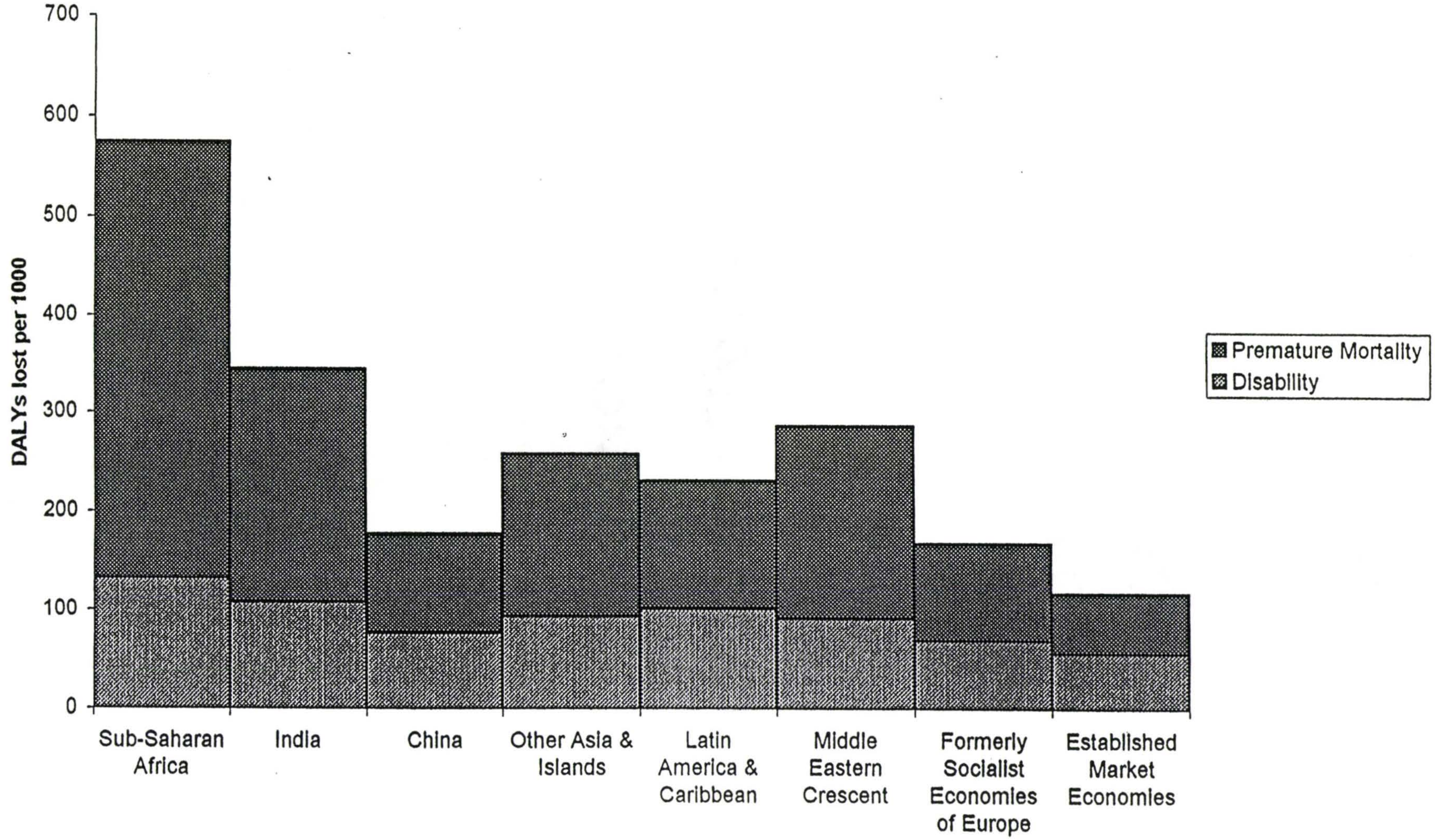


Figure 3.3

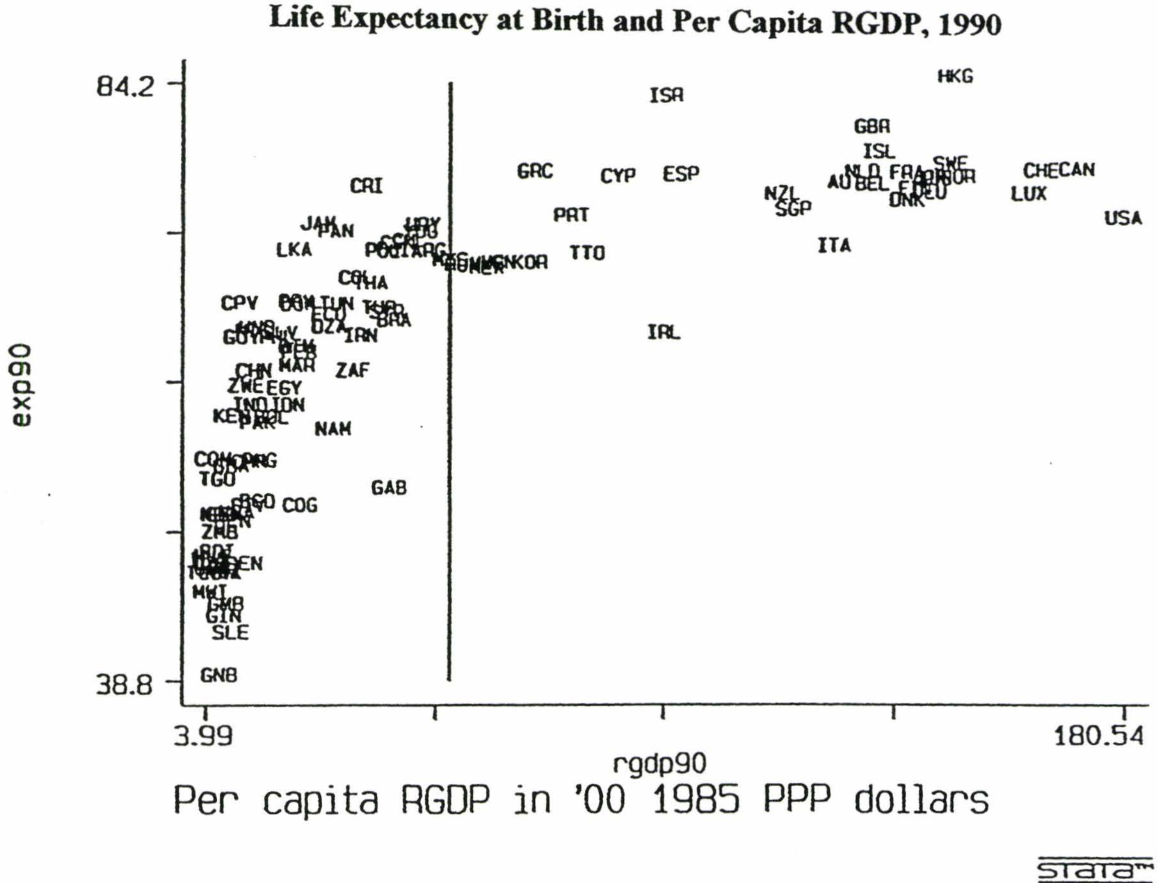

Life Expectancy at Birth and Per Capita RGDP, 1990
84.2
38.8
exp90
3.99
180.54
rgdp90
Per capita RGDP in '00 1985 PPP dollars
USA
LUX
CHECAN
HKG
ISR
GBR
ISL
ITA
IRL
GRC
CYP
ESP
PRT
TTO
CRI
LKA
GAB
COG
NAM
ZAF
SLE
GNB

Figure 3.4

Predicted Life Expectancy at Birth to Per Capita RGDP, 1990

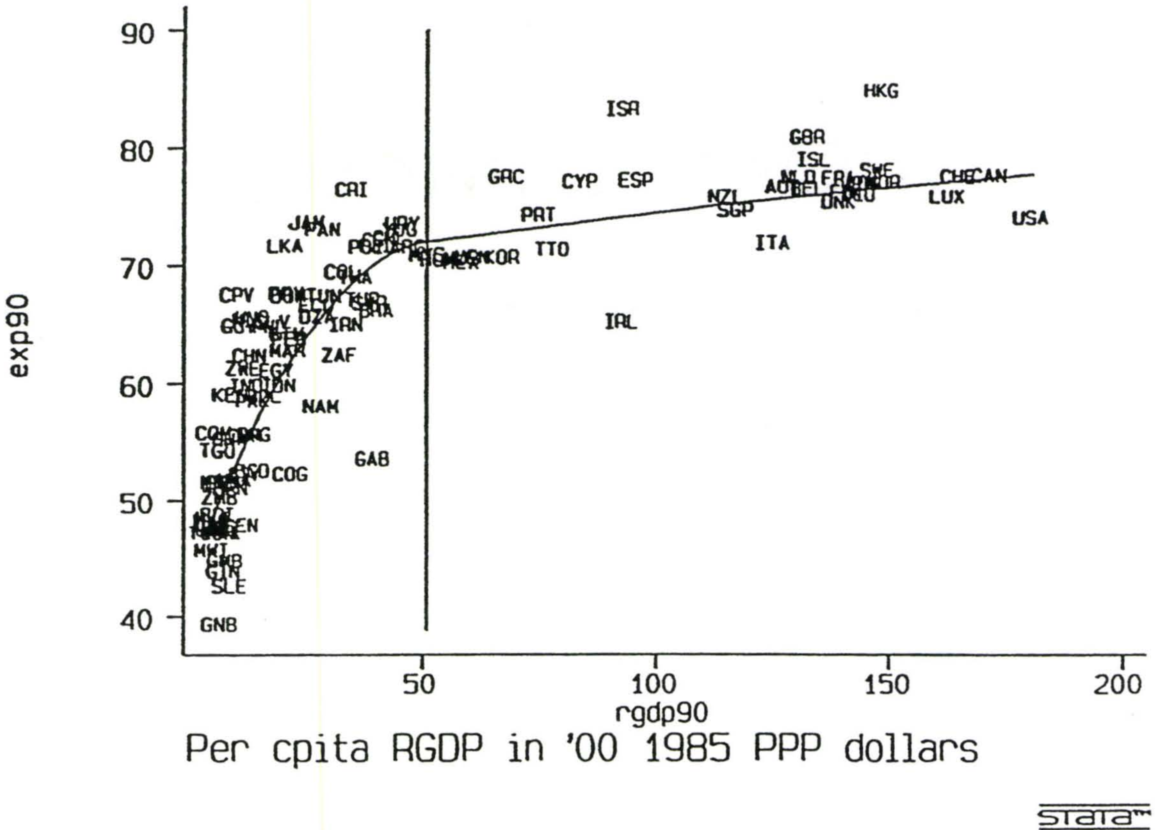


Figure 3.5

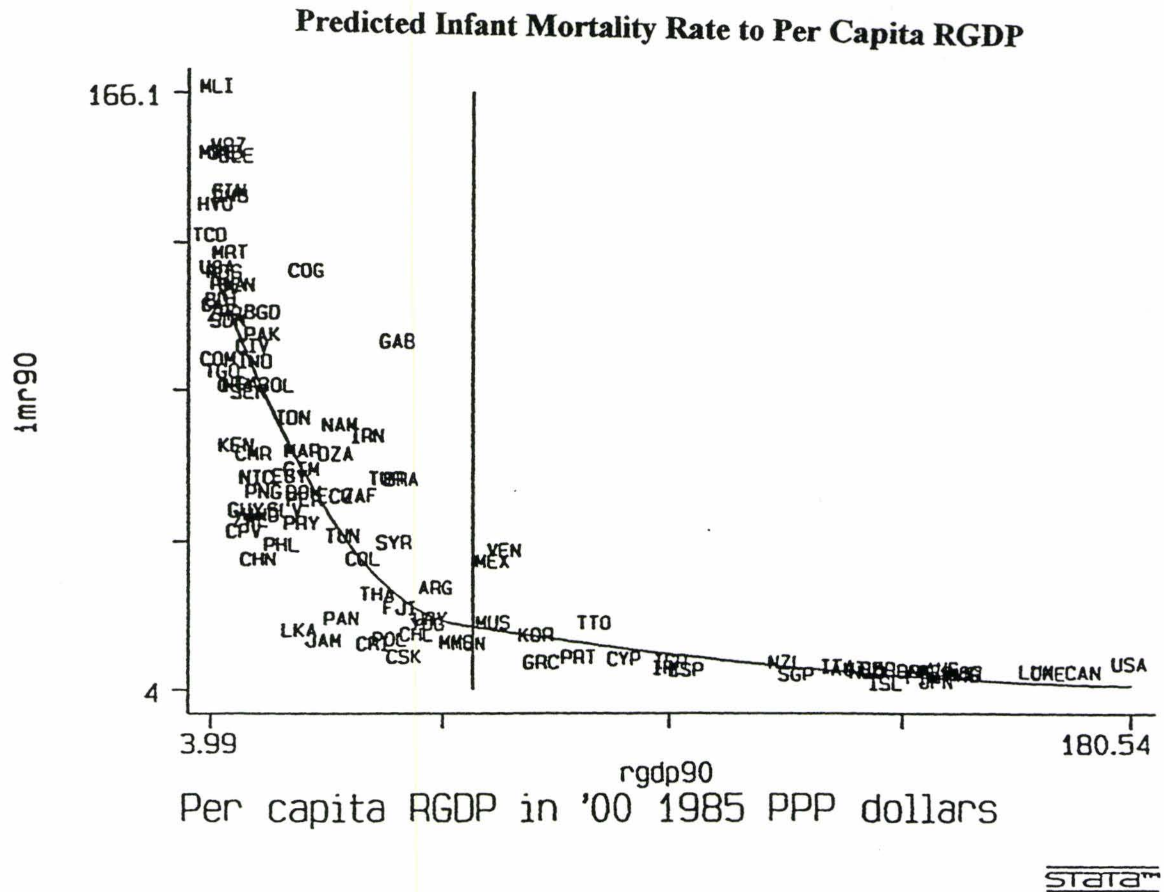

Predicted Infant Mortality Rate to Per Capita RGDP
166.1
4
imr90
3.99
180.54
rgdp90
Per capita RGDP in '00 1985 PPP dollars
MLI
TCD
MRT
COG
BGD
PAK
GAB
IND
BOL
NAM
IRN
KEN
CMR
DZA
PNG
PRY
SYR
PHL
CHN
COL
ARG
PAN
LKA
JAM
CSK
VEN
MEX
MUS
KOR
TTO
GRC
PRT
CYP
SGP
ISL
JPN
CAN
USA

Figure 3.6

Predicted Life Expectancy at Birth and Per Capita RGDP, 1970

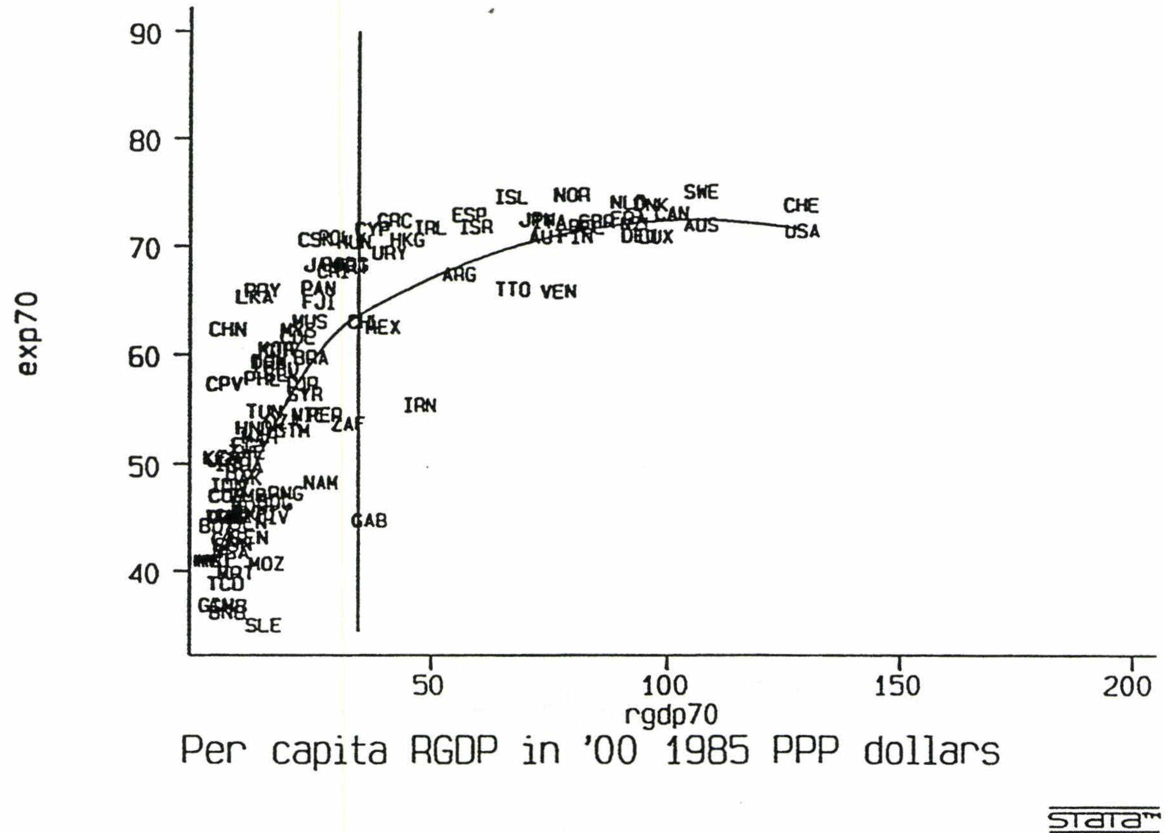


Figure 3.7

Predicted Life Expectancy at Birth and Per Capita RGDP, 1980

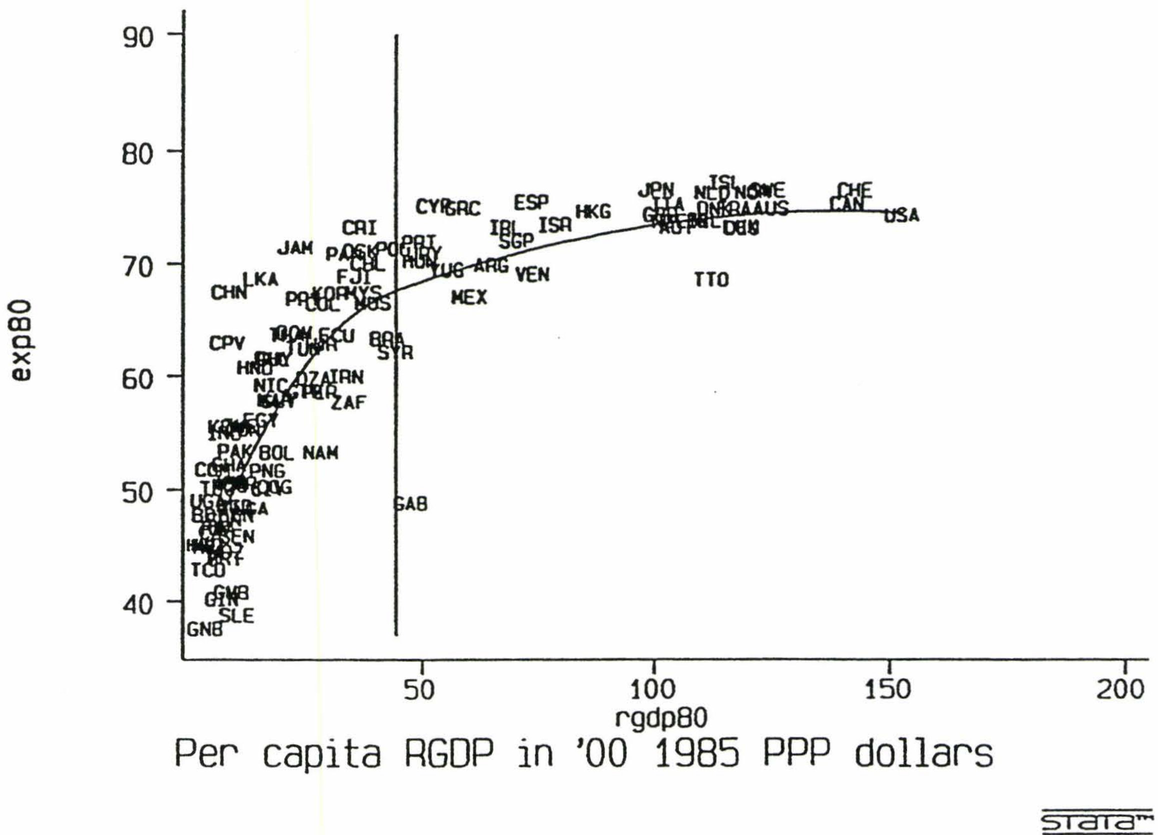
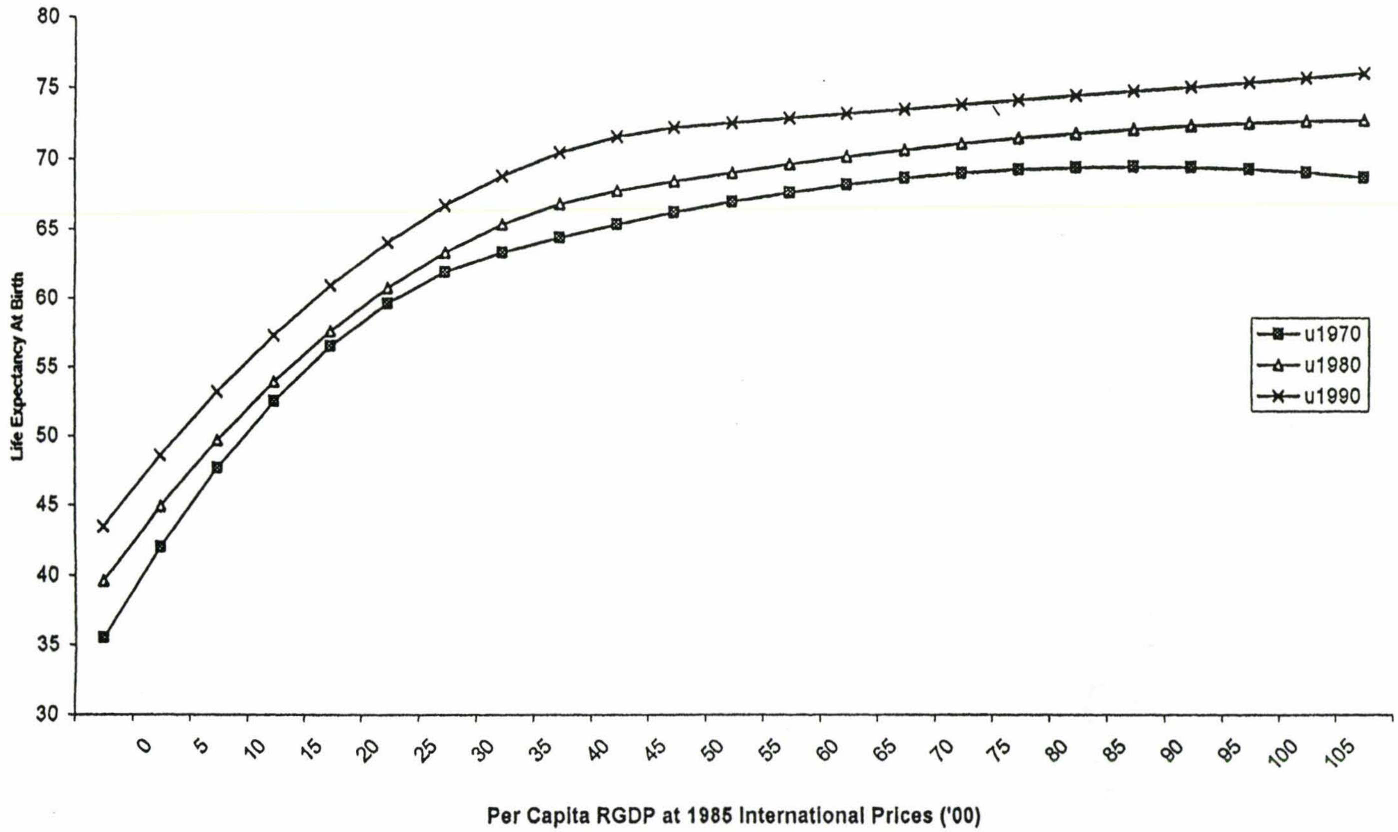


Figure 3.8 Structural Shift, 1970-1990



CHAPTER 4.

MICRO HOUSEHOLD ANALYSIS

4.1. Aims and Objectives:

In this Chapter we estimate equations informed by reduced form solutions of the household model for child health and the effects of various household socio-economic factors. Over the past decades, most developing countries have experienced drastic declines in their general mortality levels. Despite this, infant mortality and child mortality remain high within the rural areas and among the urban poor. Our major interest in this study is to shed some light on the determinants of child health at the household level. We use the case of Kenya to investigate the interplay of various child and household characteristics on child survival.

4.2 Data Sources:

We utilise data from the 1989 and 1993 Kenya Demographic and Health Surveys (KDHS89 and KDHS93). Field surveys were conducted by the National Council for Population and Development, Ministry of Home Affairs and National Heritage and the Central Bureau of Statistics in the Ministry of Planning and National Development, Kenya. Financial assistance for the surveys was provided by the United States Agency for International Development (USAID) with technical assistance from Macro International Inc., USA., who also coded and managed all demographic health survey data sets. Appendix 5 lists variables covered in the survey.

As a standard practice in these kind of surveys, three questionnaires were administered: the household schedule, the woman's questionnaire and the husband's questionnaire. The household schedule is used to identify household members and select those eligible for the other two interviews. In the husband's questionnaire a sample of males (1,116 in 1989 and 2,336 in 1993) were asked questions on their background and demographic characteristics as well as their knowledge and use of contraceptives. For the 1993 survey they were further questioned on their sexual activity and AIDS awareness. The woman's questionnaire is quite detailed. Most of our analyses will be based on the responses from it.

The individual woman's questionnaire, which is of prime interest in this study, contains the following sections:

- respondent's background
- reproductive history
- contraception knowledge and use
- health and breast-feeding practices
- marriage history
- fertility preferences
- male's background information and woman's work history.

A total of 7,150 women between the ages of 15 and 49 were interviewed for the KDHS89 while for KDHS93 the figure was 7,540. North Eastern and parts of the Eastern provinces were not covered due to security reasons and inaccessibility as a result of lack of roads and difficult terrain. The same area was not covered in the two surveys. The area covered represents 95% of the Kenyan population. The map, Figure 4.1 shows the coverage of the surveys.

The dependent variable will be based on child mortality. Neither KDHS data set contains any household income data but KDHS89 has substantial information on household asset ownership. Appendix 6 shows the distribution of asset ownership between urban and rural respondents. We use asset ownership as proxies for income. Ownership of a car, refrigerator, tractor, permanent house or even a T.V., is highly correlated with income, especially in LDCs.

4.3. Related Findings and Variable Definitions:

Behrman and Wolfe (1987) estimated the probability of child mortality for various regions in Nicaragua. Child mortality was measured as a dichotomous variable. We estimate the probability of child mortality based on whether the child died or lived to his/her first birthday and then estimate child mortality to age 3. Despite drastic declines in mortality across all countries, infant and child mortality are still high in developing countries. According to World Bank (1999), the infant mortality rate in Kenya is 57 per 1000 live births compared to the average for low income countries of 68; middle income countries, 35, and 6 for high income countries.

The micro household model discussed in Chapter 3 identifies major determinants of health. Much attention has been directed to the broader category of socio-economic determinants. In more developed countries, for example, mortality has declined to levels whereby further significant declines may only be achieved through changes in lifestyles.

In modelling determinants of health at the micro level, the household model assumes that parents maximise their utility subject to certain constraints. These actions determine optimal levels of child mortality, health and nutritional status. In the case of imperfect information, parents may only perceive a few of the variables though they will still determine their children's health status.

Wolfe and Behrman (1982) observed that it is difficult to specify the exact functional form of the utility and production functions that could be used to get the reduced forms for empirical estimation. The choice of functional form must be arbitrary. There are no theoretical reasons for preferring say a linear form to Cobb Douglas or Leontief preferences within this context. Indeed, even if we knew the form of the utility function, it may still be impractical to derive the reduced forms because we may not know the form of the production functions.

We proceed to estimate "stylized" reduced forms of the household function using the logit model. The data set contains longitudinal event history data. It is also censored.

The death of a child is normally a very painful, sad and unfortunate event. The affected families go through a grieving period beyond which they then would not like to discuss their lost child, especially with people from outside their community. Death is therefore rarely looked at as a natural occurrence that should be talked about. In Kenya, as is the case for most African cultures, mothers were traditionally socialized not to talk about their dead children, especially if they died in early infancy. There is therefore likely to be systematic underreporting of deaths, especially those that occurred in early infancy. There is however an endogeneity problem when one attempts to control for this by incorporating the age at death of the child in estimation. We only have this information if a child is dead at the time of the interview. For the censored cases we do not know when they died.

As mentioned above, we estimate the probability of child mortality, with child mortality as a categorical variable taking the value of 1 if the child died within the reference period, 0 otherwise. We use the logit model. The suitability of this model to this type of data set is discussed in the next section. Like Wolfe and Behrman (1982), we discuss the explanatory variables based on four broad categories, with the variable names given in parentheses.

(i) <u>Factors that affect efficiency in household production</u>

Mother's age (momage,momage^2)

The variable **momage** refers to the current age of the mother at the time of the survey. The survey does not report age of the mother at the birth of the child. Although it may be possible to reconstruct age at birth of child from the child's current age **(cage)** there are some problems with doing this. First of all, **cage** is only available for those children still alive at the time of the survey. For those dead at the time of the survey it might be possible to subtract **momage** from age of child at death **(agedead)** plus the time that has elapsed since the death of the child. However, reported age of child at time of death is generally very unreliable. This is due to memory lapses and cultural practices in which mothers do not generally remember when the child died, especially if this happened in early infancy.

In general, mothers may become more efficient with general experience and maturity (Wolfe and Behrman,1982). The effect of a mother's age on child health may, however, be nonlinear. Teenage mothers and those close to menopause tend to have less healthy children than those in the middle of the child bearing ages. Teenage mothers tend to have unplanned pregnancies and are thus less ready to take charge of motherhood. Mothers in the middle of the child bearing ages may have better education as they have had time to complete it, and are more prepared to provide care for their children. Mothers close to menopause run a higher risk of bearing children with health complications. For example, the risk of having a child with chromosome defects such as Down syndrome increases with the age of the mother, especially after age 35. It is for this possible nonlinearity that we incorporate the polynomial term, **momage^2.**

Mother's Education (momedu)

Education is not only a consumption good (we gain positive utility from it) but also an investment in human capital. Education, therefore, affects one's capability to implement health and nutritional decisions within the household. In the US, for example, Edwards and Grossman (1978) found that parental education was a more crucial determinant of child health than was family income. Barrera (1990) summarizes how mother's education affects child health. First, mother's education increases the economic resources available to the family through the mother's association with wealthier men, increased earnings from market efficiency gains and through an increase in full income brought about by non-market efficiency gains.

Second, Barrera argues that education improves the allocation of resources due to better knowledge and access to information. Better information allows the more educated mothers to make better informed decisions that contribute positively to the health of their children. Finally, education may be indicative of a greater willingness of the parents (especially the mother) to make human capital investments and to provide better nutrition and better training in home making which affect her ability to raise "high quality" children (Barrera (1990); p71).

The Kenyan system of education changed in 1985 from a four tier system (7-4-2-3) to a three tier (8-4-4) system. i.e. from 7 years of primary, 4 in secondary school, 2 of higher secondary school and 3 of university (except for Medicine and Architectural programs, which were 5 years) to 8 of primary, 4 of secondary and 4 of university. In our analysis, we combine secondary and higher education for two reasons. One is that, because of the change in the system, it is difficult to determine whether the respondent should be placed in the secondary or higher category. The second reason is to maintain a reasonable number of elements in each category. However, for the purpose of tabulation, the Data are presented in the following categories: 0- None, 1-Primary education, 2-Secondary, 3-Higher, 9-missing. In estimation, we use three indicator dummy variables, with one of them as the reference group. The reference group is no education. The other two categories are defined as follows: **Imomedu_1-**Primary education and **Imomedu_2-**

Secondary and Higher. Throughout, an "**I**" at the beginning of a variable name denotes a dummy variable generated in the logit estimation by the *xi*: interaction procedure in STATA. The other dummy variables not generated by this procedure are preceded by letter **d**. (See Appendix 7 on variable definitions.) Each dummy variable equals 1 if the value is true, 0 otherwise. Note that it has been established in the empirical literature that there is a wide difference in child survival rates between parents with no education and those with basic education. There is, however, not much difference in survival of children between mothers who have completed secondary education and those with university education, controlling for other factors such as income and marital status. This provides another reason for grouping secondary and post-secondary groups together.

Marital Status (dmarried)

We expect children of married couples to be less at risk of death than those from single parents (never married, separated, divorced and widowed). The husband's income may supplement the mother's income. They also are expected to offer assistance in household production and in decision making. The presence of a husband also frees up the mother's time and she is able to concentrate more on child health enhancing activities. The categories for marital status in the survey are: never married, married and living together, living together but not married, widowed, divorced, and not now living together but still married. We collapse these categories into two broad ones because our focus is to find out if there are differences in child mortality among children from single parent and those from two parent families. The reference group will be a category that comprises single parent families and includes: never married, widowed, divorced, not living together. Therefore the dummy variable, **dmarried,** groups together the married and living together and the now living together but not married categories.

Number of Marital Unions (munions)

Women respondents were asked the question: "How many other wives does he (in reference to their husband) have?" The number of other wives ranged from zero to ten. Note that there are no reported cases of polyandrous relationships (one wife-many husbands) in Kenya. Polygynous unions are, however, widespread and in some cases having more than one wife is considered a status symbol.

In order not to lose the respondents that were not married we recoded the value labels. In the original data set **munions** is coded with a missing value label if the respondent was single (never married, widowed or divorced). We recoded these observations with a value label of 0. Note that the fact that this is the same value as for the monogamous unions does not cause a problem in the estimation due to the inclusion of the two parent family dummy, which picks up the effect of being married (**dmarried=1**) and monogamous (**munions=0**) versus not married (**dmarried=0** and **munions=0**). We further include the variable age at first marriage (**agemarry).** For those respondents that were never married we change their value labels from missing to 0.

Controlling for income, however, **munions** may have a negative effect as it enlarges family size and consequently increases competition for available family resources. The husband also spends less time at each household and therefore his potential contribution to household production per household is reduced. This effect may be nonlinear. The effect of having one other wife versus zero may be larger than having 5 versus 4. We therefore will incorporate both **munions** and **munions^2.**

Husband's Education (hedu)

As in **momedu,** the reference group is "no education". The other categories are defined as follows: **Ihedu_1** equals 1 if primary education, 0 otherwise, and **Ihedu_2** equals 1 if secondary education or higher, 0 otherwise. As noted for mother's education, husband's education leads to efficiency in household production. In addition, better education may be associated with higher earnings. Although husband's occupation was covered in the survey, we do not use it as an explanatory variable due to the fact that self employed individuals were reported as unemployed.

Number of young children in household (under5)

Wolfe and Behrman (1982) argue that the number of young children in the household has a greater bearing on child mortality than household size **(hsize)**. They assert that the negative effect of household size is offset by the positive effect of economies of scale. This may hold true especially for rural areas where family labor is crucial for ensuring a steady and sufficient food supply. We use both **under5** and **hsize** in our estimations. Note that the number of children less than five years does not include the current child if the child died. Therefore in order to avoid potential endogeneity bias we redefined **under5** as number of other children less than 5, i.e. excluding the child in question. We also redefined **hsize** as the number of people in the household that are aged 5 years or more.

Residence (resid)

We stratified the sample into rural and urban; **resid** equals 1 if rural and 0 otherwise. The base case is therefore urban residence. In Kenya, as is the case for most LDCs, there is a concentration of health care facilities and providers in the urban centers. At the same time, with the rapid rate of urbanization there has been an emergence of slums in the fast growing urban areas. Overall we might expect better average health status in the urban than in rural areas due to proximity to health care providers and facilities.

Asset ownership (owncar)

Because we do not have a measure for family income, we use car ownership **(Iowncar)** as a proxy. The public transport system in Kenya in both the urban and rural areas is very underdeveloped and unreliable. Therefore, above a certain income, a car becomes essential especially in an urban setting where individuals have to commute to work and school. Note that in surveys in more industrialized countries, the number of cars owned by the family is sometimes used to proxy family income. The initial purchase price or down payment is quite high and prohibitive for low income groups. Loans are not readily available without collateral. Therefore only those in the upper middle or high income bracket are able to purchase cars. Car ownership is modeled as a dummy variable. **Iowncar** equals 1 if the household owns a car, 0 otherwise.

(ii) Nutritional factors

Length of breast feeding (breast)

Longer breast feeding may help the infant develop immunity against some common ailments. The mother-child bond developed through the breast feeding process may also help the child in both emotional and psychological development. However, breast-feeding may be related to availability of mother's milk substitutes, socio-economic class and perceived opportunity cost of mother's time. These factors may be highly influenced by mother's age and educational level. In the two surveys, the maternity history, of which breast-feeding is a part, covers up to six entries, relating to births in the five years preceding the survey. In cases where there are more than six births within the last five years (for example, where there were multiple births) only the last six are included.

There are two issues that we had to deal with in regard to this variable. One was the censoring issue for those children still breast-feeding. We do not know how long after the survey they continued breast-feeding. The other issue was redefining the variable in order to avoid a possible spurious correlation between breast-feeding duration and mortality by not using the part of breast-feeding duration that overlaps the age range over which we are estimating the logit. The length of breast-feeding will be smaller, on average, for those children who died than for the survivors, since breast-feeding will stop when the baby dies. Thus this might cause a spurious negative correlation between time breast-fed and mortality.

To deal with these issues for infant mortality we introduce an ever/never breast-fed dummy variable **(dbreast).** The dummy equals 1 if child ever breast-fed, 0 otherwise. Entries that were recorded as inconsistent (code 97), say for example where the reported length of breast-feeding is longer than the child's age, will be incorporated in the ever breast-fed category, as will be those whose reported length of breast-feeding as 0 (less than one complete month). Only those who "never breast-fed" (code 94) will have **dbreast=0.**

When estimating mortality between ages 1 to 3 years, the breast-feeding dummy is defined in a slightly different way. It equals 1 if the child breast-fed beyond 12 months, 0 otherwise.

(iii) Public environment factors

We incorporate the availability of toilet facilities to capture this. Respondents were asked what kind of toilet facility the household had. The information was recorded in five categories: flush toilet, bucket, pit latrine, other and none. Of the total respondents for KDHS89, 73% used pit latrines, 13% flush toilets, 12% had no toilet facilities and 2% used "bucket" or other toilet facilities. We introduce **toilet** dummy variables to capture each of them. The reference group is no toilet. Note that this variable may also be a proxy for income. We introduce two dummies: **Itoilet_1-**flush toilet**, Itoilet_2-**other facilities (pit latrine, bucket and other).

(iv) <u>Other variables and controls</u>

Child's sex (csex)

This will be a dummy variable which equals 1 if child is male, 0 otherwise. Genetic differences between sexes lead to better survival rates for female children than male children. However, son preference, embedded in traditional norms, cultural characteristics and practices, may offset these biological differences. Preference for male children may be based on higher expected earnings and male status within the family. Naming practices, patrilineal and male dominance, inheritance and dowry practices, for example, may also lead to a preference for male children in certain LDCs. A preference for male children may also result from the expected provision of old age security (financial or otherwise) by the child to parents especially where publicly financed old age security and/or financial markets are not well developed. Allocation of resources within the household may therefore favor male children, affecting the survival of female vis a vis male children.

Birth Order (bord and bord^2)

Mothers tend to become more efficient in household production with experience. We might therefore expect higher birth order infants to have higher chances of survival. Larger family sizes, however, (ceteri paribus) imply competition for available resources and this may therefore offset efficiency gains from experience. We thus incorporate a quadratic term (**bord^2**) to capture this.

Religion (Ireligion)

We hypothesise that religious beliefs tend to influence child care practices at the household level. We are particularly interested in whether there are any differences in child survival rates between children born to Christian versus Muslim mothers. Caldwell (1993) found higher infant mortality among Muslim mothers. Controlling for higher female literacy among Christian mothers compared to Muslim mothers, there could be additional effects of Muslim practices on child survival. We therefore include this variable to control for religion influence on child survival. The reference category is Catholics and Protestants. We define the other dummy variables as follows: **Ireligion_2-**muslims **and Ireligion_3** -other**.**

4.4. Descriptive Analysis:

Table 4.1 shows the distribution of respondents (mothers) by age at the time of interview for the two surveys. The mean current age for both surveys is 28 years. The mean age at marriage is, however, 18 years. Appendix 8 shows means and standard deviations for most variables discussed above. Table 4.2 shows the distribution of respondents by region. As noted earlier, the North Eastern and some parts of the Eastern provinces were not included in the survey. Tables 4.3 and Table 4.4 show respondent distribution by region and highest education achieved. Overall, women's levels of literacy increased during that period. The percentage of those with no education declined from about 24% in 1989 to 17% in 1993 while percentages of all the other education categories increased. From Table 4.3, women in Nairobi were the most highly educated, with over 40% reporting having secondary or higher education, followed by Central (25%), Western (22%), Rift Valley (20%), Nyanza (17%), Eastern (16%) and Coast, which is predominantly Muslim, only 14%. Over 50% of women in the Coast region reported having no education at all. Table 4.4, which summarises results for 1993, show that despite the fact that the percentage of sample for Nairobi which had a higher proportion of women in higher education declined, the education level for the overall sample increased. This could be attributed to a general increase in educational levels for all regions. In Coast province for example, the percentage of women with no education declined while there was an increase in percentages in other categories.

4.5. Logit Model:

4.5.1. Model Overview

In this section we estimate the probability of child survival conditional on certain household, parental and socio-economic factors. Logit analysis is essentially a technique that is used to predict a binary dependent variable from a set of explanatory variables. In logit regressions our main objective is to estimate the probability of an event occurring, conditional on the value(s) of some explanatory variables.

For a single variable *(X)*, the probability of event *Y* occurring conditional on *X* is denoted by :

$$P(Y/X)=exp[b_o+b_jx]/(1+exp[b_o+b_jx]) \qquad (4.1)$$

where b_0 and b_j are the coefficients to be estimated from the data while *X* is the explanatory variable and exp[ ] denotes the antilog function. The coefficients are estimated by the maximum likelihood method. In the case of multiple regressors the probability of event *Y* occurring could be generally written as:

$$P(Y|X)=e^{X'\beta}/(1+e^{X'\beta}) \qquad (4.2)$$

X is a vector of explanatory variables (e.g. child, environmental and household characteristics) and β is a vector of unknown parameters to be estimated. In this study we are interested in analysing child survival in Kenya using KDHS89 and KDHS93. The estimation will be done as follows:

First, using the KDHS89 data set, we estimate the probability of a child surviving to 12 months and to 3 years using the logit model. We focus on the last child (current birth) in order to minimize potential memory lapse problems. Mothers are more likely to recall more accurately events that occurred in the more recent past. Thus such information as breast-feeding and age of child at death is likely to be more accurate for more recent births. Note that maternity history is only available for the births that occurred in the 5 years preceding the survey. Therefore the choice of 3 years as the upper end was in order to maintain a balance between having a large enough interval to be able to study mortality and maintaining a reasonable sample size. As we shall show later when estimating mortality between time interval n and $n + a$, we exclude deaths that occurred before age n and also those children still alive but not yet $n + a$ old. The higher the interval a, the higher the number of children that are excluded. The value $a = 2$ years was deemed most appropriate for $n = 1$.

Second, using the KDHS93 data set, we estimate mortality over the same age interval and compare the results to the 1989 results. Below we will develop the logit model and discuss some of its advantages. In this model there are two possible states of health at time $n+a$. Either the child survived to time $n+a$ (H=0) given she was alive at time n, or she died (H=1). The dependent variable will therefore be dichotomous, taking the value of 0 or 1 depending on the child's health state at the reference time period $n+a$.

As noted above, we especially consider the case for n=0 (birth) and a=12 months to capture deaths between 0-12 months and n=1 year and a=2 years to estimate child mortality between ages 1-3 years. For each observational unit there are two states of health. Either the infant/child dies (H=1) or the child survives (H=0).

The logit model specifies the probability of the child dying within any time period ($n+a$) conditional on having survived period n. It is given by (4.2) with "event Y" being death

(H = 1), where to save space the conditioning on X is assumed. This probability is also referred to as a discrete hazard and measures the "probability that an event will occur at a particular time to a particular individual, given that the individual is at risk at that time" (Allison,1984, p.17).

From (4.2),

$$P(H=0)=1-P(H=1)=1-[e^{X'\beta}/(1+e^{X'\beta})]=1/(1+e^{X'\beta}) \quad (4.3)$$

The odds ratio is: $P(H=1)/P(H=0)= e^{X'\beta}$. Expressing it in terms of the log odds ratio gives log $[P(H=1)/P(H=0)]=X'\beta$. As P(H=1) varies between 0 and 1 the odds ratio will vary between 0 and infinity while the log odds ratio varies between minus and plus infinity

The log likelihood function is as follows:

$$\ln L = \sum [y_i \ln P(H_i=1) + (1-y_i) \ln P(H_i=0)] \qquad (4.4)$$

Where $y_i=1$ if $H_i=1$

$y_i=0$ if $H_i=0$,

and $H_i=1$ if child i died, $H_i=0$ otherwise, Greene (1993) and Kennedy (1992).

We redefine our sample in an attempt to overcome a problem encountered with the logit/probit model in estimation using longitudinal event history data (Allison, 1984). This is the issue of censoring. Because of the way we model the dependent variable, some censored observations are now included. For example, when we consider deaths between 0-12 months, although we omit those who are alive and not yet 12 months, we include everyone else, including those who died after the first 12 months. They enter the estimation as having survived the 0-12 month interval.

The econometric analysis will therefore be based on estimating the logit model by the maximum likelihood logit procedure.

4.5.2. Estimation of Logit Model

In this section we estimate a logit model on the last child born to each of our respondents. We first look at the probability of a child surviving to 12 months followed by survival to 3 years. As noted in the previous section, we are interested in modelling whether the last child was dead ($H = 1$) or alive ($H = 0$) at time $n + a$ given that the child survived to time n. In most of our Tables we report both odds ratios and coefficients. An odds ratio of less than 1 implies that an increase in this variable is likely to reduce the probability of the event occurring. In this case an odds ratio of less than 1 means the variable has a negative effect on the probability of the child dying. The coefficients could be re-calculated by taking the natural logs of the odds ratios. STATA statistical software was used.

As noted above an odds ratio for each variable represents how a change in the variable affects the probability of dying ($H = 1$). An odds ratio closer to 1 would therefore imply the effect of this variable is smaller with a coefficient closer to 0. We obtain the Huber or White Robust Standard Errors by using the robust option with the logit command.

4.5.2.1. Probability of Dying, 0-12 months

The logit results for deaths between 0-12 months are given in Table 4.5. We test jointly the restrictions that the following variables: place of residence, husband's education, religion, ownership of a car, household size and the nonlinear marital union term, which are individually insignificant, are jointly insignificant at 5% level of significance. The chi(2) test statistic is 3.50 as compared to a critical chi(2) with 8 degrees of freedom of 15.51. We therefore accept the null hypothesis of joint insignificance, with a caveat that the test is biased toward accepting due to our having selected the hypothesis on the basis of their individual insignificance.

In both cases (restricted and unrestricted equations), age at first marriage **(agemarry)** is positively associated with infant mortality. Note that in the summary statistics we showed that the mean age at marriage was about 18 years. Since we control for mother's age, this implies that a higher value of agemarry would show that the child was born after the mother had been married for a shorter period of time. This variable therefore may be capturing the effect of length of time married before birth of the child. Those mothers who have been married for a longer time before having children may have had a longer period to prepare for children. Those who have been married for a shorter period may possibly not have had enough time to plan for the birth of the child.

The number of marital unions **(munions)** has a significant negative impact on survival of children. An odds ratio of 1.49 for **munions** along with an insignificant odds ratio of 1.00 for **munions2** suggest that, ceteris paribus, having an additional wife on average increases the probability that the child dies by about 49 percent (e.g. from 0.4 to 0.6). The corresponding log odds ratio is 0.40 which indicates that each additional wife increases the log of the hazard (probability that the child dies) by about 40 percent.

Breast-feeding is highly negatively associated with infant mortality. As noted earlier, this might be related to the resulting immunity to diseases that is associated with breast milk. It may also be possible that women who do not breast-feed do so because of their low health status. Mothers with lower health status are also likely to have less healthy children. Availability of toilet facilities is also inversely related to infant mortality. The number of other children in the household that are under 5 also surprisingly is associated with a higher child survival. This may however be the case if it indicates the underlying characteristics of mothers. It might be that fecundity is positively associated to mother's health. Therefore, in LDCs with low contraceptive use, healthier and more fecund women may be more likely to have (more) healthy children.

Note that for mother's current age and birth order, the coefficients on the non-linear terms **(momage2)** and **(bord2)** are opposite in sign to those of the linear terms **(momage)** and **(bord)**. In order to investigate the effect of mother's age and birth order, we plotted (not shown) the first partial derivatives with respect to each variable over the relevant variable range. In the case of mother's age, infant mortality increases with current age of the mother at all ages. On the other hand it declines with birth order. Lower infant mortality associated with birth order might be related to the mother's gain in experience in child caring. Having had a child before makes the mother more efficient in household production. Infant mortality may also decrease with mother's age due to gains in child caring experiences, but this decline may be offset by the financial burden of having many children. Controlling for mother's experience, we then observe that children born by older mothers stand a higher risk of dying in infancy. There are health risks associated with having children at an older age. As discussed in the literature review, although children of teenage mothers seem to have more health problems, those born of mothers above age 35 have even more medical complications.

Mother's education (**momedu)** shows the expected relationship, i.e. infants whose mothers have had primary or higher education have better chances of surviving infancy. There was, however, no significant difference in infant mortality between mothers with primary education and those with higher education. The results once again point to the importance of literacy in improving children's health status, as already established in the aggregate analysis in the previous chapter. At the aggregate level we established that literacy level was positively associated with life expectancy at birth and negatively with infant mortality rates.

The married dummy (**dmarried**) is negative and significant. Thus, infants born in two parent families have better chances of surviving infancy than those of single parents.

Though insignificant at both 5% and 10% levels of significance, the coefficient on car ownership in the unrestricted equation is negatively associated with infant mortality. As mentioned in the literature review this might be capturing the positive effect of income (or wealth) on child survival. Owning a car is known to be highly correlated with income. This is in fact more so in less developed countries such as Kenya.

4.5.2.2. Probability of Dying, 1-3 years

We now present results on child mortality, i.e., the probability that a child dies before age 3 conditional on surviving to the first birthday. As noted in section 4.5.1, we had to trade off between having a wide enough interval to have "enough" mortality to give informative results while at the same time not having so wide an interval that we lose a lot of observations because not enough children have reached the upper end of the time interval. As in the case of infant mortality we focus on the current births. As noted in part (ii) of section 4.3 in the discussion on breast-feeding, we treat breast-feeding as a dummy variable to avoid possible spurious correlation between length of breast-feeding and mortality. We also redefine the breast-feeding dummy as 1 if breast-feeding went over 1 year (12 months), 0 otherwise.

Further, those current births that are not yet 3 years old are dropped from the sample, as are those that died before their first birthday. In addition, those current (last) births that may have died after their third birthday enter the equation as having been alive between ages 1-3. The estimated results are presented in Table 4.6. The restrictions that are supported by a Wald test for joint significance are:

Iresid_1=Imomedu_1=agemarry=Iheduc_1=Iheduc_2=bord=bord2=Itoilet_1=Itoilet _2=owncar=Irelig_2=Irelig_3=hsize=0.

This is a standard Wald test. The test statistic is 10.53 compared to a chi-squared critical value with 13 degrees of freedom of 22.36 at the 5% level of significance. We therefore accept the null hypothesis, with the same caveat as before about the size of the test.

Increases in mother's age (**momage** and **momage2**) have the same direction of effect as in infant mortality. A plot of the partial derivative showed again that the positive coefficient on **momage2** overpowers the negative coefficient on **momage** and mortality monotonically increases with mother's age in the relevant age range. Children born in two parent families again tend be healthier than those from single parents as the dummy variable (**dmarried**) shows. The breast-feeding dummy is once again significant, indicating that breast-feeding for a longer period (over 12 months) lowers mortality risk in children. Again this might be the case that mothers breast-feeding for a longer duration are healthier than those breast-feeding for a shorter period of time. Breast-feeding mothers are also less likely to have another child as it extends post-partum amenorrhea. Their children are therefore likely to be more healthy. Note too that breast-feeding also has an indirect effect on fertility. Longer periods of breast-feeding lengthen post partum amenorrhea and consequently the start of ovulation. Therefore this reduces fertility over the relevant time period implying less competition for household resources and better health for the already born children.

Note especially that a number of variables that were significant at the 5% level in the case of infant mortality, such as age at first marriage, **munions**, number of other children under 5, birth order and toilet facilities, are insignificant in the case of child mortality.

4.5.2.3. Estimates Using the 1993 Kenya Demographic Health Survey (KDHS93)

As mentioned in section 4.2, KDHS93 contains most variables covered by KDHS89. For purposes of comparison with the results in subsections 4.5.2.1 and 4.5.2.2 we retain the estimating equations used in that section. We first estimate logit equations for infant mortality (0-12 months) and "child" mortality (1-3 years. Table 4.7 summarises mortality results for the KDHS93 data set. Increases in mother's age **(momage)** and **(momage2)** have the same direction of effect as for KDHS89 (see Tables 4.5 and 4.6). Once more, a plot of partial derivatives with respect to **momage** leads to the conclusion that infant and child mortality monotonically increases with mother's age over the relevant age range. Age at first marriage **(agemarry)** is also positively associated with mortality risk, though insignificant for "child" mortality.

Various other results complement those summarised in Tables 4.5 and 4.6 for KDHS89. Children from two parent families, as captured by the dummy, **dmarried,** have better chances of surviving their first 3 years of life than those in single parent households. Breast-feeding **(dbreast)** and number of other children under 5 **(under5)** have significant negative impacts on infant and "child" mortality. Availability of alternative toilet facilities (pit latrine, bucket or other) improves the chances that the infant will survive to her first birthday. Communicable diseases still play a leading role in child morbidity and mortality in Kenya. Availability of toilet facilities may reduce their spread and burden. Therefore all these alternative forms of toilet facilities play a major role in controlling the spread of such diseases as dysentery, typhoid and cholera in both rural and high density urban areas where there is no running water and therefore no flush toilets. Interestingly, mortality risk between ages 1-3 was lower in the rural than in the urban areas, after controlling for other factors.

Further, as in the case for infant mortality estimates for KDHS89 (Table 4.5), the coefficients on **bord** and **bord2** indicate similar relationship between birth order and child survival. Considering coefficient magnitudes and partial derivatives we find that mortality (both 0-12 months and 1-3 years) decreases with birth order as in 1989. This, as indicated earlier, could be attributed to "learning by doing" on the part of mothers.

Table 4.1

Distribution of Respondents by 5 Year Age Groups

AGE GROUP	KDHS89		KDHS93	
	# of respondents	% of total respondents	# of respondents	% of total respondents
15-19	1,481	20.7	1,788	23.7
20-24	1,402	19.6	1,605	21.3
25-29	1,357	19.0	1,199	15.9
30-34	1,007	14.1	1,112	14.7
35-39	830	11.6	743	9.9
40-44	646	9.0	653	8.7
45-49	427	6.0	440	5.8
Total	7,150	100.0	7,540	100.0
Mean current age	28.20		27.63	

Table 4.2

Distribution of Respondents by Sample Region

REGION	KDHS89		KDHS93	
	# of respondents	% of total respondents	# of respondents	% of total respondents
Nairobi	859	12.0	367	4.9
Central	1,281	17.9	1,075	14.3
Coast	720	10.1	1,091	14.5
Eastern	898	12.6	1,044	13.8
Nyanza	1,265	17.7	1,264	16.8
Rift Valley	1,100	15.4	1,764	23.3
Western	1,027	14.4	945	12.5
Total	**7150**	**100.1**	**7540**	**100.1**

Table 4.3

Distribution of Respondents by Region and Highest Educational Level, 1989.
(Percentages are given in parentheses)

EDUCATION LEVEL	REGION							
	Nairobi	Central	Coast	Eastern	Nyanza	Rift Valley	Western	TOTAL
No education	73 (8.5)	193 (15.1)	376 (52.2)	203 (22.6)	336 (26.6)	298 (27.1)	223 (21.7)	1702 (23.8)
Primary	415 (48.3)	766 (59.8)	240 (33.3)	549 (61.1)	706 (55.8)	575 (52.3)	575 (56.0)	3826 (53.5)
Secondary	356 (41.4)	315 (24.6)	103 (14.3)	145 (16.1)	220 (17.4)	221 (20.1)	227 (22.1)	1587 (22.2)
Higher	13 (1.5)	3 (0.2)	1 (0.1)	0 (0)	2 (0.2)	5 (0.5)	1 (0.1)	25 (0.3)
Not known	2 (0.2)	4 (0.3)	0 (0.0)	1 (0.1)	1 (0.1)	1 (0.1)	1 (0.1)	10 (0.1)
TOTAL	**859**	**1281**	**720**	**898**	**1265**	**1100**	**1027**	**7150**

Table 4.4

Distribution of Respondents by Region and Highest Educational Level, 1993.
(Percentages are given in parentheses)

EDUCATION LEVEL	REGION							
	Nairobi	Central	Coast	Eastern	Nyanza	Rift Valley	Western	TOTAL
No education	28 (7.6)	92 (8.6)	312 (28.6)	170 (16.3)	241 (19.1))	332 (18.9)	122 (12.9)	1297 (17.2)
Primary	163 (44.4)	650 (60.5)	543 (49.8)	653 (62.6)	800 (63.3)	1066 (60.8)	574 (60.7)	4449 (59.0)
Secondary	164 (44.7)	330 (30.7)	233 (21.4)	221 (21.2)	221 (17.5)	345 (19.8)	243 (25.7)	1757 (23.3)
Higher	12 (3.3)	3 (0.3)	3 (0.3)	0 (0.0)	2 (0.2)	11 (0.6)	6 (0.6)	37 (0.5)
TOTAL	**367**	**1075**	**1091**	**1044**	**1264**	**1754**	**945**	**7540**

Table 4.5

Logit Estimates for Probability of Infant Mortality (deaths between 0-12 months) for the Last Child (birth index 1) , KDHS89, n = 2403.

	Unrestricted Model			Restricted Model		
Variable	Odds ratio	Coefficient	Asymptotic t-ratio[a]	Odds ratio	Coefficient	Asymptotic t-ratio[a]
Iresid_1	1.21	0.19	0.72	-	-	-
Momage	0.62	-0.48	-3.90	0.62	-0.48	-3.97
Momage2	1.01	0.01	3.18	1.01	0.01	3.21
Imomedu_1	0.64	-0.44	-1.75	0.66	-0.42	-1.76
Imomedu_2	0.53	-0.62	-1.72	0.56	-0.57	-1.65
Agemarry	1.08	0.08	2.50	1.08	0.08	2.51
Dmarried	0.46	-0.77	-2.31	0.46	-0.78	-2.40
Munion	1.49	0.40	2.45	1.46	0.38	3.85
Munion2	1.00	-0.001	-0.05	-	-	-
Lheduc_1	0.93	-0.07	-0.24	-	-	-
Iheduc_2	1.18	0.17	0.48	-	-	-
under5	0.47	-0.76	-5.20	0.47	-0.75	-5.19
Hsize	1.01	0.01	0.33	-	-	-
Icsex_1	1.21	0.20	1.02	1.20	0.19	0.97
Bord	1.64	0.49	2.72	1.65	0.50	2.80
bordl2	0.98	-0.02	-1.67	0.98	-0.02	-1.71
Dbreastl	0.02	-3.98	-11.59	0.02	-3.95	-11.87
Itoilet_1	0.45	-0.80	-1.54	0.38	-0.96	-2.17
Itoilet_2	0.54	-0.61	-2.29	0.54	-0.61	-2.48
Ireligion_2	1.20	0.18	0.44	-	-	-
Lreligion_3	0.78	-0.25	-0.63	-	-	-
Owncar_1	0.55	-0.60	-1.25	-	-	-
Constant	-	-	-		8.53	5.02

Pseudo R-squared for unrestricted model 0.248 restricted model 0.245

dbreastl - is a dummy variable which equals 1 if the infant was ever breast-fed, 0 otherwise

dmarried - is also a dummy variable which equals 1 for two parent families, 0 otherwise.

Notes to Table 4.5

[a] These are asymptotic t-ratios calculated using Huber-White robust standard errors. The t-ratios are for testing if the odds ratio equals 1 or the coefficient equals 0.

[1] The letter I before a variable indicates it is a dummy indicator as created in STATA program with the xi: interaction expansion command. See STATA statistics software release 5.

[2] See Section 4.3 for variable definitions and Appendix 7 for quick reference.

Table 4.6

Logit Estimates for Probability of Child Mortality (deaths between 1-3 years) for the Last Child (birth index 1) , KDHS89, n = 589.

	Unrestricted Model			Restricted Model		
Variable	Odds ratio	Coefficient	Asymptotic t-ratio[a]	Odds ratio	Coefficient	Asymptotic t-ratio[a]
Iresid_1	1.11	0.11	0.20	-	-	-
Momage	0.50	-0.70	-2.60	0.59	-0.52	-2.96
Momage2	1.01	0.01	1.84	1.01	0.01	2.44
Imomedu_1	1.44	0.36	0.54	-	-	-
Imomedu_2	0.88	-0.13	-0.14	-	-	-
Agemarry	1.03	0.03	0.37	-	-	-
Dmarried	0.23	-1.48	-2.89	0.24	-1.42	-3.05
Munions	0.40	-0.90	-1.44	0.52	-0.66	-0.98
unions2	1.36	0.31	1.93	1.30	0.26	1.58
lheduc_1	0.46	-0.77	-1.32	-	-	-
Iheduc_2	0.83	-0.18	-0.30	-	-	-
under5	0.59	-0.52	-1.71	0.68	-0.38	-1.55
Hsize	1.06	0.06	0.62	-	-	-
Icsex_1	1.20	0.18	0.44	1.23	0.21	0.51
Bord	1.85	0.61	1.65	-	-	-
bord2	0.98	-0.02	-0.90	-	-	-
Dbreast	0.29	-1.24	-2.75	0.29	-1.24	-3.10
Itoilet_1	0.48	-0.73	-0.62	-	-	-
Itoilet_2	0.55	-0.59	-0.70	-	-	-
Ireligion_2	1.59	0.47	0.64	-	-	-
Lreligion_3	0.97	-0.03	-0.04	-	-	-
owncar_1	0.81	-0.21	-0.27	-	-	-
Constant	-	10.95	2.65	-	8.43	2.81

Pseudo R-squared for: unrestricted model 0.236
restricted model 0.186

Notes to Table 4.6

[a] These are asymptotic t-ratios calculated using Huber-White robust standard errors. The t-ratios are for testing if the odds ratio equals 1 or coefficient equals 0.

[1] The letter I before a variable indicates it is a dummy indicator as created in STATA program with xi: interaction expansion command. See STATA statistics software release 5.

[2] See section 4.3 for variable definitions and Appendix 8 for quick reference.

Table 4.7

Logit Estimates for Probability of Infant Mortality (0-12 months) and "Child" Mortality 1-3years) for the Last hild (birth index 1), KDHS93

	0 - 12 months n = 2402 Pseudo R. Sq. = 0.248			1 - 3 years n = 710 Pseudo R. Sq. = 0.236		
Variable	**Odds ratio**	**Coefficient**	**Asymptotic t-ratio[a]**	**Odds ratio**	**Coefficient**	**Asymptotic t-ratio[a]**
Iresid_1	1.21	0.19	0.49	0.26	-1.33	-2.36
Momage	0.65	-0.43	-2.86	0.33	-1.11	-4.05
Momage2	1.00	0.01	2.04	1.01	0.01	3.45
Imomedu_1	1.56	0.44	1.40	0.78	-0.25	-0.47
Imomedu_2	0.84	-0.17	-0.38	0.28	-1.29	-1.47
Agemarry	1.06	0.06	1.54	1.18	0.17	2.81
Dmarried	0.50	-0.70	-2.24	0.25	-1.37	-3.02
Munions	1.37	0.31	1.49	1.05	0.05	0.09
Munions2	1.00	-0.003	-0.15	0.97	-0.03	-0.39
lheduc_1	0.74	-0.30	-0.92	0.43	-0.85	-1.64
Iheduc_2	0.84	-0.17	-0.47	0.70	-0.36	-0.69
under5	0.30	-1.19	-5.94	0.35	-1.05	-1.95
Hsize	0.93	-0.07	-1.12	0.90	-0.10	-0.76
Icsex_1	0.87	-0.14	-0.75	1.54	0.44	1.13
Bordl	1.76	0.56	3.35	6.06	1.80	3.94
bordl2	0.98	-0.02	-1.49	0.88	-0.13	-3.31
Dbreastl	0.01	-4.26	-9.98	0.13	-2.03	-4.80
Itoilet_1	0.83	-0.19	-0.34	0.30	-1.19	-1.36
Itoilet_2	0.57	-0.56	-2.43	0.53	-0.63	-1.19
Ireligion_2	0.70	-0.35	-0.67	0.27	-1.32	-1.47
Lreligion_3	0.91	-0.09	-0.12	0.34	-1.07	-1.13
Constant	-	8.47	3.79	-	13.52	3.44

Notes to Table 4.7

The KDHS93 data set does not have information on car ownership.

[a] These are asymptotic t-ratios calculated using Huber-White robust standard errors. The t-ratios are for testing if the odds ratio equals 1 or coefficient equals 0.

[1] The letter I before a variable indicates it is a dummy indicator as created in STATA program with xi: interaction expansion command. See STATA statistics software release 5.

[2] See section 4.3 for variable definitions and Appendix 8 for quick reference.

Figure 4.1

Map of Kenya Showing Coverage for KDHS89 and KDHS93
(Shaded area represents regions not included in the two surveys)

CHAPTER 5

SUMMARY AND CONCLUSIONS

Chapter 3 focused on aggregate issues relating to child health. We especially focused on estimating the relationship between health, as measured by life expectancy and infant mortality rate and per capita *RGDP*, its distribution and socio-economic indicators. To analyze these relationships we used purchasing power parity data sets for 1970, 1980 and 1990 for over 100 countries. The results showed that the relationship between income and both life expectancy and infant mortality is nonlinear. The relationship is almost linear at low incomes and high incomes.

Income has the greatest influence on these health outcomes where each money unit matters most, i.e. at very low incomes. We therefore draw the conclusion that policies aimed at raising per capita income will have the greatest impact in low income countries where the income-life expectancy/infant mortality relationship is almost linear. Comparing various econometric specifications it is shown that the quadratic spline seems to represent best the relationship between per capita *RGDP* and life expectancy/infant mortality rate. We further show, using a simulation exercise, that redistribution of income in the middle income range was found to lead to an increase in average life expectancy, while income redistribution within low or high income levels had little or no effect. Overall, income distribution does not appear to be important in the regressions.

When we incorporate other factors we find that literacy levels have a great impact in predicting a higher life expectancy and a lower infant mortality. A lower population/doctor ratio is also associated with better health. This is consistent with earlier findings.

Chapter 4 of this thesis focused on the interplay of various child and household characteristics on child survival. We estimated child mortality in the first year and also between ages 1 and 3 using Kenya National Demographic Surveys 1989 and 1993.

In the descriptive analysis, we may conclude that the general educational level for women in Kenya went up between 1989 and 1993. This was in part due to an increase in the percentage of women in higher education categories for the regions that were least educated in 1989 (Coast and Rift Valley).

Using a logit model, we established that children who breast-fed had better chances of survival. We further show that children in a two parent family have lower probability of dying both in infancy and between ages 1 and 3. We also show using the 1989 data set that mother's education is also associated with higher chances of their children surviving infancy. It was apparent, however, that there was no significant difference in the survival of children whose mothers have primary education versus those with secondary or higher education. Moving from no education to basic education was what really mattered in increasing the probability of child survival. In the aggregate analysis, Chapter 3, it was established that literacy is a major determinant of life expectancy at birth and also infant mortality rate.

In Chapter we further show that the effects of mothers' age and birth order on child survival are nonlinear. The results show that overall infant and child mortality increases with mothers' age and declines with birth order.

One conclusion we draw from the study is that policies designed to promote education, especially female education, will lower infant mortality, increase life expectancy at birth and consequently lead to better child survival. This conclusion is based on our findings in both the aggregate and household analysis. We further note that there is need for promotion and encouragement of breast-feeding especially in rural areas where supplemental food is not readily available to most mothers..

Further, we established that income still has a role to play in lowering infant mortality, especially at low and middle income ranges. This calls for programmes aimed at alleviation of absolute poverty in the less developed countries. Amelioration of poverty in LDCs would lead to improved living conditions and better health.

APPENDIX 1

Variable Definitions and Data Sources

AGR	-	% share of agriculture in GDP[a].
ACCESS	-	% population with access to health services, 1987-89[b].
DCS	-	Daily calorie supply (as % of daily requirements)[b].
EXP	-	Life expectancy at birth (Males and females)[a].
IMR	-	Infant mortality rate per 1000 live births[a]. i.e. the number of children dying between ages 0 and 1 per 1000 live births.
FILLIT	-	Female illiteracy[a,c]. Proportion of females over the age of fifteen who cannot, with understanding read and write a short, simple statement on their everyday life. World Bank 93, pp. 308.
POP	-	Total population[a].
RGDPC	-	Real Gross Domestic Product in constant dollars using Chain Index (1985 International prices in PWT 5)[d] i.e. adjusted for purchasing power parity. We refer to it in the rest of the thesis as Per Capita RGDP
POPDOC	-	Population per physician[a,c,g]. (POPDOC90[a,c], POPDOC80[a,g],POPDOC70[a,c])
Per Capita RGDP	-	Refers to Per Capita Real Gross Domestic Product. See RGDPC above.

TILLIT - Total illiteracy[a,c,e,f]. Defined as in FILLIT above but for total population. (TILLIT90[a,c], TILLIT80[a,e],TILLIT70[a,f])

SHARETOP10- % share of income of highest 10 %[c].

SHARELOW20- % share of income of lowest quintile (20%)[c].

URB - Urban population as a percentage of total population[c].

WATER - % of total population with access to safe water[b].

(Figures for Singapore and Guinea are from source a). Note that safe water refers to safe water supply which include treated surface waters or untreated but uncontaminated water such as from springs, sanitary wells and protected boreholes (HDR92).

YLOW20 - Absolute average income of the lowest quintile. This is calculated as follows: *YLOW20 = [(SHARELOW20/100)*RGDPCH]*5*

Z - RGDPC spline term defined for a given country i in year t as:

$RGDPC_{it}$- mean $RGDPC_t$ if $RGDPC_{it}$ > mean $RGDPC_t$

0 otherwise.

[a]- Social indicators of development 1993. Version 2.5. April 1992. IBRD/The World Bank.

[b]- Human Development Report 1992 (HDR92), UNDP.p.130.

[c]- World Development Report, 1993. (World Bank 93). Investing in Health

[d]- Penn World Tables, Mark 5.6 (PWT56). Available through web site: http://datacentre.chass.utoronto.ca/pwt/index.html. We use series RGDPC for 1970, 1980 and 1990.1

For a detailed description of the International comparisons project and Penn World Tables see Summers and Heston (1991).

[e]- World Development Report, 1983. World Economic Recessions and Prospects for Recovery. Management in Development. World Economic Indicators

[f]- Penn World Tables 1976.

[g]- World Development Report 1984/85. International Capital and Economic Development: World Economic Indicators.

APPENDIX 2

Purchasing Power Parity (PPP) Adjusted Per Capita RGDP In 1985 Prices (US $)

COUNTRY	WBCOY[1]	1970	1980	1990
Algeria	DZA	1826	2758	2777
Argentina	ARG	5637	6506	4706
Australia	AUS	10756	12520	14445
Austria	AUT	7510	10509	12695
Bangladesh	BGD	1280	1085	1390
Belgium	BEL	8331	11109	13232
Benin	BEN	1118	1114	920
Bolivia	BOL	1661	1989	1658
Brazil	BRA	2434	4303	4042
Burkina Faso	HVO	374	457	511
Burundi	BDI	480	550	641
Cameroon	CMR	804	1194	1226
Canada	CAN	10124	14133	17173
Cape Verde	CPV	634	934	1058
Central African Republic	CAF	747	706	579
Chad	TCD	660	528	399
Chile	CHL	3605	3892	4338
China	CHN	696	972	1324
Colombia	COL	2140	2946	3300
Comoros	COM	693	631	564
Congo	COG	1670	1931	2211
Costa Rica	CRI	2904	3717	3499
Cote D'Ivoire	CIV	1615	1790	1213
Cyprus	CYP	3753	5295	8368
Czechoslavakia	CSK	2520	3731	4095

COUNTRY	WBCOY[1]	1970	1980	1990
Denmark	DNK	9670	11342	13909
Dominican Republic	DOM	1536	2343	2166
Ecuador	ECU	1789	3238	2755
Egypt	EGY	1163	1645	1912
El Salvador	SLV	1810	2014	1824
Fiji	FJI	2592	3609	4007
Finland	FIN	8108	10851	14059
France	FRA	9200	11756	13904
Gabon	GAB	3704	4797	3958
Gambia	GMB	722	1017	799
Germany	DEU	9425	11920	14341
Ghana	GHA	1059	976	902
Great Britain	GBR	8537	10167	13217
Greece	GRC	4224	5901	6768
Guatemala	GTM	2028	2574	2127
Guinea	GIN	467	817	767
Guinea Bissau	GNB	699	471	689
Guyana	GUY	1816	1927	1094
Honduras	HND	1237	1519	1377
Hong Kong	HKG	4502	8719	14849
Hungary	HUN	3358	4992	5357
Iceland	ISL	6772	11566	13362
India	IND	802	882	1264
Indonesia	IDN	715	1281	1974
Iran	IRN	4796	3434	3392
Ireland	IRL	5015	6823	9274
Israel	ISR	6004	7895	9268
Italy	ITA	7568	10323	12488
Jamaica	JAM	2645	2362	2545

COUNTRY	WBCOY[1]	1970	1980	1990
Japan	JPN	7307	10072	14331
Kenya	KEN	586	911	911
Korea	KOR	1680	3093	6673
Luxembourg	LUX	9782	11893	16280
Madagascar	MDG	1146	984	675
Malawi	MWI	440	554	519
Malaysia	MYS	2154	3799	5124
Mali	MLI	419	532	531
Mauritania	MRT	872	885	791
Mauritius	MUS	2398	3988	5838
Mexico	MEX	3987	6054	5827
Morocco	MAR	1342	1941	2151
Mozambique	MOZ	1497	923	760
Namibia	NAM	2642	2904	2854
Netherlands	NLD	9199	11284	13029
New Zealand	NZL	9392	10362	11513
Nicaragua	NIC	2359	1853	1294
Nigeria	NGA	767	1438	995
Norway	NOR	8034	12141	14902
Pakistan	PAK	1029	1110	1394
Panama	PAN	2584	3392	2888
Papua New Guinea	PNG	1896	1779	1425
Paraguay	PRY	1394	2534	2128
Peru	PER	2736	2875	2188
Philippines	PHL	1403	1879	1763
Poland	POL	2941	4419	3820
Portugal	PRT	3306	4982	7478
Rwanda	RWA	647	757	756

COUNTRY	WBCOY[1]	1970	1980	1990
Senegal	SEN	1146	1134	1145
Sierra Leone	SLE	1435	1139	901
Singapore	SGP	3017	7053	11710
South Africa	ZAF	3254	3496	3248
Spain	ESP	5861	7390	9583
Sri Lanka	LKA	1243	1635	2096
Sudan	SDN	817	866	757
Sweden	SWE	10766	12456	14762
Switzerland	CHE	12942	14301	16505
Syria	SYR	2294	4467	3897
Thailand	THA	1526	2178	3580
Togo	TGO	618	731	641
Trinidad and Tobago	TTO	6795	11262	7764
Tunisia	TUN	1442	2527	2910
Turkey	TUR	2202	2874	3741
Uganda	UGA	647	534	554
United States	USA	12963	15295	18054
Uruguay	URY	4121	5091	4602
Venezuela	VEN	7753	7401	6055
Yugoslavia	YUG	3297	5565	4548
Zambia	ZMB	1117	971	689
Zimbabwe	ZWE	1082	1206	1182

[1] WBCOY are the standard World Bank country acronyms that we use in the figures (for example Figure 1).

APPENDIX 3

Purchasing power parity (PPP) income estimates are used to give more meaningful estimates of real income and therefore of the relationship between real income and health. Kravis, Heston and Summers (1978a) observed that the conversion of GDP of various countries to U.S. dollars using official exchange rates underestimates up to three times the real GDP of those countries.

The United Nations' International Comparison Project (ICP) was initiated to compare purchasing power of currencies and their real GDPs based on Purchasing Power Parity (PPP). PPP is defined as the number of units of a country's currency required to buy the same amount of goods and services in the domestic market as one dollar would buy in the U.S. (World Bank, 1989).

If all goods and services were traded and exchange rates between currencies were in equilibrium, purchasing power would be the same in each country. This is not, however, true in the real world. With ICP estimates a country's GDP is thus calculated based on what goods and services a country's currency will buy, compared with the purchasing power of other currencies. The origins of the PPP estimates can be traced back to two studies by Gilbert and Kravis (1954) and Gilbert and Associates (1958) in which they attempted to construct PPP for the OECD member countries. After this period the ICP project was formed at the University of Pennsylvania to provide estimates for the United Nations statistical office and was funded by the World Bank with support from the Ford Foundation and the United States Agency for International Development (USAID).

Phase I was conducted in 1970 and covered 10 countries, namely Columbia, France, Federal Republic of Germany, Hungary, India, Italy, Japan, Kenya, United Kingdom, and United States of America (Kravis, Kenessey, Heston and Summers, 1975). Phase II was conducted in 1973 and covered the 10 countries from Phase I plus six new ones, namely Belgium, Iran, Republic of Korea, Malaysia, Netherlands and Philippines (Kravis, Heston and Summers, 1978b).

Phase III, conducted in 1975, covered a total of 34 countries. These were the following: Austria, Belgium, Brazil, Colombia, Denmark, France, Federal Republic of Germany, Hungary, India, Iran, Ireland, Italy, Jamaica, Japan, Kenya, Korea, Luxembourg, Malawi, Malaysia, Mexico, Netherlands, Pakistan, Philippines, Poland, Romania, Spain, Sri Lanka, Syria, Thailand, United Kingdom, United States, Uruguay, Yugoslavia, and Zambia. (Kravis et al. 1982).

Based on the estimated relations between real gross domestic product and nominal GDP and the degree of openness of the economy measured as the average over a specified period of the ratio of exports plus imports to GDP, Summers and Heston (1984) estimated real GDP for 124 countries for the period 1950-1980.

Summers and Heston's estimates are based on three concepts: real GDP (adjusted at 1975 international prices), real GDP at current international prices and real GDP at constant 1975 international prices adjusted for changes in terms of trade. The 124 countries comprise 94% of the world's population. Notable omissions were: Kuwait, Libya, Saudi Arabia, due to large net foreign balances leading to unreliable estimates from cross sectional data, and Cuba and Lebanon, which were excluded due to insufficient information. Of the 124 countries, over half have data for 1950-1980 while the rest have data only for 1960-1980. Phase IV of the project was done in 1980.

Phase V was done in 1985, and covered 64 countries. The results are published in the World Bank (1989) Table 30, Column 1. Note also that starting with the 1989 report, the World Bank annually publishes real GDP estimates based on ICP estimates.

Phase VI was conducted in 1990. World Bank (1993) gives indexed figures for 1987 based on PPP. For the OECD member countries the results are based on ICP Phase VI figures for 1990 extrapolated backwards to 1987. For the non-OECD countries the results are based on Phase V figures for 1985 extrapolated forward to 1987.

The PPP project is ongoing and the latest data is available from National Bureau of Economics and Research (NBER) at http://datacentre.chass.utoronto.ca/pwt/index.html. We used the latest that was available, known as Penn World Tables 5.6 (PWT56).

APPENDIX 4.

Ahluwahlia (1976) did some exploratory work on the relationship between income relative inequality and the level of development as measured by per capita GDP. The study used cross-sectional data for 60 countries which comprised 40 LDCs, 14 MDCs and 6 socialist countries. Income shares of different percentile groups were regressed on real per capita GNP and a number of development indicators to control for the stage of development. He estimated the model for all the countries and also for a reduced sample of LDCs.

Using a quadratic formulation there was strong evidence in support of the Kuznets U-shaped hypothesis on the relationship between inequality and level of development. Kuznets (1955,1963) observed that inequality rose and then fell during the period of modern economic growth.

Ahluwahlia estimated the relationship between income shares of different percentile groups and the natural log of real per capita GNP (LNGNP) and its quadratic term ($LNGNP^2$). The coefficient of $LNGNP^2$ was negative and significant for both models involving the entire sample and for the reduced sample (LDCs only).

Ahluwahlia also attempted to test for the absolute impoverishment hypothesis. Although there was evidence of increasing relative inequality at early stages of development, absolute incomes of the poor improved. There was no evidence to suggest that faster development in the short run led to higher inequality. The data tended to suggest that faster growth led to a faster transition through the different levels of inequality.

APPENDIX 5

SECTION NAME	CODE	CLASS	ROWS	LENGTH	SECTION LABEL
REC01	01	S	1	66	Respondent's basic data
REC11	11	S	1	67	Respondent's basic data
REC21	21	M	20	47	Reproduction
REC22	22	S	1	62	Reproduction
REC31	31	S	1	155	Contraception table
REC32	32	S	1	121	Contraceptive knowledge and use
REC33	33	M	7	34	Contraceptive practice
REC41	41	M	6	33	Maternity
REC42	42	S	1	42	Health and breast-feeding
REC43	43	M	6	104	Health
REC44	44	M	6	67	Height and weight
REC51	51	S	1	61	Marriage/exposure
REC61	61	S	1	38	Fertility preferences
REC71	71	S	1	37	Husband's characteristics
REC81	81	S	1	40	Characteristics of interview
REC91	91	S	1	93	Various single variables
REC94	94	M	6	26	Maternity history variables
REC95	95	M	6	57	Specific history variables
REC97	97	S	1	106	Husband's questionnaire cover page and husband's background
REC98	98	S	1	80	Husband's contraceptive table
REC99	99	S	1	77	Husband's contraceptive and fertility preferences

SPSS users: RECORDS = 71
STATA users:_ lines(71)
S - Single record-e.g. demographic and socio-economic information on respondent
M - Multiple record e.g. Birth history
Length - Maximum length for each record/Number of columns.

APPENDIX 6

Urban-Rural Differentials In Asset Ownership, KDHS89: Percentages Out Of Total Respondents In Each Group.

		Urban	Rural	Total
Refrigerator	Do not own	83.4	99.0	94.8
	Own	16.4	1.0	5.1
Car	Do not own	81.7	95.5	91.8
	Own	18.1	4.4	8.1
Tractor	Do not own	97.8	98.0	97.9
	Own	1.8	1.5	1.5
Television	Do not own	77.4	98.6	92.9
	Own	22.4	1.4	7.0
Radio	Do not own	22.3	43.1	37.6
	Own	77.6	56.8	85.2
Electricity	Do not access	54.7	97.3	85.9
	Have access	45.2	2.6	14.0

APPENDIX 7

Alphabetical Listing Of Variable Definitions For KDHS Data Analysis

^{c} implies categorical variable

agemarry - age at first marriage

breast - length of breast-feeding in months.

bord - birth order. Since we focus on current child, this variable therefore indicates how many births the respondent had had at time of survey

csex^c - sex of child.

heduc^c - husband's highest education level

hsize - number of household members aged 5 or more years

momage - mother's current age

momedu^c - mother's highest education level

mstatus^c - mother's marital status

munions - number of other wives that the husband has

under5 - number of other children in the household under the age of 5

toilet^c - type of toilet facility

religion^c - religion of mother

Dummy indicators for categorical variables

csex:	**Icsex_0**	Male
	Icsex_1	Female
dbreast:	breast-feeding dummy. For estimation of deaths between 0-12 months this dummy equals 1 if child ever breast-fed, 0 otherwise. For deaths 0<A<B **dbreast** equals 1 if child breast-fed beyond A, 0 otherwise.	
dmarried:	married dummy which equals 1 for a two parent family 0, otherwise	
heduc:	**Iheduc_0**	None
	Iheduc_1	Primary
	Iheduc_2	Secondary or higher
momedu:	**Imomedu_0**	None
	Imomedu_1	Primary
	Imomedu_2	Secondary or higher
mstatus:	**Imstatus_0**	Single (Never married, widowed, divorced, not living together)
	Imstatus_1	Married and living together
resid:	place of residence	
	Iresid_0	urban
	Iresid_1	rural
toilet:	**Itoilet_0**	None
	Itoilet_1	flush
	Itoilet_2	Other (pit latrine, bucket and other)

religion:	**Ireligion_1**	Catholics and Protestants
	Ireligion_2	Muslims
	Ireligion_3	Other

owncar: whether household owns a car or not. It equals 1 if they own a car, 0 otherwise

Note: In estimation STATA always drops the smallest value of the dummy variable.

This therefore becomes the base or referral dummy indicator.

APPENDIX 8

Mean And Standard Deviations Of Major Variables In Child Survival Used In Analysis KDHS89 & KDHS93.

Mother's Characteristics	1989		1993	
	Mean	Standard Deviation	Mean	Standard Deviation
Current age	28.20	9.08	27.63	9.23
Education[c]	0.99	0.68	1.07	0.64
Children ever born	3.52	3.20	3.17	3.15
Age at first marriage	17.67	3.76	18.19	3.58
Marital Status[c]	0.98	0.95	0.98	1.07
No. of Marital[c] unions	0.25	0.69	0.16	0.60
Husband's Education	1.20	0.68	1.22	0.65
Other Variables				
Average length of breast- feeding (months)[1]	17.63	7.61	19.52	7.94
Children under 5	0.85	1.34	0.59	0.88
Household size	5.84	3.70	5.39	3.03
Preceding Birth Interval (months)[1]	35.86	21.75	36.03	21.44
Birth Order	4.57	2.92	4.41	2.89

[c] Categorical Variable

[1] These figures are for the last births that had already been weaned whether alive or dead at time of interview.

REFERENCES

Ahluwahlia, M.S. (1976) "Inequality, Poverty and Development," *Journal of Development Economics*, (3), December, pp. 307-342.

Allison, P.D.(1984) Event History Analysis. Regression for Longitudinal Event Data. Sage Publications, Beverly Hills.

Anand, S. and M. Ravallion (1993) "Human Development in Poor Countries: On the Role of Private Incomes and Public Services," *Journal of Economic Perspectives*, (7:1), pp. 133-150.

Barrera, A. (1990) "The role of Maternal schooling and its Interaction with Public Health Programs in Child Health Production," *Journal of Development Economics,* (32:1), pp. 69-91.

Behrman, J.R. and B.L. Wolfe (1987) "How Does Mother's Schooling Affect Family Health, Nutrition, Medical Care Usage, and Household Sanitation," *Journal of Econometrics,* (36:1/2),September/October, pp. 185-204.

Behrman, J.R. and A.B. Deolalikar (1988) "Health and Nutrition," in Handbook of Development Economics,(1), H. Chenery and T.H. Srinivasan,eds., Amsterdam, North-Holland, pp. 631-711.

Caldwell, J.C. (1993) "Health Transition: The Cultural, Social and Behavioural Determinants of Health in the Third World," *Social Science and Medicine*, (32:2), January, pp. 125-135.

CIAR (1991) "The Determinants of Health" Canadian Institute for Advanced Research, Working Paper No.5. Canadian Institute for Advanced Research,Toronto.

Edwards, L.N. and M. Grossman (1978) "Children's Health and the Family," National Bureau of Economic Research, Working Paper No. 256, New York.

ENHR (1991) Essential National Health Research: A Strategy for Action in Health and Human Development. Task Force on Health Research for Development.

Feeny, D. (Unpublished) Notes on a Simple Household Model. Lecture Notes, MS 722, Session 9, March 8, 1993.

Gilbert, M. and Kravis, I.B. (1954) An International Comparison of National Producers and the Purchasing Power of Currencies: A Study of the United States, the United Kingdom, France, Germany, and Italy. Organization for European Economic Cooperation, Paris.

Gilbert, M. and Associates (1958) Comparative National Product and Price Levels. Organization for European Economic Cooperation, Paris.

Greene, W.H. (1993) Econometric Analysis. MacMillian Publishing Company,Paris, 2nd Edition.

Kennedy, P. (1992) A Guide to Econometrics. The MIT Press, Cambridge, Massachusetts, 3rd Edition.

Kravis, I.B., Z. Kennessy, A. Heston and R. Summers (1975) A System of Internatioinal Comparisons of Gross Product and Purchasing Power. John Hopkins, Baltimore.

Kravis, I.B., A. Heston, and R. Summers (1978a) International Comparisons and Real Product and Purchasing Power. Johns Hopkins, Baltimore.

Kravis, I.B., A. Heston, and R. Summers (1978b) "Real GDP Per Capita for More Than One Hundred Countries," *Economic Journal*, (88), pp. 215-242.

Kravis, I.B, A. Heston and R. Summers (1982) A System of International Comparisons of Gross Product and Purchasing Power; Phase II. John Hopkins University Press, Baltimore.

Kuznets, S. (1955) "Economic Growth and Income Inequality," *American Economic Review,* (49), March 1955.

Kuznets, S. (1963) "Quantitative Aspects of the Economic Growth of Nations: VIII Distribution of Income by Size," *Economic Development and Cultural Change*, (11:2), pt.2.

Lindert, P.H. (1986) "Comment," in Stanley L. Engerman and Robert E. Gallman, eds., Long-Term Factors in American Economic Growth, National Bureau of Economic Research Conference on Research in Income and Wealth, *Studies in Income and Wealth*, (51), University of Chicago Press, Chicago, pp. 527-555.

Preston, S.H. (1975) "The Changing Relation Mortality and Level of Economic Development," *Population Studies*, (29:2), pp. 231-248.

Pritchett, L and L.H. Summers (1996) "Wealthier is Healthier," *Journal of Human Resources*, (31:4), Fall, pp. 841-868.

Ram, R. (1992) "Inter-country Inequalities in Income and Basic Needs Indicators: A Recent Perspective," *World Development*, (20:6), pp. 899-905.

Rosenzweig, M.R. and J.P. Schultz (1983) "Estimating a Household Production Function: Heterogeneity, the Demand for Health Inputs, and Their Effects on Birth Weight," *Journal of Political Economy*, (91:5), pp. 723-746.

Sen, A. (1981) "Public Action and the Quality of Life in Developing Countries," *Oxford Bulletin of Economics and Statistics*, (43:4), November, pp. 287-318.

STATA Statistical Software, Release 5. Statistics, Data Management and Graphics. Stata Press, College Station, Texas, 1997.

Summers, R. and A. Heston (1984) "Improved International Comparison of Real Product and its Composition: 1950-1980," *Review of Income and Wealth*, (30:2), June, pp. 207-262.

Summers, R. and A. Heston (1991) "The Penn World Table (Mark 5): An Expanded Set of International Comparisons, 1950-1988," *Quarterly Journal of Economics*, (106:2), pp. 327-368.

United Nations Development Report (1992) Human Development Report,1992. Oxford University Press, New York.

Waldmann, R.J. (1992) "Income Distribution and Infant Mortality," *Quarterly Journal of Economics*, (107:4), pp. 1283-1302.

Wennemo, I. (1993) "Infant Mortality, Public Policy and Inequality. A Comparison of 18 Industrialized Countries, 1950-1985," *Sociology of Health and Illness*, (15), pp. 429-446.

Wheeler, D. (1980) " Basic Needs Fulfillment and Economic Growth. A Simultaneous Model," *Journal of Development Economics*, (7), pp. 435-451.

Wilkinson, R.G. (1992) "Income Distribution and Life Expectancy," *British Medical Journal*, (304), pp. 165-168.

Wolfe, B.L. and J.R. Behrman (1982) "Determinants of Child Mortality, Health, and Nutrition in a Developing Country," *Journal of Health Economics*, (11), pp. 163-194.

World Bank (1976) World Tables, 1976. World Bank, Washington DC.

World Bank (1983) World Development Report, 1983. World Bank, Washington DC.

World Bank (1985) World Development Report, 1984/85. World Bank, Washington DC.

World Bank (1989) World Development Report, 1989. World Bank, Washington DC.

World Bank (1990) World Development Report, 1990. World Bank, Washington DC.

World Bank (1993) World Development Report, 1993. World Bank, Washington DC.

World Bank (1999) World Development Report, 1998/99. World Bank, Washington DC.